Essays in Financial Economics

Simon Mayer

Erasmus University Rotterdam and Tinbergen Institute

ISBN: 978 90 361 0659 7

Cover design: Crasborn Graphic Designers bno, Valkenburg a.d. Geul

This book is no. 784 of the Tinbergen Institute Research Series, established through cooperation between Rozenberg Publishers and the Tinbergen Institute. A list of books which already appeared in the series can be found in the back.

Essays in Financial Economics

Essays over Financiële Economie

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the
rector magnificus

Prof.dr. F.A. van der Duijn Schouten

and in accordance with the decision of the Doctorate Board.
The public defence shall be held on

Friday 3 September 2021at 13.00hrs
by

Simon Richard Mayer
born in Augsburg, Germany.

Erasmus University Rotterdam

Erasmus

Doctoral Committee:
All names with the initials and full title without stating the university

Promotors:	prof.dr. S Gryglewicz prof.dr. P Verwijmeren
Other members:	prof.dr. PAE Koudijs prof.dr. JTJ Smit dr. R. Westermann

Acknowledgements

First of all, I would like to thank my advisor Sebastian Gryglewicz for his continued guidance and support. Sebastian supported and guided me during my time at Erasmus University but also while I was abroad. He was always available in difficult times and always ready to advise and help. I also would like to thank Sebastian for teaching me crucial things in the research process and for co-authoring with me on many exciting projects. Without Sebastian, I would never have reached any of my achievements and I am really grateful for all his support.

I would also like to thank Sebastian for introducing me to Erwan Morellec. I would like to thank Erwan for his support, for his guidance, for co-authoring with me, and for hosting me at EPFL. I really enjoyed my time in Lausanne at EPFL and learnt a lot during my stay.

Next, I would like to thank Barney Hartman-Glaser for mentoring me, for supporting me, and for working with me on joint projects. I also would like to thank Barney for hosting me at UCLA. I had an incredible time at UCLA that was very productive.

I am also very grateful to Konstantin Milbradt for his continued guidance, support, and co-authorship. I would like to thank Konstantin also for hosting me at Northwestern. I really enjoyed my time at Northwestern and it was a great preparation for my job market the year after, which — I believe — helped me a lot.

I am also grateful to my colleagues at Erasmus University and my fellow PhD students at Tinbergen Institute. Especially, I would like to thank Patrick Verwijmeren who advised and supported me during my PhD. Thanks also to my fellow PhD students Ioana, Joris, Tim, Simon, and Leonard for the times at Tinbergen Institute and afterwards.

I am very grateful for the support from my friends, family, and my girlfriend Mona. My girlfriend Mona has supported me and has shared all the difficulties with me, especially during job market and application process. It has been a difficult time for us both, and I am very grateful to her for supporting me during this time even though it was tough for her and required lots of work and effort. I also would like to thank my parents, my father Richard and my mother Rita. My parents have always supported me in difficult times and without them the PhD would not have been possible. Likewise, I am grateful to my brother Stefan for his support, guidance, and helping me out. Without support of family, I would not have been here and could not have made it to this point. Last, I would like to thank all friends for sharing good times with me and supporting me during the process. I would like to thank my friends from Germany, Marvin, Tobias, Johannes, Marius, Raffael, and Florian. I would like to thank Simon, Alex, Kemal, Razvan, Jamel, Josh, and Ben for the incredible time in Amsterdam (which I will miss a lot).

Contents

Chapter 1

Introduction

Typically, a firm's insiders, such as high level executives and CEOs, have different incentives or information than a firm's outside stakeholders, such as outside investors. Crucially, this leads to agency conflicts, moral hazard, and information asymmetries and frictions between a firm's insiders and outsiders. These agency conflicts or information asymmetries then affect a firm's financing and investment decisions as well as corporate policies. A large part of the corporate finance literature studies these agency conflicts, their impact on corporate policies, and how to resolve agency conflicts. In this thesis, I present five papers that contribute to this literature in corporate finance. While each of these papers studies a different economic problem, there is one common theme among all of them, namely, agency conflicts and moral hazard.

In Chapter 2 titled "Financing Breakthroughs under Failure Risk," I study agency conflicts inherent to the financing of young and innovative firms and startups. This essay is motivated by the example of Theranos, a US-based startup developing a novel blood testing technology. After its founding in 2003, the US startup *Theranos* raised funds from venture capitalists and private investors, building on the promise of a novel method of blood testing. This resulted in a 10 billion dollar valuation at its peak in 2014. Between 2014 and 2018, the blood test technology developed by Theranos failed. However, instead of disclosing the technology's failure, the company issued false statements regarding the project's progress and continued to raise funds, until the pyramid of lies eventually collapsed in 2018. The example of Theranos highlights key difficulties inherent to financing innovative projects. Innovative projects typically i) exhibit substantial failure risk, ii) require a high level of expertise from insiders developing the project, and iii) require capital from investors with limited expertise. To resolve agency conflicts between the insiders, developing the project, and investors, financing the project, the provision of financing must be contingent on project outcomes. However, when it is hard for investors to track project progress, insiders can hide bad outcomes, such as project failure. In this paper, I study dynamic contracting in light of this tension and characterizes the optimal incentive provision for truthful disclosure of bad outcomes. To do so, I develop a dynamic principal-agent model, in which the principal financing the project cannot observe project failure and the agent developing the project, can hide or misreport failure. Time-decreasing rewards for failure and high rewards for success provide incentives for disclosure of failure. As there is a tension between incentives for disclosure of failure and project development, the optimal contract does not always incentivize disclosure of failure and consists of two stages. The provision of financing is unconditional in the first stage and performance-sensitive in the second stage. The paper also discusses implications for venture capital financing, R&D financing, and executive compensation contracts.

Chapter 3 titled "Agency Conflicts and Short- vs. Long-Termism in Corporate policies" (joint work with Sebastian Gryglewicz and Erwan Morellec) analyzes how agency conflicts within a more general types of firms affect the time horizon of corporate policies, potentially leading to short-termism and long-termism in corporate policies. The question of the optimal horizon of corporate policies has received considerable attention in recent years, with much of the discussion focusing on whether short-termism destroys value. At the same time, big tech companies — such as Amazon, Tesla, or Uber — are known for their excessive focus on the long term, i.e., their long-termism. We build a dynamic agency model in which the agent controls both current earnings via short-term investment and firm growth via long-term investment. Under the optimal contract, agency conflicts can induce short- and long-term investment levels beyond first best, leading to short- or long-termism in corporate policies. The paper analytically shows how firm characteristics shape the optimal contract and the horizon of corporate policies, thereby generating a number of novel empirical predictions on the optimality of short- versus long-termism. We show that depending of firm characteristics, agency conflicts may lead to short-termism (i.e., excessive short-term investment) or long-termism (i.e., excessive long-term investment). We also demonstrate that combining short- and long-term agency conflicts naturally leads to asymmetric pay-for-performance in managerial contracts.

In Chapter 4 "Delegated Monitoring and Contracting" (joint work with Sebastian Gryglewicz), we study optimal dynamic delegated contracting and monitoring. Financial intermediaries—such as private equity funds, hedge funds, and banks—play a crucial role in firm-level governance by both monitoring management's activities and directly influencing managerial compensation contracts. While these intermediaries may possess unique capacities to contain agency conflicts within firms, they are subject to agency frictions of their own. Thus, intermediation—specifically, delegated monitoring—is a double agency problem. We then analyze the joint presence of agency conflicts at the firm and intermediary levels and their effects on the design of incentive contracts. In a dynamic agency model, investors finance a firm run by an agent while delegating monitoring and contracting with the agent to an intermediary. The intermediary passes through part of its incentives to the agent. After good performance, the agent's incentives create an agency overhang problem, that is, the reduced incentives of the intermediary for monitoring because its benefit accrues mostly to the agent. Conversely, after poor performance, the agent's incentives generate incentives for monitoring to avoid further distress. Investors can benefit from directly contracting with the agent to prevent propagating agency conflicts between the agent and the intermediary.

In Chapter 5 titled "Optimal Financing with Tokens" (joint work with Sebastian Gryglewicz and Erwan Morellec), we study optimal design and financing with cryptocurrencies and tokens. Firms and their business models are changing continuously and there is frequent disruption by technological breakthroughs, such as the internet or more recently blockchain technology and cryptocurrencies. Many theories in corporate finance have been developed before these disruptions and thus may not be able to characterize optimal corporate decision making for firms with innovative and novel business models relying on novel technologies. Technological disruptions are reshaping the financial sector (Stulz (2019)), where FinTech has become a new paradigm. New asset classes and financial instruments like cryptocurrencies and digital tokens have gained tremendous popularity, with the market capitalization of digital currencies exceeding one trillion US dollars at the time of writing this thesis. In addition, new financial instruments also create new possibilities for firms to seek financing. Notably, initial coin offerings (ICOs) have become an important source of financing for firms that develop digital platforms (Howell, Niessner, and Yermack, 2020). By the end of 2018, over 5500 firms had attempted to raise funds using an ICO, raising over 30 billion dollars (Lyandres, Palazzo, and Rabetti, 2019) and with at least 20 ICOs taking in more than 100 million dollars (Howell et al., 2020). In an ICO, a firm raises funds by issuing cryptographically secured tokens. Because these tokens serve as the means of payment on a platform or offer access to the firm's services, they possess *utility features* and are therefore often called *utility tokens*. Despite the popularity of ICOs and the considerable growth of the academic literature on this new form of financing, a number of key questions remain open. Chief among these is whether an ICO should be preferred to alternative ways of financing, such as financing with equity or with tokens other than utility tokens. In Chapter 5, we develop a model in which a startup firm issues tokens to finance a digital platform, which creates agency conflicts between platform developers and outsiders. We show that token financing is preferred to equity financing, unless the platform expects strong cash flows, has large financing needs, or faces severe agency conflicts. Tokens are characterized by their utility features, facilitating transactions, and security features, granting cash flow rights. While security features trigger endogenous network effects and spur platform adoption, they also dilute developers' equity stake and incentives so that the optimal level of security features decreases with agency conflicts and financing needs. We also show under what circumstances, different types of token offerings, such as an ICO or a security token offering (STO), is the preferred way of financing for a firm launching a digital platform.

In Chapter 6 "Managing Stablecoins" (joint work with Ye Li), we develop a dynamic model of the issuance and design of stablecoins. A major problem of cryptocurrencies and crypto-tokens is that their inherently high price volatility invalidates their function as transaction medium or store of value. Stablecoins are supposed to mitigate this problem. Stablecoins are crypto-tokens that are designed to have a stable price relative to some assets or fiat currencies. In 2020, the number of stablecoin projects has risen sharply and, by now, stablecoins make up over 5% of the total cryptocurrency market capitalization. Typically, stablecoins are issued by private entities that ensure price stability, e.g., by managing stablecoin supply or by facilitating the convertibility of stablecoins to another asset. However, the issuer of stablecoins may have different interests than the holders of stablecoins. In addition, at the time of writing this statement, there is no systematic regulation of stablecoin issuance to reduce these conflicts of interest. We then develop a dynamic model of the issuance and design of stablecoins. In a dynamic model of stablecoins, we show that even with over-collateralization, a pledge of one-to-one convertibility to a reference currency is not sustainable in a stochastic environment. The distribution of states is bimodal – a fixed exchange rate can persist, but debasement happens with a positive probability and recovery is slow. When negative shocks drain the reserves that back stablecoins, debasement allows the issuer to share risk with users. Collateral requirements cannot eliminate debasement, because risk sharing is ex-post efficient under any threat of costly liquidation, whether it is due to reserve depletion or violation of regulation. Optimal stablecoin management requires a combination of strategies commonly observed in practice, such as open market operations, transaction fees or subsidies, re-pegging, and issuance and repurchase of "secondary units" that function as stablecoin issuers' equity. The implementation varies with user-network effects and is guided by Tobin's q of transaction data as productive capital.

Chapter 7 concludes the thesis with a brief summary and gives suggestions for further research.

Chapter 2

Financing Breakthroughs under Failure Risk[1]

2.1 Introduction

After its founding in 2003, the US startup *Theranos* raised funds from venture capitalists and private investors, building on the promise of a novel method of blood testing. This resulted in a 10 billion dollar valuation at its peak in 2014. Between 2014 and 2018, the blood test technology developed by Theranos failed. However, instead of disclosing the technology's failure, the company issued false statements regarding the project's progress and continued to raise funds, until the pyramid of lies eventually collapsed in 2018.

The example of Theranos highlights key difficulties inherent to financing innovative projects. Innovative projects typically i) exhibit substantial failure risk, ii) require a high level of expertise from insiders developing the project, and iii) require capital from investors with limited expertise. To resolve agency conflicts between the insiders, developing the project, and investors, financing the project, the provision of financing must be contingent on project outcomes. However, when it is hard for investors to track project progress, insiders can hide bad outcomes, such as project failure. This paper studies dynamic contracting in light of this tension and characterizes the optimal incentive provision for truthful disclosure of bad outcomes.

In the model, a principal finances a project developed by an agent with limited liability. Project development requires funds from the principal and, absent frictions, it is efficient to finance the project until completion. The timing of completion is uncertain and for simplicity not affected by the agent. Completion results in either success or failure, whereby the agent's hidden effort during project development determines the likelihood of success. Moral hazard arises because the agent derives private benefits from shirking. Incentives for effort are provided by paying the agent more for success than for failure. If both potential project outcomes, success and failure, are publicly observable and contractible, the principal pays the agent only for success and finances the project until completion.

This paper studies incentive provision when it is hard for the principal to observe and verify project failure (whereas success is perfectly observable and contractible). Project failure is observed by the agent but possibly not observed by the principal. Because it is efficient to terminate financing once the project fails, the principal would like the agent to disclose failure. However, the agent can hide failure to prevent the termination of financing, which allows the agent to continue the project after failure and yields private benefits from doing so. In addition, as failure and reports thereof are not verifiable by the principal, the agent can misreport (i.e., fake)

[1]This Chapter is based on Mayer (2020a).

failure before it occurs, which precipitates the termination of financing and ends project development prematurely.

To incentivize the agent to disclose failure at the time it occurs, the contract stipulates rewards (pay) for failure. These rewards for failure must decrease over time, because otherwise the agent would delay disclosure of failure. However, as the agent is rewarded for disclosing failure, he finds it tempting to misreport (fake) failure before it occurs and to seize these rewards for failure. To prevent this outcome and hence to incentivize the agent to continue project development, it becomes necessary to increase the agent's stake in the project by raising rewards for success, leading to excessive rewards for success and agency rents. That is, a tension arises between providing incentives for disclosure of failure and project development.

As a result, the principal faces the following trade-off when designing the contract. On the one hand, financing a failed project is inefficient, so the principal ideally would like the agent to disclose failure and to terminate financing upon failure. On the other hand, incentivizing disclosure of failure is costly, as it leads to excessive agency rents. In light of this trade-off, the optimal contract does not always incentivize disclosure of failure and consists of two distinct stages: an *unconditional financing stage* followed by a *disclosure stage*. During the unconditional financing stage, the principal does not incentivize disclosure of failure, so the project is potentially financed and inefficiently continued after failure. The unconditional financing stage can also be interpreted as contractual grace or probationary period during which the agent is not fired after bad outcomes. Moreover, the agent i) is not paid for failure, ii) receives low rewards for success, and iii) incurs mild punishments for delays (beyond her influence) as rewards for success decrease over time. The unconditional financing stage ends with a soft deadline at which the principal elicits a truthful progress report from the agent on whether the project has failed so far. The principal finances the project over the next stage, i.e., the disclosure stage, if and only if the progress report reveals that the project has not failed yet.

During the disclosure stage, the principal incentivizes disclosure of failure and finances the project until either completion is reported or a hard deadline is reached. Thus, financing is terminated upon failure and therefore performance-sensitive. During this stage, the agent receives high and time-decreasing rewards for failure and success and incurs harsh punishments for delays that include the threat of contract termination. The disclosure stage ends with a hard deadline at which the principal terminates financing regardless of whether the project is still profitable to pursue. In summary, the optimal contract consists of different (financing) stages, and involves tolerance towards failure through both i) pay for failure and ii) a grace period. The provision of financing becomes more performance-sensitive over time and across stages. The agent's rewards for failure and incentives are backloaded over time and across stages.

The analysis of the optimal contract has implications for the design of venture capital, R&D financing, and executive compensation contracts. The model predicts that optimal venture capital contracts involve distinct financing stages, whereby the continuation of financing becomes more performance-sensitive over time and the entrepreneur's incentives and liquidation rights in the startup (corresponding to rewards for failure) are backloaded. At the beginning of a new financing stage, information is released through progress reports, so the valuation of the startup should increase. As the provision of unconditional financing limits agency rents, optimal financing contracts for projects that are subject to severe agency conflicts involve a relatively long unconditional financing stage (and soft deadline) that is followed by a relatively short disclosure stage (and hard deadline). That is, precisely when agency conflicts in project development are severe (e.g., due to intangibility or complexity of the project), it is optimal to provide substantial unconditional financing that is independent of performance or reported progress.

Moreover, our results suggest that optimal financing of R&D projects involves several stages and milestones, whereby the continuation of financing is contingent on reported progress by insiders. Progress reports are less (more) frequent in early (later) stages of R&D financing. Optimal contracts for R&D workers stipulate back-loaded incentives and tolerance towards failure via both i) an initial probationary (grace) period and ii) dollar rewards for failure. In particular, it is optimal to pay the agent for failure that occurs at a later stage, but not for failure that occurs at an early stage. Interestingly, as progress reports are subject to moral hazard, it becomes optimal to elicit progress reports less frequently when moral hazard in project development is severe.

In the context of executive compensation contracts, the model highlights the tension that while instruments like golden handshakes, golden parachutes, and severance pay provide incentives for truthful disclosure of bad outcomes, they also undermine incentives to generate good outcomes. As a result, these instruments should not be in place in the early stages of the contract (i.e., CEO tenure). Crucially, once in place, the dollar value of these instruments should decrease over time and the CEO should be provided strong incentives through performance pay.

Next, we show that agency conflicts can induce over- or under-provision of financing (i.e., over- or under-investment) relative to the net present value (NPV) criterion. Here, under-provision of financing (under-investment) refers to a situation, wherein agency conflicts hamper financing of a project with positive NPV, and over-provision of financing (over-investment) refers to financing of projects with negative NPV. During the contract's unconditional financing stage, the principal may inefficiently finance a failed project with negative NPV, leading to over-investment. In contrast, during the contract's disclosure stage, incentive provision may require premature termination of a project with positive NPV, leading to under-investment. As a result, within the optimal contract, there can be instances of both under- and over-investment. The model implies that over-investment (under-investment) tends to arise for projects that (do not) generate results relatively quickly. In the context of venture capital financing, the model therefore predicts inefficiently high venture capital investment in projects that generate results quickly — such as information technology projects (Kerr, Nanda, and Rhodes-Kropf (2014)) — and inefficiently low investment in projects that do not generate results quickly — such as renewable energy production. Interestingly, the optimal financing contract for projects that generate results quickly stipulates substantial unconditional financing but a (relatively) short hard deadline, which is broadly consistent with a "spray and pray" investment approach adopted by venture capitalists for these types of projects (Ewens, Nanda, and Rhodes-Kropf (2018)). According to our model, this "spray and pray" approach leads to over-investment and is inefficient from a social perspective.

We study a model extension in which the project can still turn into a success after (interim) failure. Under these circumstances, the agent would like to hide (interim) failure from the principal to prevent contract termination in order to gamble on a later success following the motto "Fake it till you make it." Interestingly, the model then implies that more ambitious or risky projects are subject to less severe agency problems and thus are preferred by investors. When the project is ambitious and unlikely to turn into success after failure, it is less tempting to hide failure to gamble on success, which alleviates moral hazard. In the context of venture capital financing, the model predicts that venture capitalists seek to finance very ambitious and risky startups, i.e., potential *unicorns*, even if this choice is not supported by the net present value criterion.[2]

Finally, we analyze the role of monitoring and inspections in incentive provision. For this purpose, we assume that the principal can inspect the project's progress at a cost. Upon inspection, the principal learns whether the project has failed so far. That is, an inspection enables the principal to detect whether the agent hides failure and allows punishment of that misbehavior with termination. The optimal contract (with monitoring) features several unconditional financing and disclosure stages, whereby periodic inspections occur at deterministic dates and an unconditional financing stage is followed by a disclosure stage. At the end of a disclosure stage, the principal inspects the project and grants financing for the next stage — which is an unconditional financing stage — if the inspection reveals that the project is still profitable to pursue. Importantly, we find that inspections and contract termination are substitutes in incentive provision, generating the prediction that venture capitalists with high expertise and ability to oversee/inspect a startup's operations tend to provide financing over a longer horizon.

This paper builds on the literature that studies dynamic contracting in continuous time, starting with DeMarzo and Sannikov (2006), Biais, Mariotti, Plantin, and Rochet (2007), and Sannikov (2008). Recent contributions include He (2009a, 2011, 2012), Hoffmann and Pfeil (2010), Biais, Mariotti, Rochet, and Villeneuve (2010), DeMarzo, Fishman, He, and Wang (2012), Edmans, Gabaix, Sadzik, and Sannikov (2012), Hartman-Glaser, Piskorski, and Tchistyi (2012), Zhu (2012), DeMarzo and Sannikov (2016), Marinovic and Varas (2019a), Gryglewicz, Mayer, and Morellec (2020), and Hoffmann, Inderst, and Opp (2019). Relatedly, DeMarzo and Fishman (2007b,a) study dynamic financial contracting in a discrete time setting. Piskorski and Westerfield (2016), Orlov (2018), and Malenko (2019a) analyze incentive provision with optimal dynamic contracts and monitoring. Likewise, Halac and Prat (2016), Varas, Marinovic, and Skrzypacz (2020a), and Marinovic and Szydlowski (2019) characterize optimal monitoring in dynamic settings but do not focus on optimal contracts. None of these papers studies optimal incentive provision for disclosure of bad outcomes.

Our framework is similar to Green and Taylor (2016) and Varas (2017). Green and Taylor (2016) study contracting for a multistage project when intermediate progress is privately observed. They derive an optimal contract that involves a period during which financing is guaranteed whereby the agent is rewarded for good but never for bad outcomes. Varas (2017) studies managerial short-termism in a dynamic model and finds that the threat of termination and deferred payouts for success provide efficient incentives. Our analysis differs from these papers along several dimensions. First, in our model project development may lead to two types of adverse outcomes, failure and completion delays, whereas in Green and Taylor (2016) and Varas (2017) only completion delays may arise. Second, our paper studies incentives to truthfully disclose, i.e., not to hide or to fake, bad outcomes. In Green and Taylor (2016) and Varas (2017) the moral hazard problem is different in that the agent is tempted to fake but never tempted to hide good outcomes. Third, in Green and Taylor (2016) and Varas (2017), the agent is not rewarded for failure.

Our paper builds on the literature on optimal incentive schemes for venture capital, such as Bergemann and Hege (1998, 2005), or for financing innovation in a broader sense, such as Holmstrom (1989), Aghion and Tirole (1994), and Nanda

[2] A unicorn is a privately held startup company with a valuation that exceeds one billion dollars. Recent unicorns include Airbnb, Uber and Pinterest.

and Rhodes-Kropf (2017).[3] Conceptually, our paper is related to Manso (2011), who studies incentive provision in a static multi-armed bandit setting in which the agent can either exploit existing technologies or explore new technologies through experimentation. Our model is dynamic and differs from Manso (2011) mainly in the following aspects. First, we study incentive provision for information disclosure rather than for experimentation. Second, we demonstrate that to provide incentives it can be sometimes optimal to reward failure but it is always necessary to penalize delays; delays do not arise in Manso (2011).

Last, our paper is related to the literature on dynamic adverse selection — see e.g. Daley and Green (2012, 2016, 2019), Morellec and Schürhoff (2011), and Grenadier and Malenko (2011) — and to the literature on dynamic information disclosure (without optimal contracts) — see e.g. Jovanovic (1982), Acharya, DeMarzo, and Kremer (2011), Guttman, Kremer, and Skrzypacz (2014), Marinovic and Varas (2016), Bertomeu, Marinovic, Terry, and Varas (2019), and Halac and Kremer (2019).[4]

2.2 Model setup

Time t is continuous and defined over $[0, \infty)$. A principal (she) finances a project that is developed by an agent (he) with limited liability and zero wealth. The principal has deep pockets, can fully commit to any long term contract, and possesses all bargaining power when signing a contract with the agent.[5] Both the principal and the agent are risk neutral, do not discount payoffs, and have an outside option equal to zero. Per unit of time, project development requires funds $\kappa > 0$ from the principal. The principal can always terminate financing and project development and does so at an endogenous time $T_0 \in [0, \infty]$.

The project completion time τ is uncertain and arrives according to a Poisson process N_t, in that $\tau = \inf\{t \geq 0 : dN_t = 1\}$. The intensity of N_t equals $\Lambda > 0$ for $t \leq T_0$ and zero for $t > T_0$, meaning that the project cannot be completed anymore after financing is terminated. This specification implies that the expected time to completion equals $\frac{1}{\Lambda}$ provided the project receives sufficient financing. Project completion results in one of two possible outcomes, called *success* and *failure*. Upon completion at time τ, with probability $a_\tau p$, the project is successful and yields terminal payoff $\mu > \frac{2\kappa}{\Lambda p}$ to the principal. Otherwise, with probability $1 - a_\tau p$, the project fails and yields no terminal payoff. Here, $a_\tau \in \{0, 1\}$ is the agent's effort (just before time τ), and $p \in (0, 1)$ is exogenous.[6] Notably, without frictions and costless effort, the project has net present value (NPV) $p\mu - \frac{\kappa}{\Lambda} > 0$ and it is efficient to finance the project until time τ.

Over a short period of time $[t, t + dt]$, the project succeeds with probability $a_t p \Lambda dt$, fails with probability $(1 - a_t p)\Lambda dt$, and does not complete with probability $1 - \Lambda dt$. Similar to Board and Meyer-ter Vehn (2013) and Hoffmann and Pfeil (2021), the

[3]Kerr and Nanda (2015) and Lerner and Nanda (2020) provide excellent reviews of the academic literature on financing entrepreneurship and innovation.

[4]Papers on optimal information disclosure in static settings include Fishman and Hagerty (1989, 1990), Diamond and Verrecchia (1991) and, more recently, Szydlowski (2019). For a literature review on this topic, we refer to Beyer, Cohen, Lys, and Walther (2010).

[5]In the context of venture capital financing, the commitment power of the principal representing the venture capitalist can be motivated and micro-founded using the findings of Donaldson, Piacentino, and Thakor (2020).

[6]The terminal payoff can be interpreted broadly and, for instance, can represent immediate monetary payoffs, future expected cash flows, or the principal's private value of a technological achievement.

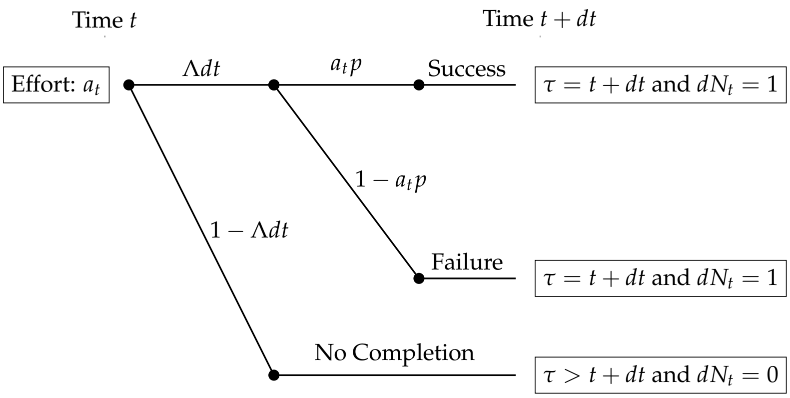

FIGURE 2.1: Heuristic timing over $[t, t + dt)$. The branches of the tree contain the probabilities of the respective random event over $[t, t + dt)$.

agent chooses effort $a_t \in \{0, 1\}$ before the random event $dN_t \in \{0, 1\}$ realizes over $[t, t + dt)$. Figure 1 illustrates the heuristic timing over $[t, t + dt)$.

Project development is subject to the following frictions. First, project failure is hard for the principal to observe or verify. When the project fails, failure is *publicly* observed (i.e., observed by both principal and agent) with probability $\pi \in [0, 1]$. Publicly observed failure is contractible. Otherwise, with probability $1 - \pi$, failure is *privately* observed by the agent but not observed by the principal, who also cannot verify failure. For instance, $\pi \in (0, 1)$ captures the fact that the insiders developing the project can hide certain bad outcomes from outside investors, while it is difficult or even impossible to hide all types of bad outcomes. The agent may report (i.e., disclose) failure, but the principal cannot verify the reported failure. As a result, the agent can hide or misreport (fake) failure.[7] However, because it is efficient to terminate financing after failure at time τ, the principal would ideally like the agent to disclose failure truthfully. Unlike failure, success is publicly observable, verifiable, and contractible. This assumption reflects that the agent is tempted to hide bad rather than good outcomes and that it is harder to fake good rather than bad outcomes.

Second, effort a_t is not observed by the principal and the agent derives private benefits from shirking $(1 - a_t)\phi$ for $t \leq T_0$, giving rise to moral hazard. After financing is terminated, private benefits are zero. Even though not explicitly modeled here, private benefits (from shirking) may arise from the inefficient diversion of funds, as e.g. in DeMarzo and Sannikov (2006), and, therefore, pertain as long as the principal allocates funds to the project (i.e., only for $t \leq T_0$). Crucially, effort after time τ is redundant so that the agent optimally chooses $a_t = 0$ for $t \in (\tau, T_0]$. This leads to private benefits ϕ from operating the project after time τ, which only pertain if the agent hides project failure and averts termination after time τ. We assume $0 \leq \phi < \kappa$, implying that shirking is inefficient. Last, certain modeling assumptions

[7]Publicly and privately observed failure can be interpreted as two different types of bad outcomes. That is, publicly observed failure is a "very bad" outcome the agent cannot hide and privately observed failure is a bad outcome the agent can hide. I would like to thank Neng Wang for pointing out this interpretation.

— e.g., observable success and exogenous completion — and the robustness of our findings are discussed in Section 2.7.2.

Contracting problem

At time zero, the principal offers the agent a contract $\mathcal{C} = (c, T)$ specifying cumulative payments (i.e., wages) c and a deadline T. At the deadline, project development and financing are terminated. For notational convenience, payments c do not include project development costs κ, which are paid by the principal up to time T_0. Facing the contract $\mathcal{C}$, the agent chooses effort a and time τ^A when he reports completion. This implies that the principal finances the project until either completion is reported or the deadline is reached. That is: $T_0 = T \wedge \tau^A$. We focus on contracts $\mathcal{C}$ that induce full effort $a_t = 1$ at all times $t \in [0, T \wedge \tau]$. To solve the model, we conjecture that the (optimal) deadline T is deterministic. Lemma 9 in Appendix 2.10.4 verifies that the (optimal) deadline is indeed deterministic, in that stochastic termination is not optimal.[8]

We characterize the agent's strategy using the time at which he reports completion (rather than using the time at which he reports failure). Because success is publicly observable, this notation implies for the case of success $\tau^A = \tau$. Likewise, we adopt the notation $\tau^A = \tau$ when failure is publicly observed. Conversely, $\tau^A \neq \tau$ means that the agent misreports failure. In particular, the agent can misreport failure in two ways. First, he can *hide failure*, in which case $\tau^A > \tau$. Second, because reported failure is not verifiable, the agent can *fake failure* and misreport failure even if it has not occurred, in which case $\tau^A < \tau$.[9] An interpretation is that the agent liquidates, exits, and sells the project prematurely to new investors, although the agent could still add value to the project and project development is not complete. Alternatively interpreted, the agent causes or precipitates failure and termination by not continuing to work hard in project development.

As the principal and the agent do not discount, it is optimal to pay the agent only at time τ^A, when reported success or failure is informative about the agent's effort. Because the completion timing as such is not informative about effort, payments before time τ^A do not help to incentivize effort. In fact, they even render it attractive to hide failure and so generate dis-incentives. Thus

$$dc_t = \alpha_t \mathbf{1}_{\{\text{Success at time } t\}} + \beta_t \mathbf{1}_{\{\text{Failure report at time } t\}} + \gamma_t \mathbf{1}_{\{\text{Failure observed at time } t\}} \tag{2.1}$$

where α_t is the payment to the agent if the project succeeds at time t (in which case $\tau = \tau^A = t$). The agent's payment for failure depends on whether the principal observes failure. Specifically, β_t is the agent's pay when he reports failure at time t, while γ_t is the agent's pay when failure is publicly observed at time t. As will become clear later, payments for *privately* observed and reported failure $\beta_t > 0$ are necessary to incentivize disclosure of failure. In contrast, payments for *publicly* observed failure are not necessary to incentivize disclosure of failure and, in addition, do not motivate effort. Hence, we conjecture and verify that it is optimal not to pay the agent for observed failure, in that $\gamma_t = 0$.

[8] I would like to thank an anonymous referee for encouraging me to add Lemma 9 and its formal proof.

[9] This is akin to assuming that the agent can destroy payoff, which is a frequent assumption in the financial contracting literature (see e.g. Innes (1990)).

The agent maximizes his payoff, stemming from wage payments and private benefits:

$$W_0 := \max_{a,\tau^A} \mathbb{E}^A \left[\int_0^{T\wedge\tau^A} (dc_t + \phi(1-a_t)dt) \right]. \tag{2.2}$$

Here, superscript A indicates that the expectation is taken to be conditional on the agent's information, which may differ from the principal's information. The principal chooses contract $\mathcal{C}$ to maximize expected project payoff net of the cost of financing project development and compensating the agent:

$$F_0 := \max_{\mathcal{C}} \mathbb{E} \left[\int_0^{T\wedge\tau^A} (\mu dS_t - \kappa dt - dc_t) \right]. \tag{2.3}$$

where the expectation is conditional on the principal's information. Here, $dS_t = 1$ indicates project success at time t. Because the agent is protected by limited liability, it follows that $dc_t \geq 0$, i.e., $\alpha_t, \beta_t, \gamma_t \geq 0$, for all $t \geq 0$.

2.3 Model solution

We solve the model in several steps. First, to provide a starting point for our analysis, we analyze the benchmark, in which project failure is publicly observable (i.e., $\pi = 1$) and contractible, yet the moral hazard problem with respect to hidden effort remains. Second, we focus on *full disclosure contracts* that always incentivize the agent to disclose failure truthfully. Third, we argue why full disclosure contracts are in general not optimal and show how they can be improved through the provision of *unconditional financing*.

2.3.1 Second-best benchmark

We start by analyzing the second-best benchmark, in which project failure is observable and contractible in that $\pi = 1$. With observable failure, the optimal contract takes a simple form: there is no deadline (i.e., $T = \infty$) and the agent is only paid for success. That is, $\alpha^{SB} = \frac{\phi}{\Lambda p} > \beta^{SB} = \gamma^{SB} = 0$.

The intuition behind this result is as follows. First, note that the agent controls the project's propensity to succeed at time τ but not the completion timing τ. Therefore, the completion timing τ as such is not informative about the agent's effort, implying that the agent's compensation should not be contingent on τ. Thus, the agent is not punished for completion delays and not fired before project completion.

Second, to motivate effort, it is necessary to pay the agent more for success than for failure, leading to the incentive constraint

$$\alpha_t - \gamma_t \geq \frac{\phi}{\Lambda p}. \tag{2.4}$$

To derive (2.4), suppose the agent shirks over a short period of time $[t, t+dt)$. Then, the agent derives private benefits ϕdt and the project completes with probability Λdt, resulting in failure and pay γ_t. In contrast, if the agent exerts effort $a_t = 1$ over $[t, t+dt)$, the agent does not derive any private benefits and the project completes with probability Λdt, resulting in failure and pay γ_t with probability $1-p$ and in success and pay α_t with probability p. Thus, exerting effort over $[t, t+dt)$ is optimal if and only if

$$\underbrace{((1-p)\gamma_t + p\alpha_t)\Lambda dt}_{\text{Payoff with } a_t=1} \geq \underbrace{\phi dt + \gamma_t \Lambda dt}_{\text{Payoff with } a_t=0},$$

which simplifies to (2.4).

Note that because of $\pi = 1$, the choice of β becomes redundant. Incentives are captured by (net) rewards for success $\alpha_t - \gamma_t$. Payments for observed failure motivate shirking and hence are optimally set to $\gamma_t = 0$. In addition, (2.4) is tight to minimize agency costs. Incentives $\alpha_t - \gamma_t$ must be stronger, if moral hazard is more severe. Moral hazard is severe when private benefits from shirking ϕ or the expected duration of the project development phase $1/\Lambda$ are large.

Proposition 1 *Suppose that failure is publicly observable and contractible, in that $\pi = 1$. Then, the optimal contract C^{SB} is stationary with $T^{SB} = \infty$ and $\alpha^{SB} = \frac{\phi}{\Lambda p} > \beta^{SB} = \gamma^{SB} = 0$. The principal's payoff equals $F^{SB} = p\mu - \frac{\phi+\kappa}{\Lambda}$. The agent's payoff equals $W^{SB} = \frac{\phi}{\Lambda}$.*

With this clean benchmark in hand, we now analyze the problem with $\pi \in [0, 1)$.

2.3.2 Full disclosure contracts

Because it is efficient to terminate financing at time τ, it is natural to study *full disclosure contracts* that incentivize truthful disclosure of failure. That is, a full disclosure contract induces $\tau^A = \tau$.

Truth telling incentives

We start by characterizing the agent's incentives to disclose failure truthfully. Note that $\beta_t(1-\pi)$ is the agent's expected payment upon failure at time t, which is proportional to β_t. We thus refer to β_t simply as "rewards for failure" rather than "rewards for reported failure." Define the agent's continuation payoff under truthful reporting (i.e., $\tau^A = \tau$) and full effort as

$$W_t = \mathbb{E}_t \left[\int_t^{T\wedge\tau} dc_s \right] = \int_t^T e^{-\Lambda(s-t)} \Lambda \big((1-p)(1-\pi)\beta_s + p\alpha_s \big) ds, \tag{2.5}$$

where the (second) equality uses integration by parts, and recall that $\gamma_s = 0$. Hence:

$$\dot{W}_t = \Lambda W_t - \Lambda((1-p)(1-\pi)\beta_t + p\alpha_t), \tag{2.6}$$

where "dot" denotes the time derivative (i.e., $\dot{W}_t = \frac{dW_t}{dt}$). Hence, $\dot{W}_t < 0$ means that the agent is *punished for delays* (beyond his influence).

First, as reported failure is not verifiable, the agent can always misreport (fake) failure before it actually occurs, which precipitates the termination of financing and ends project development. Faking failure leads to a loss of the continuation value under truth telling W_t and a reward for reporting failure β_t and, therefore, is suboptimal if and only if

$$W_t \geq \beta_t. \tag{2.7}$$

That is, to provide incentives not to misreport failure before it occurs and to incentivize project development, it is necessary to grant the agent a sufficiently large stake W_t in the project. Note that the optimal contract always incentivizes the agent not to misreport failure before it occurs. If the agent finds it optimal to fake failure at some time t with $\beta_t > 0$, the agent reports failure and hence precipitates termination at time t, while being paid $\beta_t > 0$. Then, the principal could improve her payoff by stipulating termination with zero severance pay for the agent at time t.

Second, let us analyze the agent's incentives to hide failure. To obtain some intuition, suppose the project fails at time t and failure is privately observed by the agent. Reporting failure at time t yields pay β_t. In contrast, reporting failure at time $t + dt$ not only yields pay β_{t+dt} but also benefits from operating the project over $[t, t + dt]$, ϕdt. Hence, the agent is better off disclosing failure truthfully at time t if and only if

$$\beta_t \geq \beta_{t+dt} + \phi dt.$$

Letting $dt \to 0$ yields

$$\dot{\beta}_t \leq -\phi. \tag{2.8}$$

Because failure can potentially occur at any time $t \in [0, T]$, a full disclosure contract must satisfy (2.8) for all $t \in [0, T]$. Thus, we can integrate (2.8) to obtain

$$\beta_t \geq \beta_s + (s - t)\phi \quad \text{for all} \quad s \in [t, T]. \tag{2.9}$$

Indeed, it follows that the agent prefers to report failure at time t rather than at any time $s > t$.

Importantly, the agent's limited liability requires $\beta_t \geq 0$. Once $\beta_t = 0$, the contract cannot provide truth telling incentives anymore by decreasing β_t. Under these circumstances, the only way to ensure truth telling is contract termination, in that $T = \inf\{t \geq 0 : \beta_t = 0\}$. Contract termination also implies that the agent never profits from hiding failure completely. Hiding failure (that occurs at time t) completely allows the agent to derive private benefits up to time T, i.e., $\phi(T - t)$, which by (2.9) is smaller than the reward for failure β_t.

In the following, it is convenient to keep track of the agent's (off-equilibrium) continuation payoff w_t, in case he has privately observed failure at some time t' with $t' \leq t$ but not reported it yet:

$$w_t := \max_{\tau^A \in [t,T]} \left[\phi(\tau^A - t) + \beta_{\tau^A}\right]. \tag{2.10}$$

Note that a full disclosure contract implies for any $t \in [0, T]$ that w_t is maximized for $\tau^A = t$, leading to $\beta_t = w_t$ and $\dot{w}_t = \dot{\beta}_t$ for all $t \in [0, T]$. With the agent's expected pay for failure $(1 - \pi)w_t$, we obtain the incentive condition w.r.t. effort a_t:

$$\alpha_t \geq (1 - \pi)w_t + \frac{\phi}{\Lambda p}. \tag{2.11}$$

The derivation of condition (2.11) is analogous to the derivation of condition (2.4) upon replacing γ_t with $(1 - \pi)w_t$.

Solving for the optimal full disclosure contract

Using integration by parts and $\tau^A = \tau$, we can rewrite the principal's continuation payoff for any $t \in [0, T \wedge \tau^A]$ as

$$F_t = \mathbb{E}\left[\int_t^{T\wedge\tau^A} (\mu dS_s - \kappa ds - dc_s)\right] \tag{2.12}$$

$$= \int_t^T e^{-\Lambda(s-t)}\Big(\Lambda p\mu - \kappa - \Lambda\big((1-p)(1-\pi)w_s + p\alpha_s\big)\Big)ds. \tag{2.13}$$

The optimal full disclosure contract maximizes (2.12) subject to the incentive constraint w.r.t. effort (2.11) and the incentive constraints w.r.t. truthful information disclosure (2.7) and (2.8) for all $t \in [0, T]$. The incentive condition (2.8) constrains the control of the level of β (or equivalently w) and thus β (or equivalently w) enters the principal's dynamic optimization problem as state variable, while the change in β or w (i.e., $\dot{\beta}$ or $\dot{w}$) is a control variable.

To minimize agency rents, the incentive condition (2.7) is tight, in that $W_t = w_t = \beta_t$ for all $t \in [0, T]$ and hence the principal's optimization features a single state variable w. This implies that one can express the principal's payoff as function of w, $F(w)$. Differentiating (2.12) with respect to time t and using $\frac{dF_t}{dt} = \frac{dF(w_t)}{dw_t}\frac{dw_t}{dt} = F'(w_t)\dot{w}_t$, we obtain that the principal's value function solves the following HJB equation on the endogenous state space $[0, w_0]$:

$$\Lambda F(w) = \max_{\dot{w},\alpha} \left\{ \Lambda p\mu - \kappa - \Lambda((1-p)(1-\pi)w + p\alpha) + F'(w)\dot{w} \right\} \tag{2.14}$$

subject to all relevant incentive constraint. Some observations are in order. First, because it always possible to shorten the financing deadline and to reduce w, it holds in optimum that $F'(w) \geq 0$ for all $w \leq w_0$. Then, the maximization w.r.t. $\dot{w}$ yields $\dot{w} = -\phi < 0$ so that condition (2.8) is binding. Second, because of the agent's limited liability, the contract is terminated once $w = \beta = 0$, yielding $f(0) = W(0) = 0$. Third, the value function is strictly concave, reflecting that termination is inefficient. The concavity of the value function also implies that randomized termination of a full disclosure contract is not optimal (for details and a more general statement see Lemma 9).

Importantly, rewards for success $\{\alpha_t\}$ determine the size of the agent's stake in the project W_t and hence his incentives to fake bad outcomes (see (2.7)). To minimize agency costs, (2.7) is tight and $W_t = w_t = \beta_t$ for all $t \in [0, T]$, leading to[10]

$$\alpha_t = (1-\pi)w_t + \frac{\phi}{\Lambda p} + \frac{\pi}{p}w_t = \left(1 - \pi + \frac{\pi}{p}\right) w_t + \frac{\phi}{\Lambda p}. \tag{2.15}$$

Thus, the incentive condition for effort (2.11) is slack, when $\pi > 0$.[11] The reason is that the agent requires rewards $\beta > 0$ to disclose failure. Because rewards for failure provide the agent with incentives to fake failure, it becomes necessary to increase his stake in the project by raising rewards for success beyond what is needed to motivate effort. Also observe that both rewards for success and failure decrease over time, inducing punishment for delays $\dot{W}_t = -\phi < 0$.

Plugging (2.15) and $\dot{w} = -\phi$ back into the HJB equation (2.14), one obtains a linear first order ODE that admits the closed form solution

$$F(w) = \left(\mu p - \frac{\kappa}{\Lambda}\right)\left(1 - \exp\left(\frac{-w\Lambda}{\phi}\right)\right) - w. \tag{2.16}$$

[10]For a derivation, note that $W_t = w_t = \beta_t$ for all $t \in [0, T]$ implies that $\dot{W}_t = \dot{w}_t$ for all $t \in [0, T]$. Hence, by (2.6) with $w_t = \beta_t$: $\dot{W}_t - \dot{w}_t = \Lambda(W_t - (1-\pi)w_t - p(\alpha_t - (1-\pi)w_t)) + \phi = \phi - \Lambda p(\alpha_t - (1-\pi)w_t) + \Lambda\pi w_t = 0$. Thus, $\alpha - (1-\pi)w_t = \phi/(\Lambda p) + \pi w_t/p$.

[11]Expected rewards for failure $\beta_t(1-\pi)$ increase in $1-\pi$ and higher (expected) rewards for failure require higher rewards for success to motivate effort. When $\pi = 0$, incentives for effort require particularly high rewards for success that are at the same time sufficient to incentivize the agent not to fake failure.

The starting value w_0 is chosen to maximize payoff at time zero, $F(w_0)$. Thus, w_0 solves the first-order optimality condition $F'(w_0) = 0$ so that

$$w_0 = \phi T = \frac{\phi}{\Lambda} \ln\left(\frac{\Lambda \mu p - \kappa}{\phi}\right). \tag{2.17}$$

We summarize our findings in the following proposition.

Proposition 2 *Under the optimal full disclosure contract $\mathcal{C}$, at time t with $w_t = w$, the principal's value $F(w)$ solves* (2.14). *The contract $\mathcal{C}$ stipulates $\dot{w}_t = -\phi$ and termination at time $T = \inf\{t \geq 0 : w_t = 0\}$. Payments satisfy* (2.15), *$\beta_t = w_t = W_t$, and $\gamma_t = 0$. The value w_0 solves $F'(w_0) = 0$.*

The optimal full disclosure contract differs from the second-best contract mainly in three aspects: i) rewards for failure, ii) excessive rewards for success beyond what is needed to motivate effort, and iii) punishments for delays, including the threat of contract termination. Rewards for failure and punishments for delays incentivize the agent not to *hide* bad outcomes, whereas sufficiently high rewards for success incentivize the agent not to *fake* bad outcomes.

The optimal full disclosure contract is not unique. The reason is that rewards for observed failure boost the agent's stake in the project and hence generate incentives not to fake bad outcomes. As a result, there exists an optimal full disclosure contract that pays the agent for observed failure and pays the agent less for success. This contract maintains incentive compatibility but stipulates weaker incentives to exert effort. We focus without loss of generality on the optimal (full disclosure) contract that maximizes incentives. Notably, all optimal (full disclosure) contracts share the same key characteristics and stipulate punishments for delays, rewards for failure, and excess pay for success (i.e., $\alpha_t > w_t(1-\pi) + \frac{\phi}{\Lambda p}$). We present the generalization of Proposition 2 in the Appendix in Proposition 5.

Notably, the value function $F(w)$ does not depend on π. The reason is that within a full disclosure contract, agency costs are independent of π. An increase in π makes it easier to incentivize the agent not to hide failure, leading to lower rewards for failure $\beta_t(1-\pi)$. This, however, reduces agency rents W_t, which generates incentives to misreport failure before it occurs and so requires higher rewards for success α_t. In other words, a tension arises between incentivizing disclosure of failure and incentivizing project development. In light of this tension, a full disclosure contract is not optimal and, therefore, the optimal contract does not always incentivize disclosure of failure.

2.3.3 The optimal contract

Suppose that the contract stipulates no rewards for failure over $[t_0, t_1)$, in that $\beta_t = \gamma_t = 0$ for $t \in [t_0, t_1)$. Consequently, the agent never reports and never fakes failure over $[t_0, t_1)$. As the contract does not incentivize disclosure of failure over $[t_0, t_1)$, it is also not terminated over $[t_0, t_1)$ due to disclosure of failure. In other words, the principal provides *unconditional financing* over $[t_0, t_1)$. Crucially, the provision of unconditional financing generates incentives not to fake failure (i.e., relaxes incentive constraint (2.7)) and hence limits agency rents but comes at the expense that the project may be continued and financed after failure, which is inefficient. Although a contract could stipulate unconditional financing over several distinct time intervals, we focus in the following discussion on the latest interval in time denoted by $[t_0, t_1)$. Thereafter, we verify that under the optimal contract there is only one (connected) time interval with unconditional financing.

Incentives

Consider that $\beta_t = 0$ for $t \in [t_0, t_1)$ and that the contract incentivizes truthful information disclosure from time t_1 onwards up to a deadline $T > t_1$. If the project fails at some time $t \in [t_0, t_1)$ and failure is privately observed by the agent, the agent does not report failure up to time t_1, when he receives pay $\beta_{t_1} > 0$. Thus, the agent's continuation payoff after failure at time t is:

$$w_t := (t_1 - t)\phi + \beta_{t_1} \text{ for } t \in [t_0, t_1],$$

so that $\dot{w}_t = -\phi < 0$. To incentivize effort, the agent's payoff after success must sufficiently exceed his payoff after failure in that (2.11) holds, that is, $\alpha_t \geq (1 - \pi)w_t + \frac{\phi}{\Lambda p}$.

Unconditional financing

Next, we heuristically determine when it is optimal to provide unconditional financing. First, consider $t_0 = 0$. Over the time interval $[0, t_1)$, the agent merely requires incentives for effort and it is optimal to provide minimal incentives, in that

$$\alpha_t = (1 - \pi)w_t + \frac{\phi}{\Lambda p} \quad \text{for all} \quad t \in [0, t_1). \tag{2.18}$$

That is, the provision of unconditional financing facilitates the stipulation of low rewards for success. Low rewards for success in turn reduce both punishments for delays and the agent's stake in the project, thereby reducing incentives and agency costs relative to a full disclosure contract. Formally, $W_t < w_t$ and $0 > \dot{W}_t > \dot{w}_t = -\phi$ for $t < t_1$. Overall, unconditional financing limits agency rents W_t but may lead to inefficient financing of a failed project.[12]

Second, we argue that unconditional financing starting from $t_0 > 0$ is sub-optimal. If $t_0 > 0$, the contract incentivizes truthful disclosure of failure just before time t_0 and just after time t_1, which by (2.7) requires that $W_{t_i} \geq w_{t_i}$ for $i = 0, 1$. That is, the principal cannot reduce the agent's stake W_{t_0} (i.e., agency costs) by providing unconditional financing after time t_0, while incentivizing truth telling before time t_0. More intuitively, unconditional financing after time t_0 dilutes truth telling incentives before time t_0. To avoid this inefficient dilution of incentives, it must be that $t_0 = 0$. Formally, at any time t with $W_t \geq w_t$, the optimal continuation contract is a full disclosure contract, as characterized in Proposition 2. It readily follows that there is maximally one connected time interval, during which the optimal contract stipulates unconditional financing.

As a result, the optimal contract stipulates unconditional financing over some time period $[0, t_1)$. After time t_1, the optimal contract incentivizes truthful disclosure of failure and hence becomes a full disclosure (continuation) contract with deadline $T \geq t_1$, yielding payoff $F(w_{t_1})$ to the principal and payoff $w_L := w_{t_1} = W_{t_1}$ to the agent. Also note that $w_t = \beta_t = W_t$ for all $t \geq t_1$.

[12]Note that $W_t = \int_t^T e^{-\Lambda(s-t)}\Lambda((1-p)(1-\pi)w_s + p\alpha_s)ds$. Hence: $\dot{W}_t = \Lambda W_t - \Lambda(w_t(1-\pi) + p(\alpha_t - w_t(1-\pi)))$ and $\dot{W}_t - \dot{w}_t = \Lambda W_t - \Lambda(w_t(1-\pi) + p(\alpha_t - w_t(1-\pi))) - \dot{w}_t = \Lambda(W_t - w_t) + \Lambda\pi w_t$, where we plugged in $\alpha_t = w_t(1-\pi) + \phi/(\Lambda p)$ and $\dot{w}_t = -\phi$ for $t < t_1$. Integrating this ODE for $t < t_1$ subject to $W_{t_1} = w_{t_1}$ yields $W_t - w_t = -\int_t^{t_1} e^{-\Lambda(s-t)}\Lambda\pi w_s ds < 0$ and hence also $0 > \dot{W}_t > \dot{w}_t = -\phi$.

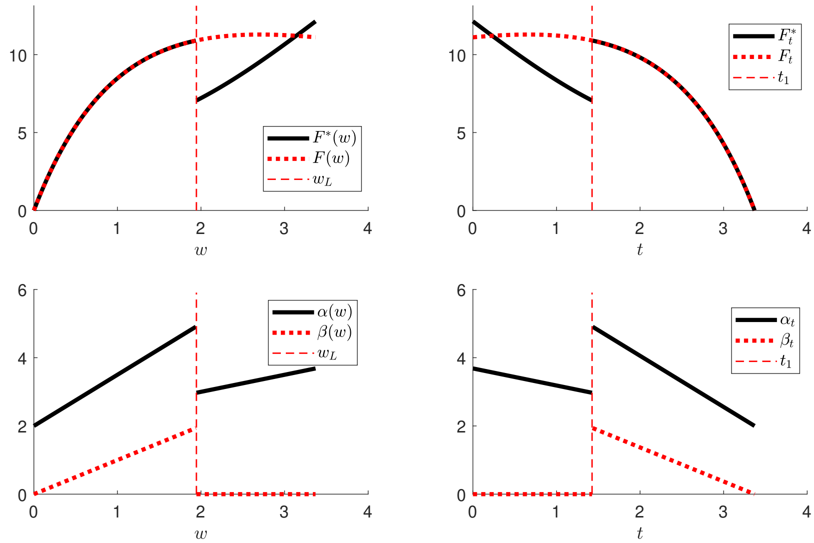

FIGURE 2.2: Numerical example in both the "w-space" and "time space". The parameters are $\mu = 50$, $\kappa = 10$, $\Lambda = 1$, $p = 0.5$, $\pi = 0.5$, and $\phi = 1$.

Solving for the optimal contract

The principal incentivizes disclosure of failure at time t_1 and forms a belief, q_t, of whether the project has failed so far, for times $t < t_1$. One can derive that:[13]

$$q_t = q(w_t) = 1 - e^{-\Lambda(1-p)(1-\pi)t} = 1 - e^{-\frac{\Lambda(1-p)(1-\pi)(w_0-w_t)}{\phi}}. \tag{2.19}$$

Intuitively, the principal faces in addition to the moral hazard problem also an *adverse selection* problem when providing unconditional financing over $[0, t_1)$. In particular, the principal may inefficiently extend financing to a failed project (which is the case with probability q_t).

Given the (previously derived) optimal controls $\beta_t = 0$ and $\alpha_t = w_t(1-\pi) + \frac{\phi}{\Lambda p}$ for $t \in [0, t_1)$ and the fact that w_t co-moves with time t via $\dot{w}_t = -\phi$, one can express the agent's continuation value over the time interval $[0, t_1)$ (i.e., for $w > w_{t_1} = w_L$) as function of w, $W(w)$.[14] As the principal's belief $q(w)$, the agent's continuation value $W(w)$, and rewards for success (failure) $\alpha(w)$ ($\beta(w)$) are along the optimal path functions of w, the state variable w summarizes all contract relevant information. Thus, it is possible to also express the principal's value function over the time interval $[0, t_1)$ (i.e., for $w > w_{t_1} = w_L$) as function of w, $f(w)$.

[13]For a derivation, note that Bayes' rule implies $q_{t+dt} = q_t + (1-q_t)(1-p)(1-\pi)\Lambda dt$, which simplifies in the limit $dt \to 0$ to $\dot{q}_t = (1-q_t)(1-p)(1-\pi)\Lambda$. This linear first order ODE is solved subject to $q_0 = 0$, yielding solution (2.19).

[14]Conjecture that W is a function of w, in that $W_t = W(w_t)$. Recall that $\dot{W}_t = \Lambda W_t - \Lambda(p\alpha_t + (1-p)(1-\pi)\beta_t)$. Using $\beta_t = \beta(w_t) = 0$ and $\alpha_t = \alpha(w_t) = w_t(1-\pi) + \frac{\phi}{\Lambda p}$, it follows that $\dot{W}_t = \Lambda(W_t - w_t) + \Lambda\pi w_t - \phi$. Using $\dot{W}_t = \frac{dW_t}{dw_t}\frac{dw_t}{dt} = W'(w_t)\dot{w}_t$, one obtains that the agent's continuation value solves $W'(w)\dot{w} = \Lambda(W(w) - w) + \Lambda\pi w - \phi$ for $w > w_{t_1} = w_L$ subject to $W(w_L) = w_L$. This confirms that W can be expressed as function of w.

During the unconditional financing stage, over a short period of time $[t, t+dt)$, the principal incurs funding costs κdt, and the following two observable outcomes that trigger termination are possible: i) (with probability $(1-q(w))\Lambda p dt$) the project has not failed so far and succeeds over $[t, t+dt)$ yielding payoff $\mu p - \alpha(w)$, or ii) (with probability $(1-q(w))\Lambda(1-p)\pi dt$) the project has not failed so far but fails over $[t, t+dt)$ and failure is observed which yields payoff zero. This leads to the HJB equation:

$$(1-q(w))\Lambda\big(p+(1-p)\pi\big)f(w) = (1-q(w))\Lambda p(\mu-\alpha(w)) - \kappa + f'(w)\dot{w}, \quad (2.20)$$

with $\dot{w} = -\phi < 0$. At the end of the unconditional financing stage at time t_1 when $w_t = w_L$, the principal asks the agent whether the project has already failed. With probability $q(w_L)$, the project has failed and the principal must compensate the agent for failure $\beta_{t_1} = w_L$. With probability $1-q(w_L)$, the project has not failed yet and the principal realizes the continuation payoff $F(w_L)$, leading to the value matching condition

$$f(w_L) = (1-q(w_L))F(w_L) - q(w_L)w_L. \quad (2.21)$$

In addition, optimal w_L is pinned down by the smooth pasting condition:

$$f'(w_L) = \frac{\partial}{\partial w_L}\left((1-q(w_L))F(w_L) - q(w_L)w_L\right). \quad (2.22)$$

Taking stock, on the state interval $[0, w_L]$ (i.e., the time interval $[t_1, T]$), the value function is characterized by (2.16). On the state interval $(w_L, w_0]$ (i.e., the time interval $[0, t_1)$), the value function is characterized by function $f(w)$, solving (2.20). That is:

$$F^*(w) = \begin{cases} F(w) & \text{for } w \in [0, w_L], \\ f(w) & \text{for } w \in (w_L, w_0]. \end{cases} \quad (2.23)$$

The optimization of the initial payoff $F(w_0)$ with respect to w_0 determines the optimal deadline $T = \inf\{t \geq 0 : w_t = 0\}$. Lemma 2 in Appendix 2.9 presents a closed-form expression for $f(w)$ and Figure 2.2 provides a numerical example of the optimal contract. The upper two panels display the value function under the optimal contract (solid black line) and the value function under the optimal full disclosure contract (dotted red line) both in dependence of w and time, t. Note that the value function exhibits a jump at time t_1 (i.e., at $w = w_L$), when the agent makes a progress report and uncertainty is resolved. The lower two panels display the agent's rewards for success and failure both in dependence of w and time, t. Rewards for success (and failure) decrease during a given financing stage but increase once a new financing stage begins at time t_1. Importantly, we find numerically that the deadline under the optimal full disclosure contract is inefficiently short/tight compared with the optimal contract, so that a full disclosure contract is not optimal. In other words, the provision of unconditional financing leads to a longer financing deadline, thereby improving surplus.[15]

Proposition 3 *The optimal contract does not incentivize disclosure of failure over some time period* $[0, t_1)$ *and becomes a full disclosure contract, characterized in Proposition 2, after time* t_1.

[15]To see this, note that in the upper left panel of Figure 2.2 the peak of the value function under the full disclosure contract (dotted red line) determines the optimal starting value w_0 and deadline $T = w_0/\phi$ under the optimal full disclosure contract. It turns out that the optimal starting value (and deadline) is lower than the optimal starting value (and deadline) under the optimal contract.

1. *With $w_{t_1} = w_L$, the optimal time t_1 is characterized by* (2.21) *and* (2.22)*, while the value function F^* is characterized by* (2.23)*. In addition, $t_1 \to 0$ as $\pi \to 0$ and $t_1 \to T$ as $\pi \to 1$.*

2. *The contract is terminated at time $T = \inf\{t \geq 0 : w_t = 0\}$ and w_0 (and equivalently T) maximizes the principal's initial payoff $F(w_0)$.*

3. *$\alpha_t = w_t(1-\pi) + \phi/(\Lambda p) \geq \beta_t = \gamma_t = 0$, $\dot{w}_t = -\phi < \dot{W}_t < 0$, and $W_t < w_t$ for all $t \in [0, t_1)$.*

4. *$\alpha_t = w_t(1 - \pi + \pi/p) + \phi/(\Lambda p)$, $w_t = \beta_t = W_t$, and $\gamma_t = 0$ for all $t \in [t_1, T)$.*

Taking stock, the optimal contract involves two stages: an *unconditional financing stage* $[0, t_1)$ and a *disclosure stage* $[t_1, T]$. During the unconditional financing stage, the contract does not incentivize disclosure of failure and financing is guaranteed. The unconditional financing stage can also be interpreted as contractual grace or probationary period during which the agent is not fired after bad outcomes. During that stage, the agent i) is not paid for failure, ii) receives relatively low rewards for success, and iii) incurs mild punishments for delays.[16] The unconditional financing stage ends with the soft deadline t_1 at which the principal incentivizes a truthful progress report of whether the project has failed yet. The principal continues financing for the next stage if and only if the progress report reveals that the project is still profitable to pursue (i.e., has not failed).

After time t_1, during the disclosure stage, the principal incentivizes truthful disclosure of failure and finances the project until either completion is reported or the hard deadline T is reached. Note that at the hard deadline T financing is terminated regardless of whether the project has failed yet. During the disclosure stage, the contract stipulates high but time-decreasing rewards for success and failure, inducing harsh punishments for delays. That is, the agent's incentives and performance pay — captured by (net) rewards for success $\alpha_t - w_t(1-\pi)$ and punishments for delays $-\dot{W}_t$ — are stronger during the disclosure stage and hence increase following completion delays. In summary, within the optimal contract, financing is staged for pure incentive purposes, even though there is only a single milestone to complete the project. The contract stipulates tolerance for failure through both i) pay for failure and ii) a grace period (unconditional financing stage). The provision of financing becomes more performance-sensitive over time and across stages. Likewise, the agent's incentives are backloaded, in that they become stronger over time and across stages.

Discussion: uniqueness of the optimal contract

The optimal contract and in particular the choice of α and γ are not unique, precisely because the optimal full disclosure contract is not unique. Proposition 3 describes the optimal contract that maximizes incentives, and its generalization is presented in Proposition 6 in the Appendix. In what follows, we propose a model extension in which the contract from Proposition 3 is (strictly) optimal and the choice of γ is uniquely pinned down. Specifically, we consider that when the agent hides failure from the principal over $[t, t+dt)$, with probability $\lambda^F dt$, project failure becomes publicly observed or detected, leading to project termination and rewards for publicly

[16] Because the principal does not ask the agent to disclose failure, the exact value of β_t is not payoff relevant and it is without loss of generality to set $\beta_t = 0$ over $[0, t_1)$. One could stipulate another value for β_t at times $t < t_1$ as long as $\beta_t \leq \min\{W_t, w_t\}$, so that the agent never prefers to disclose failure or to fake failure for $t < t_1$.

observed failure for the agent γ_t. Note that $\lambda^F > 0$ may reflect that bad news and failure hidden by insiders can leak out to the public (e.g., via whistleblowing), which for instance was the case for Theranos.

We study the agent's incentives to hide failure and to delay disclosure after she has privately observed failure at time t during the disclosure stage. Disclosure of failure at time t yields β_t dollars. If the agent discloses failure at time $t + dt$, she receives private benefits of shirking ϕdt. With probability $1 - \lambda^F dt$, the principal does not observe (detect) failure over $[t, t + dt)$ so that the agent receives β_{t+dt} dollars from reporting failure at time $t + dt$. With probability $\lambda^F dt$, the principal observes failure over $[t, t + dt)$, leading to rewards for publicly observed failure γ_t. Thus, the agent prefers to disclose failure at time t if and only if

$$\beta_t \geq \phi dt + (1 - \lambda^F dt)\beta_{t+dt} + \lambda^F dt \cdot \gamma_t.$$

Taking $dt \to 0$ and simplifying yields the disclosure incentive constraint

$$\dot{\beta}_t \leq \lambda^F \beta_t - \phi - \lambda^F \gamma_t. \tag{2.24}$$

Note that the right-hand-side of (2.24) decreases with γ_t, reflecting that rewards for (public) failure $\gamma_t > 0$ raise the temptation to hide failure and tighten the incentive compatibility constraint (2.24). Thus, it becomes strictly optimal to stipulate $\gamma_t = 0$ during the disclosure stage. Then, in the limit $\lambda^F \downarrow 0$, the incentive constraint (2.24) reduces to (2.8), and the uniquely (strictly) optimal contract (optimal full disclosure contract) is described in Proposition 3 (Proposition 2).

In addition, Appendix 2.12.3 studies a model variant in which the agent exerts effort a_t against convex (private) costs, whereby $a_t \in [0, 1]$. In general, the principal then maximizes effort incentives by stipulating $\gamma_t = 0$, which leads to high-powered effort incentives during the disclosure stage and pins down the choice of α_t and γ_t. Finally, Appendix 2.11 presents another model variant in which the optimal contract may stipulate $\gamma_t = \beta_t$. In some applications, $\gamma_t = \beta_t$ might ease the interpretation of the contract quantities, for instance, when rewards for failure are interpreted as golden parachutes that are activated both when failure is revealed or disclosed.

2.4 Analysis

2.4.1 Applications

Venture capital financing. In this context, the principal represents the venture capitalist (VC) and the agent represents the enterpreneur(s) or founder(s) of the startup financed by the venture capitalist. Consistent with the empirical findings of Kaplan and Strömberg (2003, 2004), the model implies that optimal venture capital contracts involve distinct financing stages, whereby the entrepreneur's incentives and compensation are backloaded over time. In addition, the agent's rewards for failure can be broadly interpreted as liquidiation rights in the startup. Thus, our findings imply that entrepreneurs' liquidation rights should be backloaded, in that entrepreneurs should have no liquidation rights in the early financing stages. While existing models of dynamic contracting feature several stages (e.g., Green and Taylor (2016)), the unique feature of our model is that the contract stages and their length arise endogenously to provide efficient incentives for disclosure of failure and project development. Thus, our model has novel implications for the optimal length and design of these stages and staged financing (as discussed in Section 2.4.2).

Any financing stage concludes with a deadline, but depending on the financing stage a different type of deadline is optimal. At a soft deadline, the entrepreneur makes a progress report and receives financing for the next stage if and only if the progress report reveals that the project is still profitable to pursue. At a hard deadline, financing is terminated regardless of whether the project is still profitable to pursue. The model highlights that these different types of financing deadlines are crucial to incentivize disclosure of failure and efficient project termination, which is different in Green and Taylor (2016) and Varas (2017) where deadlines are needed to incentivize project completion and effort.

Moreover, the model predicts an upward jump in startup valuation (i.e., project valuation) at the beginning of a new financing stage, as new information is released and uncertainty is resolved through a progress report. Figure 2.2 illustrates that the principal's value function F_t^*, representing the valuation of the project (startup), jumps at time t_1 either up, if the agent does not disclose failure at t_1, or jumps down to zero, if the agent discloses failure at t_1. In addition, note that the entrepreneur's dollar rewards for success decrease during a given financing stage yet increase whenever a new financing stage begins. That is, the entrepreneur is effectively rewarded for reaching a new financing stage. Importantly, both the jumps in valuation and rewards for failure and success arise due to endogenous information revelation, a feature that is typically not present in dynamic models of optimal contracting.

R&D financing. Our model also has implications for optimal R&D financing within more general types of firms as well as for the design of R&D projects and compensation contracts of R&D workers (insiders). In particular, optimal financing of R&D projects involves several stages, whereby insiders make occasional progress report and the continuation of financing is contingent on the outcomes of these reports. Crucially, it is optimal for the financier to elicit less frequent progress reports at the early stages of project development and more frequent progress reports at later stages. Over time, the provision of financing becomes more sensitive to reported progress and performance.

The compensation of R&D workers should involve high rewards for success, penalties for delays in project completion, and tolerance for failure via both i) pay for failure and ii) a grace period. That is, R&D workers should have an initial probationary (grace) period that corresponds to the unconditional financing stage in the contract and their incentives should be backloaded. Importantly, R&D workers should be rewarded and paid for failure at later stages, but not for failure at early stages. Note that backloaded incentives and rewards for failure are optimal to jointly optimize the agent's incentives for project development and disclosure of failure. This mechanism leading to backloaded incentives/compensation is different to the one in Sannikov (2014) or Marinovic and Varas (2019a) where backloaded compensation is necessary to provide incentives for actions with persistent (i.e., long-run) impact on firm performance.

Related to our paper, Manso (2011) finds that to incentivize experimentation (which is crucial for R&D), it may be optimal to reward failure. Note that the mechanism leading to rewards for failure in our model is different from the one in Manso (2011), as rewards for failure are needed to incentivize disclosure for failure. In addition, considering a dynamic framework, we obtain the novel findings that i) it is optimal to only reward failure that occurs at a later stage and ii) if there are rewards for failure they should decrease over time to incentivize the agent not to delay disclosure of failure.

Executive compensation contracts. Alternatively, one can interpret the agent as the manager (i.e., CEO) of a firm while the principal represents the firm's investors or outside shareholders. Golden parachutes, golden handshakes, and severance pay are instruments that induce tolerance for failure in executive compensation contracts (Edmans and Gabaix (2016)). Note that in our setting, the optimal contract implements tolerance towards failure through both i) pay (rewards) for failure and ii) a probationary or grace period (that corresponds to the unconditional financing stage) during which the CEO is not fired after bad outcomes. Importantly, our model highlights that while instruments like golden parachutes, golden handshakes, and severance pay incentivize disclosure of failure, they also provide incentives to fake or generate bad outcomes on purpose or not to work hard. This may lead to a tension in incentive provision similar to the one in Inderst and Mueller (2010) where severance pay may be needed to incentivize a "bad" CEO to quit but at the same time severance pay undermines incentives for effort. The novel feature in our model is that because the contract does not always incentivize disclosure and consists of an unconditional financing stage followed by a disclosure stage, our results imply golden parachutes (golden handshakes) and severance pay should not be in place in the early stages of the CEO's contract (i.e., CEO tenure) but only later on. Once in place, the dollar size of these instruments should decrease over time and the CEO should be provided strong incentives through performance pay.

2.4.2 Project characteristics and financing contracts

We study how project characteristics shape the design of optimal financing contracts. In particular, we analyze the optimal length of the different stages and the optimal provision of unconditional financing, which are unique features of our model. Figure 2.3 plots the financing deadline T, the dollar reward for failure at the beginning of the disclosure stage β_{t_1}, and the relative length of the unconditional financing stage t_1/T against $1/\Lambda$ (upper three panels) and ϕ (lower three panels). Observe that β_{t_1} is the highest reward for failure attainable and proxies overall/average dollar rewards for failure within the optimal contract.

Notably, the financing deadline T is hump-shaped in the (average) duration of the project development phase $1/\Lambda$, as shown in Corollary 1 and illustrated in Figure 2.3. The intuition is as follows. Projects with a long development phase (i.e., projects characterized by a high value of $1/\Lambda$) naturally require financing over a longer horizon. However, projects characterized by a high value of $1/\Lambda$ are at the same time subject to more severe moral hazard, which makes financing such projects over a long horizon less attractive.

Figure 2.3 also illustrates that projects, generating (preliminary) results quickly and characterized by a low value of $1/\Lambda$, are more suitable (or likely) to receive unconditional financing early on. The reason is that these projects are financed with a short deadline, which curbs the temptation to hide bad outcomes and hence facilitates the provision of unconditional financing. As a result, the model implies that for projects characterized by a low value of $1/\Lambda$, the provision of financing is less sensitive to reported progress yet subject to a strict financing deadline. In the context of venture capital financing, these findings imply that venture capitalists provide relatively more unconditional financing for projects that generate results quickly. In addition, dollar rewards for failure (as captured by β_{t_1}) tend to be higher for moderate values of $1/\Lambda$. In particular, optimal financing contracts for projects that generate results quickly (i.e., with low $1/\Lambda$) stipulate low dollar rewards for failure, but

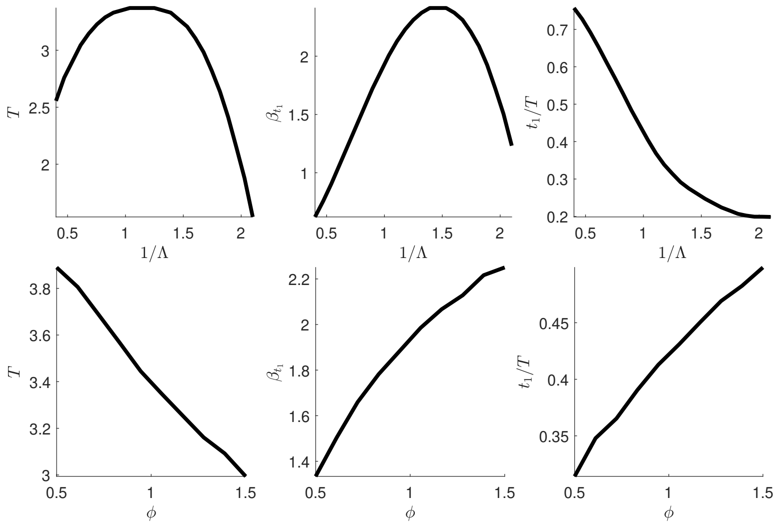

FIGURE 2.3: Comparative statics w.r.t. length of project development phase $1/\Lambda$ and the agency parameter ϕ. The baseline parameters are $\mu = 50$, $\kappa = 10$, $p = 0.5$, $\pi = 0.5$, and $\phi = 1$.

exhibit substantial tolerance towards failure through the provision of unconditional financing or through a long probationary (grace) period for the agent.

Last, the financing deadline T decreases in ϕ, while rewards for failure β_{t_1} and the relative length of the unconditional financing stage t_1/T increase in ϕ. As a result, optimal financing contracts for projects that are subject to severe agency conflicts involve a relatively long unconditional financing stage (and soft deadline) that is followed by a relatively short disclosure stage (and hard deadline). This finding also implies that the provision of financing is less performance-sensitive, when agency conflicts are more severe. The intuition is that because the provision of unconditional (not performance-sensitive) financing limits agency rents, unconditional financing is especially valuable when agency conflicts are severe and ϕ is high. In other words, our model generates the novel and at first glance counter-intuitive result that because self-reported progress is subject to moral hazard, it is optimal to provide substantial unconditional financing that is independent of performance or reported progress, precisely when agency conflicts in project development are severe (e.g., due to high intangibility or complexity of the project). We conclude this section with the following corollary that provides analytical results regarding the effects discussed above.

Corollary 1 *Let $\pi > 0$ be sufficiently small. Then, the following holds:*

1. *T and β_{t_1} increase in $1/\Lambda$ for $1/\Lambda$ sufficiently small and decrease in $1/\Lambda$ for $1/\Lambda$ sufficiently large.*

2. *T decreases in ϕ and β_{t_1} decreases in ϕ for ϕ sufficiently large.*

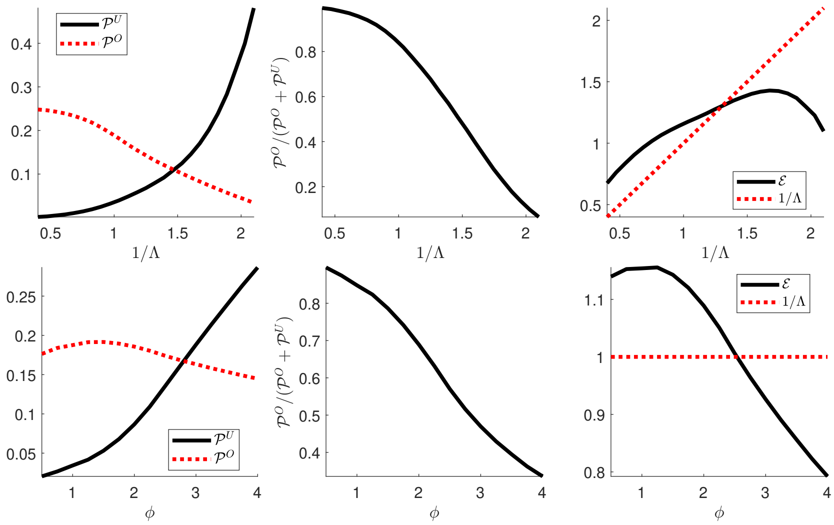

FIGURE 2.4: Over- vs. under-investment. The baseline parameters are $\mu = 50$, $\kappa = 10$, $p = \pi = 0.5$, and $\phi = 1$.

2.4.3 Over- and under-provision of financing

In our model, moral hazard can lead to under-provision of financing (i.e., under-investment), when contract termination during the disclosure stage before time τ precludes financing of a project with positive NPV. Moral hazard can also lead to over-provision of financing (i.e., over-investment), which refers to financing a project with negative NPV. Over-provision of financing may arise when the principal inefficiently extends financing to a failed project (with negative NPV) during the unconditional financing stage.

To assess financing efficiency, we calculate the ex-ante probability of under-investment

$$\mathcal{P}^U := \mathbb{P}(T < \tau) = e^{-\Lambda T},$$

i.e., the likelihood that the project is terminated before completion. Likewise, we calculate the ex-ante probability of over-investment

$$\mathcal{P}^O = \mathbb{P}\big(\text{Unobserved Failure during } [0, t_1)\big) = (1 - e^{-\Lambda t_1})(1 - p)(1 - \pi).$$

Another measure of financing efficiency is the expected length of the financing period $\mathcal{E} := \mathbb{E}[T \wedge \tau^A]$, which captures the principal's investment horizon. This quantity equals $1/\Lambda$ in the second-best case and without frictions. Hence, $\mathcal{E} < 1/\Lambda$ indicates under-provision of financing and $\mathcal{E} > 1/\Lambda$ indicates over-provision of financing. Appendix 2.12.4 shows how to calculate $\mathcal{E}$. Note that we have optimal financing (i.e., $\mathcal{P}^O = \mathcal{P}^U = \mathcal{E} - 1/\Lambda = 0$) in the second-best environment (when $\pi = 1$) but under-investment (i.e., $0 = \mathcal{P}^O < \mathcal{P}^U$ and $\mathcal{E} < 1/\Lambda$) under a full disclosure contract. Conversely, the optimal contract with distinct financing stages or, broadly interpreted, stage financing as such is more likely to cause over-investment (i.e., $\mathcal{P}^O > 0$ and $\mathcal{E} > 1/\Lambda$). Note that over-investment in our model is an extreme case in that it corresponds to financing a project that does not produce a payoff at all.

The reason is that we have normalized payoffs after failure to zero. Section 2.5 considers a model variant in which the project may generate payoff, when it receives financing despite failure (which might be inefficient). Under these circumstances, over-investment implies that the project generates more payoffs than is efficient according to the NPV criterion.

Figure 2.4 displays the ex-ante probabilities of under- and over-investment $\mathcal{P}^U$ and $\mathcal{P}^O$, the probability of over-investment conditional on under- or over-investment $\mathcal{P}^O/(\mathcal{P}^O + \mathcal{P}^U)$, and the average length of the financing period $\mathcal{E}$ in dependence of both $1/\Lambda$ (upper three panels) and ϕ (lower three panels). Figure 2.4 shows that, on average, there is under-provision of financing for projects that do not generate results quickly (i.e., projects with a long development phase and high $1/\Lambda$), whereas there is over-provision of financing for projects that generate results quickly (i.e., projects with a short development phase and low $1/\Lambda$). That is, investors tend to terminate financing for long-term (short-term) projects inefficiently early (late). In other words, moral hazard increases the principal's investment horizon (i.e., financing deadline) for short-term projects but decreases the investment horizon for long-term projects, relative to the benchmark without frictions. Likewise, Figure 2.4 also illustrates that mild agency conflicts induce over-provision of financing (i.e., over-investment) but severe agency conflicts induce under-provision of financing (i.e., under-investment).

In the context of venture capital financing, the model predicts inefficiently high venture capital investment in projects that generate results quickly (i.e., projects with low $1/\Lambda$) — such as information technology projects in which the cost of experimentation is generally low (Kerr et al. (2014)) — and inefficiently low venture capital investment in projects that do not generate results quickly — such as renewable energy production in which the cost of experimentation is generally high. These findings are consistent with those of Nanda, Younge, and Fleming (2014) who document relatively low venture capital investment in renewable energy production. Also note that the parameter ϕ, capturing the severity of agency conflicts, may be (inversely) related to the venture capitalist's oversight in the project or expertise in financing these types of projects, in that both oversight and expertise are likely to reduce agency conflicts. As such, the model predicts inefficiently high (low) venture capital investment, when the venture capitalist's expertise or oversight is strong (weak) and ϕ is low (high).

Recall that according to Figure 2.3, projects generating results quickly are financed with a relatively long unconditional financing stage followed by a short disclosure stage, leading to a low value of T and a high value of t_1/T. Note that this contract comes close to a "simple" contract in which the principal provides financing up to a certain deadline T without providing incentives for disclosure of failure so that $t_1 \simeq T$. Under these circumstances, the principal is likely to end up financing a failed project, which is over-investment. Importantly, venture capitalists' over-investment and excessive provision of unconditional financing (with a short financing horizon and limited involvement in governance) for projects generating results quickly is broadly consistent with a "spray and pray" investment approach adopted by venture capitalists for these types of projects (Ewens et al. (2018)). Different to the theory model in Ewens et al. (2018), the mechanism leading to this result in our framework relies on optimal incentive provision for information disclosure rather than experimentation in venture capital investments. However, according to our model, this "spray and pray" investment approach leads to over-investment and is inefficient from a social perspective.

2.5 Financing unicorns and project risk

We relax the assumption that a failed project does not generate payoff. Consider that a failed project may also generate (terminal) payoff μ. However, unlike a successful project, which generates payoff μ immediately at time τ, a failed project generates payoff μ with delay at time $\tau^\lambda > \tau$. Terminal payoff μ is observable and contractible. As in the baseline model, failure at time τ is observed by the agent but observed by the principal only with probability $\pi \in [0,1]$. Thus, the principal may not be able to distinguish between whether it is immediate success at time τ or "failure" and later "success" at time τ^λ that has led to the terminal payoff. Therefore, the agent is paid α_t dollars when payoff μ realizes at time t (and failure has not been reported by the agent or observed by the principal prior to time t). In addition, the principal cannot verify failure or reports thereof. Note that this setting is relevant when the terminal payoff represents a certain goal or milestone in project development, whereby failure captures interim bad outcomes that are hard for the principal to observe or verify. Accordingly, this setting can also be interpreted as a multi-stage project.

When the principal finances the project after its failure, the time τ^λ arrives at exogeneous rate $\lambda > 0$. That is, over $[\tau, \tau^\lambda)$ the principal still incurs financing costs κ so that delay is costly and the payoff of a failed project is lower than the payoff of a successful project.[17] On average, a failed project takes $1/\lambda$ units of time to generate payoff and, therefore, possesses at time τ *after completion* value (i.e., NPV):

$$\mu^f := \mu - \frac{\kappa}{\lambda},$$

while the value of a successful project equals μ. The NPV of the project *before completion* is

$$NPV^\lambda = p\mu + (1-p)\max\{\mu^f, 0\} - \frac{\kappa}{\Lambda},$$

whereby the principal ideally would like to terminate financing upon failure when $\mu^f < 0$. Importantly, the agent derives flow benefits ϕ as long as the principal finances the project. Thus, when the principal finances the project over the time interval $[\tau, \tau^\lambda)$ after failure, the agent incurs flow benefits ϕ, which undermines his incentives to exert effort. We assume that parameters are such that the principal optimally implements full effort and terminates financing once failure becomes known to her (e.g. via a report or via observed failure).

To characterize the agent's incentives to disclose failure, suppose the project has failed at time t. Reporting failure immediately yields payoff β_t. Delaying disclosure for dt units of time not only yields benefits from running the project ϕdt but also entails the chance, i.e., λdt, that the failed project generates payoff μ, leading to a reward α_t. Hence, the agent is better off disclosing failure at time t if and only if

$$\beta_t \geq \phi dt + \beta_{t+dt}(1 - \lambda dt) + \alpha_t \lambda dt,$$

which becomes for $dt \to 0$:

$$\dot{\beta}_t \leq -(\phi + \lambda(\alpha_t - \beta_t)) \tag{2.25}$$

Notably, the incentive condition w.r.t. effort (2.11) implies $\alpha_t > \beta_t$ and hence an increase in λ tightens the incentive condition (2.25). Intuitively, the prospect of a

[17] The cost of delay may alternatively arise due to displacement risk from other technologies or due to investors' time preference.

future breakthrough despite (interim) failure provides the agent with incentives to hide failure to avert termination and to continue project development to gamble on later success under the motto "Fake it till you make it." Appendix 2.11 presents a more detailed discussion of this model specification and presents the formal solution.

A measure of project risk is given by the variance of potential project outcomes, i.e., the difference in project value after success and failure $\mu - \mu^f = \kappa/\lambda$. This measure of project risk clearly decreases in λ (and does not depend on μ). An implication is that more ambitious or risky projects (characterized by a low value of λ or a high value of $1/\lambda$) are easier to incentivize and hence easier for the principal to finance. Formally:

Corollary 2 *Suppose $\pi > 0$ and $\lambda > 0$ are sufficiently small. A mean preserving spread, increasing μ but decreasing λ while holding NPV^λ fixed, increases the principal's payoff.*

The underlying reason is that higher $1/\lambda$ (lower λ) relaxes the incentive constraint (2.25). When the project is sufficiently risky or ambitious as charcterized by a high value of $1/\lambda$, the agent is less inclined to continue operating the project in the hope of reaching a breakthrough in the future, which improves incentives to disclose failure. As a result, our model predicts that venture capitalists seek to finance very ambitious and risky start ups, i.e., potential unicorns, even if this choice is not necessarily supported by the NPV criterion. Interestingly, Gornall and Strebulaev (2020) and Gompers, Gornall, Kaplan, and Strebulaev (2020) find evidence that unicorn startups are frequently over-valued. This is broadly consistent with the notion that venture capital investors seek investments with high potential but high risk, thereby boosting the valuation for such startup firms (possibly above the fundamental value). In other words, our model predicts that risk-taking and risky investments reduce agency costs and therefore might be optimal for venture capitalists. This implication is broadly consistent with the findings of Nanda and Rhodes-Kropf (2013) who document venture capitalists' risk-taking in "hot markets."

2.6 Optimal contracts with monitoring and inspections

In this section, we introduce the possibility that the principal can inspect (i.e., monitor) project progress at a cost $K > 0$. Upon an inspection at time t, the principal learns without error whether the project has failed so far. Thus, monitoring takes the form of a costly state verification in the spirit of Townsend (1979). The outcome of an inspection is public knowledge and contractible. But, the principal cannot commit to an inspection policy, in that inspection dates are not contractible. In the context of venture capital financing, the assumption that inspections are not contractible is similar to assuming that a venture capitalist's effort or advising is not contractible, as, e.g., in Casamatta (2003) or Schmidt (2003). Thus, the principal conducts an inspection at time t only if she finds it privately optimal to do so. With F_t denoting the principal's payoff when there is no inspection at time t and $\hat{F}_t$ denoting the principal's payoff after an inspection at time t yields that the project has not failed, the principal finds it privately optimal to conduct an inspection at time t if and only if $\hat{F}_t(1 - q_t) - K \geq F_t$.

We study how inspections generate incentives to disclose failure truthfully, i.e., truth telling incentives. For this sake, let us consider that the contract incentivizes disclosure of failure so that $w_t = \beta_t$. When the principal inspects the project and learns that the agent is hiding failure, she can punish the agent. The threat of punishment generates truth telling incentives. Because of the agent's limited liability, these truth telling incentives are maximized, when the principal terminates financing and fires the agent upon detecting misbehavior. Suppose that the project has failed at time t and that failure is privately observed by the agent. The agent can hide failure maximally up to the next inspection date τ_t^M after time t. Notably, we conjecture and verify that the principal inspects the project at deterministic dates whenever w reaches zero, in that $\tau_t^M = \inf\{s \geq t : w_s = 0\}$ is deterministic. Anticipating the next inspection at the (deterministic) date τ_t^M, the agent prefers to disclose failure truthfully at time t if and only if $\beta_t \geq \beta_s + (s-t)\phi$ for all $s \in (t, \tau_t^M)$. Taking the limit $s \to t$ yields $\dot{\beta}_t \leq -\phi$, i.e., (2.8). As failure can potentially occur at any time t, the incentive condition (2.8) must hold for all times t at which there is no inspection. That is, the principal provides incentives to disclose failure both through inspections at deterministic dates and through time-decreasing rewards for failure at all other dates. Also, contract terms *after* time τ_t^M do not affect the agent's incentives to disclose failure *before* time τ_t^M.

In addition, the principal incentivizes the agent not to fake failure. As she cannot commit to inspection dates, the principal cannot do so by inspecting the project after a failure report. When the agent anticipates that the principal inspects the project after a failure report (with some probability) and therefore has incentives not to fake failure, any reported failure is true. Thus, any inspection after a failure report yields that the project indeed has failed, leading to termination. However, given that reported failure is true, the principal would like to deviate by not inspecting the project after a failure report to save monitoring costs.

Whenever there is no inspection, w drifts deterministically down, in that $\dot{w} = -\phi$. When the principal incentivizes disclosure of failure, $\beta = w$, leading to time-decreasing rewards for failure. Once w reaches zero, the principal cannot provide sufficient incentives anymore to the agent who is protected by limited liability. As such, the principal principal must either terminate financing or inspect the project. Note that anticipating an inspection or termination, the agent has no incentives to hide failure just before this event. As a result, when there is an inspection at $w = 0$, the inspection yields that the project has not failed yet. Therefore, after an inspection, the principal faces the same decision problems as at time zero, so that w is reset to w_0 and the principal's continuation payoff becomes F_0. Note that the principal finds it (privately) optimal to inspect the project at $w = 0$ if and only if $F_0 \geq K$; otherwise, the principal terminates the project. Thus, the principal's payoff at $w = 0$ just before an inspection is $\max\{F_0 - K, 0\}$. Unless otherwise mentioned, we assume that monitoring is optimal, i.e., $F_0 \geq K$. The case where $F_0 < K$ is described in the baseline where monitoring is not possible. The following Proposition summarizes the findings of this section.

Proposition 4 *The principal inspects the project at all times t with $w_t = 0$ if and only if $F_0 \geq K$. Otherwise, if and only if $F_0 < K$, the principal never inspects the project, and terminates the project once w reaches zero. After an inspection, w is reset to w_0 and the principal's continuation payoff becomes F_0. Inspections occur at deterministic dates, there are no inspections when $w > 0$, and absent an inspection, w drifts down with $\dot{w} = -\phi$. The optimal contract incentivizes disclosure of failure whenever $w \in [0, w^c]$ and does not incentivize disclosure of failure whenever $w \in (w^c, w_0]$ for $w^c \in [0, w_0]$. The contract always incentivizes disclosure of failure and $w^c = w_0$, if and only if $W_0 \geq w_0$, which is the case when $K > 0$ is sufficiently small. The contract does not always incentivize disclosure of failure and $w^c < w_0$, if and only if $W_0 < w_0$, which is the case when $K > 0$ is sufficiently large.*

In summary, when $K \leq F_0$, the optimal contract features several (possibly degenerate) unconditional financing stages (i.e., $(w^c, w_0]$) and disclosure stages (i.e., $[0, w^c]$), whereby periodic inspections occur at deterministic dates and an unconditional financing stage is followed by a disclosure stage. If the project is not completed by the end of a disclosure stage when $w = 0$, the principal inspects the project and grants financing for the next unconditional financing stage if and only if an inspection reveals that the project is still profitable to pursue. Then, w is reset to w_0, so that the agent is effectively rewarded for positive inspection outcomes.

Notably, with inspections at $w = 0$, there is no financing deadline, and project development is not terminated before completion. Thus, punishments for delays — such as contract termination — and monitoring are substitutes for the provision of truth telling incentives. That is, the principal can now discipline the agent not to hide failure forever by occasionally inspecting the project instead of terminating the project. The model therefore predicts that venture capitalists with high expertise and ability to oversee or inspect a startup's operations tend to provide financing over a longer horizon. Removing the threat of termination also increases the agent's continuation payoff (i.e., the agent's stake) W_t, thereby incentivizing the agent not to fake failure and boosting incentives for project development. The additional incentives for project development brought about by inspections reduce the need for excessively high rewards for success to provide incentives for disclosure, which lowers the cost of incentivizing disclosure of failure. As a consequence, the optimal contract with monitoring is a full disclosure contract provided the monitoring costs K are sufficiently small in which case inspections occur frequently. We numerically calculate that under our baseline parameters, a full disclosure contract is optimal when $K \leq 0.899 \equiv \overline{K}$, in which case the unconditional financing stages $(w^c, w_0]$ vanish (i.e., $w^c = w_0$).

Conversely, when the monitoring costs K are large and satisfy $K \in (\overline{K}, F_0]$, the optimal contract does not always incentivize disclosure of failure so that $w^c < w_0$. Then, the optimal contract consists indeed of several unconditional financing stages and disclosure stages, whereby the principal inspects the project at the end of a disclosure stage and elicits a truthful progress report at the end of an unconditional financing stage. Finally, when the monitoring costs are so large that monitoring is no more optimal (i.e., when $K > F_0$), then the optimal contract is described in Proposition 3 and does not always incentivize disclosure of failure.

2.7 Further results and robustness

2.7.1 Moral hazard versus adverse selection

In our model, *moral hazard* arises because effort is hidden and costly. The severity of moral hazard is captured by the agent's private benefits from shirking ϕ. In addition, failure is hard for the principal to observe or verify. Imperfectly observable failure induces another agency problem that can be interpreted as an *adverse selection* problem and its severity is captured by $1 - \pi$. The reason is that the principal provides truth telling incentives to the agent under the assumption that the agent (already) has privately observed project failure.

Proposition 3 shows that $t_1 \to 0$ as $\pi \to 0$ and $t_1 \to T$ as $\pi \to 1$ and Figure 2.3 shows that t_1/T increases in ϕ. That is, the provision of unconditional financing is valuable when moral hazard is severe (i.e., when ϕ is large) but adverse selection concerns are mild (i.e., when $1 - \pi$ is low). In other words, the provision of unconditional financing is suitable for dealing with moral hazard but less suitable for

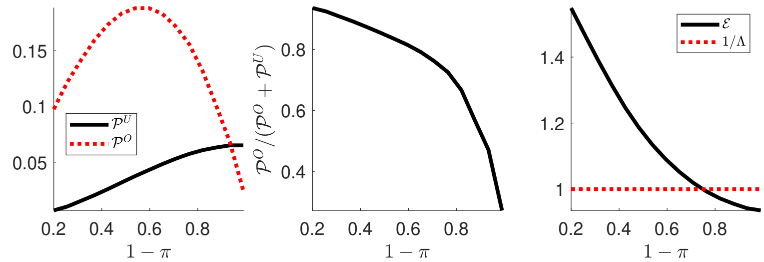

FIGURE 2.5: Over- vs. under-investment. The baseline parameters are $\mu = 50$, $\kappa = 10$, $p = 0.5$, and $\phi = 1$.

dealing with adverse selection. As a result, the provision of financing is more (less) performance sensitive when adverse selection (moral hazard) is more severe.

The lower three panels of Figure 2.4 plot the (scaled) probabilities of over- and under-investment and the average length of the financing period against ϕ, capturing the severity of moral hazard, while Figure 2.5 plots the same quantities against $1-\pi$, capturing the severity of adverse selection. Figures 2.4 and 2.5 demonstrate that mild adverse selection concerns (moral hazard problems) induce over-provision of financing (i.e., over-investment) whereas severe adverse selection concerns (moral hazard problems) induce under-provision of financing (i.e., under-investment).

2.7.2 Robustness

Our model entails a number of assumptions that are mainly designed to enhance simplicity and to facilitate a clear analysis of the main forces in a tractable model. Below, we discuss these assumptions and the robustness of the results.

Exogenous project completion. Existing dynamic contracting papers — such as Mason and Välimäki (2015), Green and Taylor (2016), and Varas (2017) — study how to motivate the agent to complete a project. In these papers, completion always corresponds to “success" and the only bad outcomes that can arise are completion delays. Notably, the agent is never tempted to hide project completion/success, but — in Green and Taylor (2016) and Varas (2017) — the agent is tempted to fake success (i.e., good outcomes). The innovation in our paper is that it considers the risk of project failure when the agent can hide or fake project failure (i.e., bad outcomes).

In our model, the assumption of an exogenous completion rate is made for simplicity and theoretical clarity and is not consequential. It offers the advantage that we obtain a clean second-best benchmark when failure is observable and contractible. Departing from this benchmark, we are able to clearly identify how imperfect observability of failure affects incentive provision. Appendix 2.12.1 shows that we obtain the same results employing an alternative framework, in which the agent controls project completion while the project is subject to failure risk. Likewise, one could extend our baseline model to a model of multi-tasking in which the agent controls both project completion and the project’s propensity to succeed or fail.

Unobservable success. Throughout this paper, we have assumed that success is perfectly observable and contractible. This assumption intuitively reflects that the agent is tempted to conceal bad outcomes rather than good outcomes. Even though not modelled explicitly, disclosure of good outcomes could yield private benefits to the agent, e.g., related to the agent's reputation or career concerns, whereas disclosure of bad outcomes could yield private dis-utility.

We show that our findings do not change substantially when success is imperfectly observable to the principal and the agent cannot fake success. That is, the agent can hide success (i.e., delay disclosure of success and report success at a later time than at which it has occurred), but the agent cannot mis-report success before it occurs. We consider that both failure and success are (publicly) observed by the principal only with probability π and privately observed by the agent otherwise. To obtain a non-trivial solution, at least one of the two possible outcomes, success and failure, must be verifiable. We assume that success is verifiable, as it is likely to be more difficult to fake good outcomes rather than bad outcomes.

To characterize the agent's incentives to disclose success, note that the agent can always delay disclosure for a unit of time and derive private benefits ϕdt. By the same arguments leading to (2.8), we obtain that the agent prefers to disclose success truthfully if and only if $\dot{\alpha}_t \leq -\phi$. This condition is obviously not satisfied in the optimal contract from Proposition 3 because α_t exhibits a jump at time t_1 when the unconditional financing stage ends. However, during the disclosure stage and within an (optimal) full disclosure contract, it follows that $\dot{\alpha}_t \leq -\phi$. This implies that the disclosure stage of the optimal contract is unaffected by whether success is observable.[18]

Appendix 2.12.2 extends this intuition and demonstrates that the optimal contract does not change much when success is imperfectly observable: it features i) an unconditional financing stage $[0, t_1)$ during which the agent discloses neither success nor failure (yet is rewarded if success is observed), and ii) a disclosure stage that looks similar to the disclosure stage of Proposition 3.

2.8 Conclusion

We study a dynamic contracting model in which a principal hires an agent to develop an innovative project. Crucially, project failure is hard for the principal to observe or verify and the agent can hide or fake failure, leading to a tension in incentive provision. The optimal contract consists of two distinct stages: i) an unconditional financing stage and ii) a disclosure stage. During the unconditional financing stage, financing is guaranteed and the contract does not incentivize disclosure of failure. During the disclosure stage, the contract incentivizes disclosure of failure and the principal finances the project until either a deadline is reached or the project is completed. Then, the agent receives time decreasing rewards for success and failure and incurs harsh punishments for delays. In the optimal contract, the provision of financing becomes more performance sensitive following completion delays, i.e., following poor performance. Our results also imply that moral hazard may lead to over- or under-provision of financing relative to the net present value criterion. Last, we characterize optimal dynamic monitoring and study the role of monitoring in incentive provision. The paper generates a set of implications for venture capital financing, R&D financing, and the design of executive compensation contracts.

[18] Recall that — by Propositions 2 and 3 — for $t \geq t_1$: $\alpha_t = (1 - \pi + \pi/p)w_t + \phi/(\Lambda p)$, so that $\dot{\alpha}_t \leq \dot{w}_t = -\phi$.

Appendix

2.9 Closed form expressions

Lemma 1 *The ODE* (2.14) *subject to* $F(0) = 0$ *and with* (2.15) *and* $\dot{w} = -\phi$ *has the closed form solution* (2.16). *In addition, the value* w_0 *from* (2.17) *satisfies* $F'(w_0) = 0$.

We just verify that the proposed function indeed is the desired solution to the ODE (2.14) subject to the stipulated boundary conditions $F(0) = F'(w_0) = 0$.

Take (2.16):

$$F(w) = \left(\mu p - \frac{\kappa}{\Lambda}\right)\left(1 - \exp\left(\frac{-w\Lambda}{\phi}\right)\right) - w$$

and differentiate to obtain

$$F'(w) = \left(\mu p - \frac{\kappa}{\Lambda}\right)\frac{\Lambda}{\phi}\exp\left(\frac{-w\Lambda}{\phi}\right) - 1.$$

Define for any function $F(w)$ the operator

$$\mathcal{D}F(w) = \Lambda F(w) + F'(w)\phi - (\Lambda p\mu - \kappa - \Lambda w - \phi),$$

and note that $\mathcal{D}F(w) = 0$ if and only if $F(w)$ solves (2.14) under the optimal controls $\dot{w} = -\phi$ and $\alpha(w) = (1 - \pi + \pi/pi)w + \phi/(\Lambda p)$ (i.e., (2.15)). We use above expressions for $F(w)$ and $F'(w)$ and obtain

$$\begin{aligned}\mathcal{D}F(w) =& \Lambda\left[\left(\mu p - \frac{\kappa}{\Lambda}\right)\left(1 - \exp\left(\frac{-w\Lambda}{\phi}\right)\right) - w\right] \\ &-(\Lambda p\mu - \kappa - \Lambda w - \phi) - \phi\left[\left(\mu p - \frac{\kappa}{\Lambda}\right)\frac{\Lambda}{\phi}\exp\left(\frac{-w\Lambda}{\phi}\right) - 1\right] = 0.\end{aligned}$$

Next, we verify that $F(0) = 0$ and $F'(w_0) = 0$ with w_0 from (2.17). It is immediate to see that $F(0) = 0$. Next recall that

$$w_0 = \frac{\phi}{\Lambda}\ln\left(\frac{\Lambda\mu p - \kappa}{\phi}\right),$$

so that

$$F'(w_0) = \left(\mu p - \frac{\kappa}{\Lambda}\right)\frac{\Lambda}{\phi}\frac{\phi}{\Lambda\mu p - \kappa} - 1 = 0.$$

The proof is complete by virtue of the Picard-Lindeloef theorem, ensuring uniqueness of the solution.

Lemma 2 *Fix* $w_0 > w_L > 0$. *The ODE* (2.20) *with* $\dot{w} = -\phi$, $\alpha - w(1-\pi) = \phi/\Lambda p$, *and* (2.21) *has the following closed form solution on* $(w_L, w_0]$:

$$f(w) = \left(e^{-B(w_L)}(F(w_L)(1 - q(w_L)) - q(w_L)w_L)\right)e^{B(w)} + e^{B(w)}\int_{w_L}^{w} e^{-B(x)}a(x)dx.$$

with

$$F(w) = \left(\mu p - \frac{\kappa}{\Lambda}\right)\left(1 - \exp\left(\frac{-w\Lambda}{\phi}\right)\right) - w \quad \textit{and}$$
$$B(w) = \frac{-(p+\pi(1-p))}{(1-p)(1-\pi)} e^{-\frac{\Lambda(1-p)(1-\pi)(w_0-w)}{\phi}}$$
$$a(w) = \frac{(1-q(w))\Lambda p(\mu - w(1-\pi)) - (1-q(w))\phi - \kappa}{\phi}$$
$$\textit{and} \quad q(w) = \left(1 - e^{-\frac{\Lambda(1-p)(1-\pi)(w_0-w)}{\phi}}\right)$$

After substituting the optimal controls $\alpha = w(1-\pi) + \phi/(\Lambda p)$ and $\dot{w} = -\phi$, the ODE to solve becomes

$$f'(w) = \frac{1}{\phi}\Big((1-q(w))\Lambda p(\mu - \alpha(w)) - \kappa - \Lambda(p + (1-p)\pi)(1-q(w))f(w)\Big).$$

This is a first order linear ODE of the general form

$$f'(w) = a(w) + b(w)f(w)$$

with

$$a(w) = \frac{(1-q(w))\Lambda p(\mu - \alpha(w)) - \kappa}{\phi}$$
$$= \frac{(1-q(w))\Lambda p(\mu - w(1-\pi)) - (1-q(w))\phi - \kappa}{\phi}$$

and

$$b(w) = \frac{-\Lambda(1-q(w))(p+(1-p)\pi)}{\phi} = \frac{-\Lambda(p+(1-p)\pi)e^{-\frac{\Lambda(1-p)(1-\pi)(w_0-w)}{\phi}}}{\phi}$$

Note that

$$B(w) = \frac{-(p+\pi(1-p))}{(1-p)(1-\pi)} e^{-\frac{\Lambda(1-p)(1-\pi)(w_0-w)}{\phi}}$$

is anti-derivative of $b(w)$ in that $B'(w) = b(w)$. The fundamental theorem of calculus implies that

$$A(w) = \int_{w_L}^{w} e^{-B(x)} a(x)dx$$

is anti-derivative of $e^{-B(w)}a(w)$.

It is well known that the first order linear differential equation of form $f'(w) = a(w) + b(w)f(w)$ admits the general solution

$$f(w) = Ce^{B(w)} + e^{B(w)} \int_{w_L}^{w} e^{-B(x)} a(x)dx$$

with constant C. We solve for C, using the boundary condition

$$f(w_L) = F(w_L)(1 - q(w_L)) - q(w_L)w_L,$$

which yields

$$C = e^{-B(w_L)}(F(w_L)(1-q(w_L)) - q(w_L)w_L),$$

as desired.

2.10 Omitted Proofs

To start with, note that, because $dc_t = 0$ for $t > \tau^A$, wage payments c (i.e., (2.26)) can be rewritten as

$$dc_t = \Big(\alpha_t \mathbf{1}_{\{\text{Success at time } t\}} + \beta_t \mathbf{1}_{\{\text{Failure report at time } t\}} + \gamma_t \mathbf{1}_{\{\text{Failure observed at time } t\}}\Big)\mathbf{1}_{\{t \le \tau^A\}}. \tag{2.26}$$

That is, because the project yields maximally once failure or success, the values of α_t, β_t, and γ_t after time τ^A (i.e., for $t > \tau^A$) are irrelevant. This observation is convenient since we do not (always) have to explicitly distinguish between the two scenarios $\tau^A < t$ and $\tau^A > t$, when describing $\alpha_t, \beta_t, \gamma_t$.

Throughout, we define the agent's expected reward for failure at time t as

$$r_t := (1 - \pi) w_t + \pi \gamma_t, \tag{2.27}$$

where w is defined as in (2.10). To ease the exposition, we refer to "the project fails at time t but failure is not observed by the principal" as "hidden failure". Likewise, "the project fails at time t and failure is observed by the principal" is called "public failure."

Last, as is standard in the literature on optimal contracts, we call a contract $\mathcal{C}$ *incentive compatible* if $\mathcal{C}$ induces full effort (i.e., $a_t = 1$ for $t \le T \wedge \tau$) and truthful disclosure of failure, whenever the principal asks the agent to disclose failure.

2.10.1 Agent's incentive compatibility

Lemma 3 *A contract $\mathcal{C}$ induces truthful disclosure of failure (i.e., $\tau^A = \tau$ with certainty) from time t' onwards if and only if* (2.7) *and* (2.8) *hold for all $t \in [t', T]$ and $T < \infty$. It induces full effort $a_t = 1$ for all $t \in [0, T \wedge \tau]$ if and only if $\alpha_t \ge r_t + \phi/(\Lambda p)$ for all $t \in [0, T \wedge \tau]$.*

Without loss of generality, we consider for the proof $t' = 0$. First, consider any time $t \ge \tau$ and that the project has failed already at time τ and failure has been privately observed by the agent. Then, if the agent has not reported failure yet up to time t, his (continuation) payoff becomes

$$w_t := \max_{\tau^A \in [t,T]} \left[\phi(\tau^A - t) + \beta_{\tau^A}\right], \tag{2.28}$$

given a contract with deadline $T \ge t$. The above expression for w_t is maximized for $\tau^A = t$, only if $\frac{\partial w_t}{\partial \tau^A} = \dot{\beta}_{\tau^A} + \phi \le 0$ for $\tau^A = t$, which is equivalent to (2.8) and a necessary condition for truthful disclosure of failure. Also note that since the project may complete at any time $t \in [0, T]$, truthful disclosure of failure requires $\frac{\partial w_t}{\partial \tau^A}|_{\tau^A = t} \le 0$ for all $t \in [0, T]$. That is, $\dot{\beta}_t = \dot{w}_t \le -\phi$ must hold for all $t \in [0, T]$. After integrating, we obtain $\beta_s \le \beta_t - (s - t)\phi$ for all $s \in [t, T]$. Truthful disclosure of failure also requires that $T < \infty$, as otherwise the agent would hide failure forever (i.e., set $\tau^A = \infty$) and derive infinite payoff from doing so.

On the other hand, if (2.8) holds for any $t \in [0, T]$ with $\beta_0 < \infty$ and $T < \infty$, it follows that $\beta_t \ge \beta_s + (s - t)\phi$ for any $t \le T$ and $s \in [t, T]$, so that w_t is maximized for $\tau^A = t$ for any $t \in [0, T]$ and the contract induces $\tau^A \le \tau$. Hence, $w_t = \beta_t$ for any $t \in [0, T \wedge \tau]$.

Third, take now $t < \tau$, let $r_t = (1-\pi)w_t + \pi\gamma_t$ and note that the agent's continuation payoff reads

$$W_t := \int_t^T e^{-\int_t^s \Lambda(s-t)} \Big(\Lambda\big((1-pa_s)r_s + pa_s\alpha_s\big) + \phi(1-a_s)\Big)ds \tag{2.29}$$

$$= \int_t^T e^{-\int_t^s \Lambda(s-t)} \Big(\Lambda\big(r_s + pa_s(\alpha_s - r_s)\big) + \phi(1-a_s)\Big)ds, \tag{2.30}$$

if he chooses $\tau^A \geq \tau > t$. On the other hand, deviating and reporting $\tau^A = t < \tau$ yields payoff β_t, so that truthful disclosure of failure, i.e., $\tau^A = \tau$ with certainty, requires (2.7) (i.e., $W_t \geq \beta_t$) to hold for any t, with W_t defined in (2.29).

Fourth, note that at any time $t < \tau$, effort $\{a_s\}_{\{s\in[t,T\wedge\tau]\}}$ maximizes W_t (see (2.29)) if and only if it maximizes pointwise (i.e., for all $s \geq t$) the (scaled) integrand of the expression for W_t:

$$\Lambda\big(r_s + pa_s(\alpha_s - r_s)\big) + \phi(1-a_s)$$

for all $s \in [t,T]$. As a result, $a_s = 1$ for all $s \geq t$ and therefore $a_t = 1$ for all $t \in [0, T\wedge\tau]$ if and only if $\alpha_t \geq r_t + \phi/(\Lambda p)$, i.e., with $\beta_t = w_t$, if and only if (2.11) holds for all $t \in [0, T\wedge\tau]$. This holds for any $T \geq 0$, even for $T = \infty$. The proof is now complete.

2.10.2 Proof of Proposition 1

Throughout, due to the observability of failure it follows that $w_t = \beta_t$. The previous Lemma tells us that incentive compatibility requires $\alpha_t \geq \gamma_t + \phi/(\Lambda p)$, where $\gamma_t = r_t$.

For any $t < \tau$, the principal's payoff can be written as

$$F_t = \int_t^T e^{-\Lambda(s-t)} \Big(\Lambda p(\mu - \alpha_s) - \Lambda(1-p)\gamma_s - \kappa\Big)ds.$$

Given a deadline T, the payoff F_t is maximized if $\{\alpha_s, \gamma_s\}_{s\geq t}$ maximize pointwise the integrand, while respecting incentive compatibility and limited liability. Subject to these constraints, the integrand is maximized pointwise for $\gamma_s = 0 < \alpha_s = \phi/(\Lambda p)$, so that the principal's payoff becomes $F_t = \int_t^T e^{-\Lambda(s-t)}(\Lambda p\mu - \phi - \kappa)ds$. Because parameters satisfy $\phi \leq \kappa$ and $\mu p > 2\kappa/\Lambda$, the integrand is positive for any $s \geq t$, so that the principal's payoff is maximized by setting $T = \infty$. Hence, the principal's payoff becomes $\mu p - \frac{\phi+\kappa}{\Lambda}$, while the agent's payoff equals $W_t = \frac{\phi}{\Lambda}$. As a result, the proposed contract is optimal and incentive compatible (i.e., induces full effort).

2.10.3 Proof of Proposition 2

We prove a more general version of Proposition 2.

Proposition 5 *Under the optimal full disclosure contract $\mathcal{C}$, at time t with $w_t = w$, the principal's value is given by* (2.14)*. The contract $\mathcal{C}$ stipulates $\dot{w}_t = -\phi$ and termination at time $T = \inf\{t \geq 0 : w_t = 0\}$. Payments satisfy*

$$\alpha_t - r_t = \frac{\phi}{\Lambda p} + \frac{\pi}{p}(w_t - \gamma_t) \quad \textit{with} \quad \gamma_t \in [0, w_t], \tag{2.31}$$

and $\beta_t = w_t = W_t$. The value w_0 solves $F'(w_0) = 0$.

The claim of Proposition 2 is attained by setting $\gamma_t = 0$. A full disclosure contract $\mathcal{C} = (c, T)$ — by definition — induces $\beta_t = w_t$ for all $t \in [0, T]$. Denote the overall (expected) surplus by $\mathcal{S}$, which is given at time t by

$$\mathcal{S}_t = \mathcal{S}(w_t) = \int_t^T e^{-\Lambda(s-t)}(\Lambda p a_s \mu - \kappa + \phi(1 - a_s))ds = \int_t^T e^{-\Lambda(s-t)}(\Lambda p \mu - \kappa)ds,$$

with $T = \inf\{t \geq 0 : w_t = 0\}$ and full effort $a_s = 1$ for $s \geq t$ in optimum (as shirking is inefficient). The surplus is split between the agent and the principal, so that the principal's payoff $\hat{F}_t$ (at any time t) is given by:

$$\hat{F}_t = \mathcal{S}_t - W_t \leq \mathcal{S}_t - w_t =: F_t = F(w_t),$$

where we used the incentive compatibility condition (2.7), $W_t \geq w_t = \beta_t$. As a result, a full disclosure maximizes the principal's continuation payoff at any time $t \geq 0$ so that $\hat{F}_t = F_t$ and therefore is a optimal full disclosure contract, if it maximizes for any time $t \geq 0$ the continuation surplus $\mathcal{S}_t$ subject to the incentive constraints (2.8), (2.11), subject to the agent's limited liability (leading to $T = \inf\{t \geq 0 : w_t = \beta_t = 0\}$), and achieves $W_t = w_t$.

Notably, for any $t \leq T$, $\mathcal{S}_t$ does not depend α_t and monotonically increases in T. Incentive compatibility requires $\dot{w}_s \leq -\phi$. Limited liability requires $w_s \geq 0$, which yields combined with $\dot{w}_s \leq -\phi$ that $T - t \leq w_t/\phi$ with equality if and only $\dot{w}_s = -\phi$ for all $s \geq t$. That is, setting $\dot{w}_s = -\phi$ for all $s \geq t$ and hence binding the constraint (2.8) maximizes the deadline and continuation surplus $\mathcal{S}_t$ at time t. Note that the proposed contract from Proposition 5 sets $\dot{w}_s = -\phi$ for all $s \geq t$ and therefore maximizes for any t with given value w_t the (time to) deadline T and hence continuation surplus $\mathcal{S}_t$.

Next, for any w_t, the proposed contract from Proposition 5 stipulates payments according to (2.31); that is:

$$\alpha_t - r_t = \frac{\phi}{\Lambda p} + \frac{\pi}{p}(w_t - \gamma_t) \quad \text{and} \quad \gamma_t \in [0, w_t]$$

and, therefore, the incentive constraint w.r.t. effort (2.11) is met for all $t \in [0, T]$. Differentiate (2.29) to obtain

$$\dot{W}_t = \Lambda W_t - \Lambda\big(r_s + p(\alpha_t - r_t)\big) = -\phi,$$

whereby the second equality follows from plugging in (2.31) and $r_t = (1 - \pi)w_t + \pi\gamma_t$. Thus, $\dot{W}_t = \dot{w}_t = -\phi$ and, therefore, due to $W_T = w_T = 0$ it follows that $W_t = w_t$ for all $t \in [0, T]$ (so that (2.7) is met). The starting value w_0 is determined to maximize the principal's ex-ante value F_0. As a result, the proposed contract from Proposition 5 is optimal in the class of full disclosure contracts, i.e., is a optimal full disclosure contract.

The principal's value can be written as

$$F_t = \mathbb{E}_t\left[\int_t^{T\wedge\tau^A}(\mu dS_t - \kappa dt - dc_t)\right] = \int_t^T e^{-\Lambda(s-t)}\Big(\Lambda[p(\mu-\alpha_s)-(1-p)r_s]-\kappa\Big)ds$$
$$= \int_t^T e^{-\Lambda(s-t)}\Big(\Lambda[p(\mu-(\alpha_s-r_s))-r_s]-\kappa\Big)ds$$
$$= \int_t^T e^{-\Lambda(s-t)}\Big(\Lambda(p\mu-w_s)-\phi-\kappa\Big)ds,$$

where the first equality uses integration by parts and the second equality plugs in $\alpha_t - r_t = \frac{\phi}{\Lambda p} + \frac{\pi}{p}(w_t - \gamma_t)$. Hence, the value function F_t indeed does not depend on the exact values of α_t, γ_t as long as (2.31) and $\gamma_t \in [0, w_t]$ hold.

Differentiating for $t \in [0, T)$ yields

$$\dot{F}_t = \Lambda F_t - \Lambda p(\mu - w_t) + \phi + \kappa.$$

One arrives at the ODE (2.14) (under the optimal controls) using

$$\dot{F}_t = \frac{dF_t}{dt} = \frac{dF_t}{dw_t}\frac{dw_t}{dt} = F'(w_t)\dot{w}_t. \tag{2.32}$$

The closed-form solution is provided in Lemma 1 in Appendix 2.9. Clearly, $F(w)$ is concave, so that the first order condition $F'(w_0) = 0$ is sufficient in determining optimal w_0. The proof is now complete.

2.10.4 Proof of Proposition 3

We prove a more general version of Proposition 3.

Proposition 6 *The optimal contract does not incentivize disclosure of failure over some time period $[0, t_1)$ and becomes a full disclosure contract with deadline T, as characterized in Proposition 2, after time t_1.*

1. *With $w_{t_1} = w_L$, the optimal time t_1 is characterized by (2.21) and (2.22), while the value function $F^*(w)$ is characterized by (2.23). In addition, $t_1 \to 0$ as $\pi \to 0$ and $t_1 \to T$ as $\pi \to 1$.*

2. *The contract is terminated at time $T = \inf\{t \geq 0 : w_t = 0\}$ and w_0 (and equivalently T) maximizes the principal's initial payoff $F^*(w_0)$.*

3. *$\alpha_t = w_t(1-\pi) + \phi/(\Lambda p) \geq \beta_t = \gamma_t = 0$, $\dot{w}_t = -\phi < \dot{W}_t < 0$, and $W_t < w_t$ for all $t \in [0, t_1)$.*

4. *$w_t = \beta_t = W_t$ and $\alpha_t - r_t = \phi/(\Lambda p) + \pi/p(w_t - \gamma_t)$, and $\gamma_t \in [0, w_t]$ for all $t \in [t_1, T)$.*

The claim of Proposition 3 is attained by setting $\gamma_t = 0$.

The proof of Proposition 6 involves several steps, that are — for a better overview — separately presented in the following Lemmata in this Section. Importantly, in Steps I through IV of the proof, we conjecture that the optimal contract termination time (i.e., financing deadline) T is deterministic. Then, Lemma 9 in Step V of the proof verifies that the optimal contract termination time T is indeed deterministic in that stochastic termination is not optimal.

We give a brief overview on how we proceed. Step I demonstrates that whenever $W_t \geq w_t$, the optimal contract incentivizes disclosure of failure from time t onward (until termination at time T). Step II shows that a full disclosure contract is not optimal when $\pi > 0$. Step III characterizes the agent's incentives, and the agent's pay for success and failure. In addition, Step III shows that when $\pi > 0$, then the optimal contract involves exactly two stages: i) an unconditional financing stage, in which the agent is not incentivized to disclose failure, followed by a ii) disclosure stage, in which the agent is incentivized to disclose failure. Step IV characterizes the principal's value function. Step V verifies that the optimal financing deadline T is deterministic, in that stochastic termination is not optimal.

Step I

Lemma 4 *At any time $t < \tau$ with $W_t \geq w_t > 0$, the optimal continuation contract is a full disclosure contract with $T - t = w_t/\phi$, $\beta_s = w_s = W_s$, $\alpha_s - r_s = \frac{\phi}{\Lambda p} + \frac{\pi}{p}(w_s - \gamma_s)$, and $\gamma_s \in [0, w_s]$ for $T \wedge \tau > s \geq t$. As a result, within the proposed (optimal) contract, there does not exist a time t with $W_t > w_t$.*

Take any time $t < T \wedge \tau$ with $W_t \geq w_t > 0$. Given a deadline T, the continuation surplus is clearly maximized only if i) the agent does not shirk (i.e., $a_s = 1$ for $s \in [t, T \wedge \tau]$) and ii) financing is terminated once the project completes, in that $\tau = \tau^A$ and $T_0 = T \wedge \tau$. This is because shirking is inefficient, i.e., $\phi \leq \kappa$.

Note that $\dot{w}_s \leq -\phi$ and limited liability, $w_s \geq 0$, imply that $T - t \leq w_t/\phi$, whereby the continuation surplus increases in $T - t$. As a result, the maximum continuation surplus attainable at time t is given by

$$\mathcal{S}_t = \mathcal{S}(w_t) = \int_t^T e^{-\Lambda(s-t)}(\Lambda p \mu - \kappa)ds \quad \text{with} \quad T - t = \frac{w_t}{\phi}.$$

The continuation surplus at time t is split between the agent and the principal so that

$$F_t + W_t \leq \mathcal{S}_t \implies F_t \leq \mathcal{S}_t - W_t. \tag{2.33}$$

A full disclosure continuation contract from t onwards achieves the continuation surplus $\mathcal{S}_t$, since it i) precludes shirking, ii) induces $\tau^A = \tau$ and $T_0 = T \wedge \tau$, and iii) optimally sets $\dot{w}_s = -\phi$ for all $s \geq t$ and hence maximizes the time to deadline $T - t$. That is, (2.33) holds in equality, in that $F_t = \mathcal{S}_t - W_t$.

Note that the principal's payoff F_t only depends on $\{\alpha_s, \beta_s, \gamma_s\}_{s \geq t}$ via

$$W_t = \int_t^T e^{-\Lambda(s-t)}\Lambda(p\alpha_s + (1-p)r_s))ds.$$

Because $W_t \geq w_t$, one way to deliver continuation payoff to the agent is by setting $\alpha_s - r_s \geq \frac{\phi}{\Lambda p} + \frac{\pi}{p}(w_s - \gamma_s)$, $\beta_s = w_s$, and $\gamma_s \in [0, w_s]$ for $s \in [t, T]$, where the inequality holds in equality for all times s with $W_s = w_s$. This is because Proposition 5 shows that within a full disclosure contract, one way to deliver continuation payoff $W_t = w_t$ to the agent is to set $\alpha_s - r_s = \frac{\phi}{\Lambda p} + \frac{\pi}{p}(w_s - \gamma_s)$, $\beta_s = w_s$, and $\gamma_s \in [0, w_s]$ for $s \in [t, T]$. Hence,(weakly) higher pay for success is needed to deliver a (weakly) higher value $W_t \geq w_t$ to the agent within a full disclosure contract. More formally, one can calculate

$$\dot{W}_s - \dot{w}_s = \Lambda W_s - \Lambda(p\alpha_s + (1-p)r_s) + \phi \leq \Lambda W_s - \Lambda r_s = \Lambda(W_s - w_s) + \Lambda \pi w s,$$

where the first inequality uses incentive comptability, $\alpha_s \geq r_s + \phi/(\Lambda p)$, and the second inequality uses $r_s = (1-\pi)w_s + \pi\gamma_s \geq (1-\pi)w_s$. That is, $\dot{W}_s - \dot{w}_s \geq 0$ can be achieved if $W_s \geq w_s$, while respecting effort incentive compatibility $\alpha_s \geq r_s + \phi/(\Lambda p)$. Hence, it follows that pay for success and failure $\{\alpha_s, \gamma_s\}_{s\geq t}$ can be chosen such that $W_s \geq w_s = \beta_s$ for all $s \in [t, T]$, so the agent does not want to fake failure and the incentive condition (2.7) is satisfied. Moreover, due to $\beta_s = w_s$ for all $s \in [t, T]$, it follows that $\dot{w}_s = \dot{\beta}_s = -\phi$, so that (2.8) is met for all $s \in [t, T]$. Thus, the requirement to deliver value W_t to the agent implies all incentive constraints that are relevant to incentivize truthful disclosure of failure and effort. As a result, a full disclosure continuation contract from t onwards maximizes $\mathcal{S}_t$ and — because it achieves (2.33) to hold in equality — the principal's continuation payoff F_t subject to all relevant incentive constraints, limited liability, and the requirement to deliver payoff W_t to the agent.

Let $t = \inf\{s \geq 0 : W_s = w_s\}$. Because the optimal full disclosure continuation contract from t onwards induces $W_s = w_s$ (when $W_t = w_t$) for $s \geq t$ (see Proposition 5 and its proof) and because a full disclosure continuation contract becomes optimal at time t once $W_t = w_t$, follows that within the proposed (optimal) contract there is no time s with $W_s > w_s$. It follows that a full disclosure continuation contract becomes optimal the first time t with $W_t = w_t$, in which case $\alpha_s - r_s = \frac{\phi}{\Lambda p} + \frac{\pi}{p}(w_s - \gamma_s)$, $\gamma_s \in [0, w_s]$, $\dot{w}_s = \dot{W}_s = -\phi$ and $W_s = w_s = \beta_s$ for $s \geq t$.

Step II

Lemma 5 *Let $\pi \in [0,1)$. A full disclosure contract is not optimal if and only if $\pi > 0$. Put differently, if and only if $\pi > 0$, there exists time $t_1 > 0$ such that the optimal contract does not incentivize disclosure of failure over $[0, t_1)$.*

Let $\pi > 0$ and suppose to the contrary that a (optimal) full disclosure contract $\mathcal{C}^0$ with payoff F^0 for the principal is optimal. By Proposition 5 and its proof, this full disclosure contract (optimally) sets $w_t = W_t = \beta_t$, $\dot{\beta}_t = -\phi$, $T = \inf\{t \geq 0 : \beta_t = 0\}$ and $\alpha_t = \beta_t(1 - \pi + \pi/p) + \phi/(\Lambda p)$ and $\gamma_t = 0$ for all $t \in [0, T \wedge \tau]$. This particular choice of $\{\alpha_t, \gamma_t\}_{t\geq 0}$ simplifies the notation.

Take some Δ satisfying $T > \Delta > 0$ and define a contract $\tilde{\mathcal{C}}$ that does not incentivize disclosure of failure for times $t \in [0, \Delta)$ as follows. $\tilde{\mathcal{C}}$ sets rewards for success $\tilde{\alpha}_t = w_t(1-\pi) + \phi/(\Lambda p)$, and stipulates no rewards for (publicly and privately observed) failure for $t < \Delta$. If the project is not completed by time Δ, it switches at time Δ to the (optimal) full disclosure contract in state $w_\Delta = w_0 - \Delta\phi$, with deadline $T = w_\Delta/\phi$, $\beta_t = w_t$, $\dot{w}_t = -\phi$, $\gamma_t = 0$ and $\alpha_t = w_t(1 - \pi + \pi/p) + \phi/(\Lambda p)$ for all $t \in [\Delta, T \wedge \tau]$. Note that the agent optimally reports failure at time Δ, if the project fails before time Δ and failure is privately observed, implying that $w_t = (\Delta - t)\phi + \beta_\Delta$ for $t \leq \Delta$.

We calculate for any $t \in [0, \Delta]$ the likelihood of privately observed failure up to time t, given by (2.19):

$$q_t = \left(1 - e^{-\Lambda(1-p)(1-\pi)t}\right) = q_0 + \dot{q}_0 t + o(t^2) = \Lambda(1-p)(1-\pi)t + o(\Delta^2),$$

where we used a Taylor expansion around $t = 0$ and that $o(t^2) = o(\Delta^2)$. Hence, the likelihood of privately observed failure over $[0, \Delta]$ equals $\Lambda(1-p)(1-\pi)\Delta + o(\Delta^2)$, the likelihood of public failure over $[0, \Delta]$ equals $\Lambda(1-p)\pi\Delta + o(\Delta^2)$, and the likelihood that the project succeeds over $[0, \Delta]$ equals $\Lambda p \Delta + o(\Delta^2)$. In addition, for all $t \in [0, \Delta]$

$$\alpha_t = \alpha_s + \dot{\alpha}_s(t-s) + o(\Delta^2) = \alpha_s + o(\Delta) \text{ for any } s \in [0, \Delta),$$

as $|s-t| = o(\Delta)$.

It follows that

$$\bar{\alpha}_\Delta := \mathbb{E}[\alpha_\tau | \text{Success over } [0,\Delta]] = \alpha_s + o(\Delta) \text{ for any } s \in [0, \Delta),$$

which is the expected compensation in case of success over $[0, \Delta]$. Likewise, one calculates that the expected financing costs over $[0, \Delta]$ equal

$$\bar{\kappa}_\Delta = \kappa(\Delta - \mathbb{P}(\text{Success or failure over } [0,\Delta])\mathbb{E}[\Delta - \tau | \text{Success or failure over } [0,\Delta]]).$$

Hence, $\bar{\kappa}_\Delta = \kappa\Delta - o(\Delta^2)$. Also note that

$$\bar{\beta}_\Delta := \mathbb{E}[\beta_\tau | \text{Privately observed failure over } [0,\Delta]] = \beta_s + o(\Delta) \text{ for any } s \in [0, \Delta).$$

Here, $\bar{\beta}_\Delta$ is the expected compensation for privately observed failure under the full disclosure contract.

Take arbitrary $s \in [0, \Delta)$. We can write the payoff under the full disclosure contract as

$$\begin{aligned} F^0 &= (1 - e^{-\Lambda\Delta})\left(p(\mu - \bar{\alpha}_\Delta) - (1-p)(1-\pi)\bar{\beta}_\Delta\right) - \bar{\kappa}_\Delta + e^{-\Lambda\Delta}F(w_\Delta) + o(\Delta^2) \\ &= \Lambda\Delta(p(\mu - \alpha_s) - (1-p)(1-\pi)w_s) - \kappa\Delta + (1 - \Lambda\Delta)F(w_\Delta) + o(\Delta^2) \\ &= \Lambda\Delta\left(p\mu - w_\Delta - \frac{\phi}{\Lambda}\right) - \kappa\Delta + (1-\Lambda\Delta)F(w_\Delta) + o(\Delta^2), \end{aligned}$$

where it was used in the last inequality that $\alpha_s = \phi/(\Lambda p) + w_\Delta(1 - \pi + \pi/p) + o(\Delta)$ and $\beta_s = w_s$ for $s \in [0, \Delta)$ under the full disclosure contract.

Likewise, the contract $\tilde{\mathcal{C}}$ implements the same deadline T as $\mathcal{C}^0$ and yields, by argument similar to the ones above, payoff at time zero

$$\begin{aligned} F^1 &= \Lambda\Delta(p(\mu - \tilde{\alpha}_s) - (1-p)(1-\pi)w_\Delta) - \kappa\Delta + (1-\Lambda\Delta)F(w_\Delta) + o(\Delta^2) \\ &= \Lambda\Delta\left(p\mu - w_\Delta(1-\pi) - \frac{\phi}{\Lambda}\right) - \kappa\Delta + (1-\Lambda\Delta)F(w_\Delta) + o(\Delta^2), \end{aligned}$$

where it was used in the last inequality that $\tilde{\alpha}_s = \phi/(\Lambda p) + w_\Delta(1-\pi) + o(\Delta)$ for $s \in [0, \Delta)$ under the contract $\tilde{\mathcal{C}}$.

Hence:

$$F^1 - F^0 = \Lambda\Delta w_\Delta \pi + o(\Delta^2),$$

which exceeds zero for Δ sufficiently small, contradicting the optimality of $\mathcal{C}^0$. Note that $w_\Delta > 0$ as $T > \Delta$. Thus, a full disclosure contract is not optimal.

On the other hand, if $\pi = 0$, it follows, due to $\alpha_t \geq w_t + \phi/(\Lambda p)$, that

$$\dot{W}_t = \Lambda(W_t - w_t - p(\alpha_t - w_t)) \leq \Lambda(W_t - w_t) - \phi = \Lambda(W_t - w_t) + \dot{w}_t.$$

Thus,

$$\dot{W}_t - \dot{w}_t \leq \Lambda(W_t - w_t)$$

and, therefore, due to $W_T = w_T = 0$ that $W_t \geq w_t$. Hence, a full disclosure contract is optimal by virtue of Lemma 4.

Step III

Lemma 6 *Define* $t_1 = \inf\{t \geq 0 : W_t \geq w_t\}$. *Then, within the optimal contract,* $t_1 > 0$ *if* $\pi > 0$, *and* $\alpha_t = w_t(1-\pi) + \phi/(\Lambda p)$, $\beta_t = \gamma_t = 0$ *for* $t < t_1$. *And,* $\alpha_t - r_t = \frac{\phi}{\Lambda p} + \frac{\pi}{p}(w_t - \gamma_t)$, $\gamma_t \in [0, w_t]$, *and* $\beta_t = w_t$ *for* $t \geq t_1$. *In addition,* $W_t < w_t$ *and* $0 > \dot{W}_t > \dot{w}_t$ *for* $t < t_1$ *and* $W_t = w_t$ *and* $\dot{W}_t = \dot{w}_t$ *for* $t \geq t_1$

Lemma 4 implies that the optimal continuation contract from time t_1 up to the deadline T is a full disclosure contract. If $t_1 = 0$, the optimal contract is a full disclosure contract from time zero so that $W_{t_1} = w_{t_1}$; otherwise, if $t_1 > 0$, continuity implies $W_{t_1} = w_{t_1}$ for $t_1 > 0$. Proposition 5 yields the second claim of the Lemma; that is, $\alpha_t - r_t = \frac{\phi}{\Lambda p} + \frac{\pi}{p}(w_t - \gamma_t)$, $\gamma_t \in [0, w_t]$, and $\beta_t = w_t$ for $t \geq t_1$. In addition, recall that by Lemma 5 a full disclosure contract is not optimal if and only if $\pi > 0$. As a full disclosure contract is optimal from time t_1 onward (see Lemma 4), it follows that $t_1 > 0$ if and only if $\pi > 0$.

To prove the first claim, fix a time t_1 after which the contract implements a full disclosure contract, while the contract does not incentivize disclosure of failure over $[0, t_1)$ and accordingly sets $\beta_t = 0$ for $t < t_1$. Define

$$\begin{aligned}\bar{\kappa}_{t_1} = \kappa(t_1 - \mathbb{P}(&\text{Success or public failure over } [0, t_1])\\ &\mathbb{E}[t_1 - \tau | \text{Success or public failure over } [0, t_1)]).\end{aligned}$$

In optimum, the contract implements full effort $a_t = 1$ for all $t \in [0, T \wedge \tau]$. The surplus at time zero generated by such a contract equals

$$\mathcal{S}_0 = e^{-\Lambda t_1}(F(w_{t_1}) + w_{t_1}) - \bar{\kappa}_{t_1} + (1 - e^{-\Lambda t_1})\Big(p\mu + \phi(1-p)(1-\pi)\mathbb{E}(t_1 - \tau | \tau < t_1)\Big),$$

which does not depend on $\{\alpha_t, \gamma_t\}_{t \leq t_1}$. Also note that the continuation surplus at time t_1 is split between the principal and the agent and — due to $w_{t_1} = W_{t_1}$ — equals $F(w_{t_1}) + w_{t_1}$. The principal's payoff equals:

$$F_0^* = \mathcal{S}_0 - W_0,$$

and, therefore, is maximized by the choice of $\{\alpha_t, \gamma_t\}_{t \leq t_1}$, that minimizes W_0 while $W_{t_1} = w_{t_1}$ and incentive compatibility $\alpha_t \geq \phi/(\Lambda p) + w_t(1-\pi) + \pi\gamma_t$. It is clear that W_0 is minimized upon promising zero rewards for failure and the lowest rewards for success possible that induce effort. That is, for $t < t_1$ setting $\alpha_t = \phi/(\Lambda p) + w_t(1-\pi)$ and $\beta_t = \gamma_t = 0$ is optimal.

Hence, for $t < t_1$

$$W_t = \int_t^T e^{-\Lambda(s-t)}\Lambda\Big((1-p)(1-\pi)w_s + p\alpha_s\Big)ds.$$

Differentiating yields

$$\begin{aligned}\dot{W}_t &= \Lambda W_t - \Lambda((1-p)(1-\pi)w_t + p\alpha_t)\\ &= \Lambda W_t - \Lambda(w_t(1-\pi) + p(\alpha_t - w_t(1-\pi)))\end{aligned}$$

and therefore

$$\dot{W}_t - \dot{w}_t = \Lambda W_t - \Lambda(w_t(1-\pi) + p(\alpha_t - w_t(1-\pi))) - \dot{w}_t = \Lambda(W_t - w_t) + \Lambda\pi w_t,$$

where we plugged in $\alpha_t = w_t(1-\pi) + \phi/(\Lambda p)$ and $\dot{w}_t = -\phi$ for $t < t_1$. Integrating this ODE for $t < t_1$ subject to $W_{t_1} = w_{t_1}$ yields $W_t - w_t = -\int_t^{t_1} e^{-\Lambda(s-t)}\Lambda\pi w_s ds < 0$ and hence also $0 > \dot{W}_t > \dot{w}_t = -\phi$.

Step IV

Lemma 7 *The value function is characterized by* (2.23) *and solved subject to* (2.21) *and* (2.22).

Given the previous Lemmata, all that remains is to determine the optimal deadline T and the first time at which the contract incentivizes truthful disclosure of failure, $t_1 = \inf\{t \geq 0 : \beta_t > 0\}$ or equivalently $t_1 = \inf\{t \geq 0 : W_t \geq w_t\}$. These two quantities are determined to maximize the principal's ex-ante payoff F_0^*.

Note that w_t — in optimum — perfectly co-moves with time t (before completion) and therefore can be taken as state variable. Hence, we can equivalently maximize F_0^* over $w_{t_1} = w_L$ and w_0, which uniquely pins down T and t_1 due to

$$\dot{w}_t = -\phi \quad \forall \quad t \in [0, T \wedge \tau^A] \implies w_t = w_0 - \phi t.$$

We solve the maximization problem sequentially: we first fix $T > 0$, which is equivalent to fixing $w_0 = \inf\{t \geq 0 : w_t = 0\}$ due to $\dot{w}_t = -\phi$ for all $t \in [0, T \wedge \tau^A]$, and maximize over t_1 (or equivalently w_L). We then obtain t_1 in dependence of T and thereafter maximize over T.

Let now

$$q_t = \mathbb{P}_t(\{\text{Hidden failure before } t\}) = 1 - e^{-\Lambda(1-p)(1-\pi)t}, \tag{2.34}$$

the principal's belief, formed over $[0, t_1)$ that the project has hiddenly failed. For $t < t_1 \wedge \tau^A$, we can rewrite the principal's payoff F_t^* (under the optimal contract) as

$$\begin{aligned}
F_t^* &= \mathbb{E}_t\left[\int_t^{T\wedge\tau^A} (\mu dS_s - \kappa ds - dc_s)\right] \\
&= \mathbb{E}_t\left[\int_t^{t_1\wedge\tau^A} (\mu dS_s - \kappa ds - dc_s)\right] \\
&\quad + \mathbb{P}_t(\tau^A \geq t_1)\Big(\mathbb{P}_t(\{\text{No Failure before } t_1\})F_{t_1}^* - \mathbb{P}_t(\{\text{Failure before } t_1\})\beta_{t_1}\Big) \\
&= \int_t^{t_1} e^{-\Lambda(p+(1-p)\pi)\int_t^s (1-q_u)du}\big(\Lambda(1-q_s)p(\mu-\alpha_s) - \kappa\big)ds \\
&\quad + e^{-\Lambda(p+(1-p)\pi)\int_t^{t_1}(1-q_u)du} \qquad (2.35) \\
&\Big(\mathbb{P}_t(\{\text{No Failure before } t_1\})F_{t_1}^* - \mathbb{P}_t(\{\text{Failure before } t_1\})\beta_{t_1}\Big) \\
&= \int_t^{t_1} e^{-\Lambda(p+(1-p)\pi)\int_t^s (1-q_u)du}\big(\Lambda(1-q_s)p(\mu-\alpha_s) - \kappa\big)ds \qquad (2.36) \\
&\quad + e^{-\Lambda(p+(1-p)\pi)\int_t^{t_1}(1-q_u)du}\Big((1 - \lim_{t\uparrow t_1} q_t)F_{t_1}^* - \lim_{t\uparrow t_1} q_t\beta_{t_1}\Big).
\end{aligned}$$

Differentiating (2.35) on $t \in [0, t_1)$ yields

$$\dot{F}_t^* = \Lambda(p + (1-p)\pi)(1-q_t)F_t^* - \Lambda p(1-q_t)(\mu - \alpha_t) + \kappa.$$

Utilizing

$$\dot{F}_t^* = \frac{dF_t^*}{dt} = \frac{dF_t^*}{dw_t}\frac{dw_t}{dt},$$

one arrives at the ODE (2.20), whereby $F_t^* = f(w_t)$ for $t \in [0, t_1)$. The closed form solution — given $w_0 > w_L \geq 0$ — is derived in Lemma 2 in Appendix 2.9.

Taking the limit $t \uparrow t_1$ in (2.35) yields

$$\lim_{t \uparrow t_1} F_t^* = (1 - \lim_{t \uparrow t_1} q_t)F_{t_1}^* - \lim_{t \uparrow t_1} q_t \beta_{t_1}, \tag{2.37}$$

which leads after substituting $\beta_{t_1} = w_{t_1} = w_L$, $\lim_{t \uparrow t_1} q_t = q(w_L)$, $\lim_{t \uparrow t_1} F_t^* = f(w_L)$, and $F_{t_1}^* = F(w_L)$ to the value matching condition (2.21).

As is standard for dynamic optimization problems, a necessary optimality condition is the smooth pasting condition

$$\frac{\partial}{\partial t_1} \lim_{t \uparrow t_1} F_t^* = \frac{\partial}{\partial t_1}\left((1 - \lim_{t \uparrow t_1} q_t)F_{t_1}^* - \lim_{t \uparrow t_1} q_t \beta_{t_1}\right). \tag{2.38}$$

If (2.38) did not hold, the principal could improve her payoff by increasing or decreasing t_1, implying that optimal t_1 must adhere to (2.38). Condition (2.38) is equivalent to condition (2.22), which is obtained after substituting $\beta_{t_1} = w_{t_1} = w_L$, $\lim_{t \uparrow t_1} q_t = q(w_L)$, $\lim_{t \uparrow t_1} F_t^* = f(w_L)$, and $F_{t_1}^* = F(w_L)$. We have solved for optimal t_1 or equivalently optimal w_L, given w_0 or equivalently T, which results in a value F_0^*. That is, w_L is a function of w_0. The optimization is complete after maximizing F_0^* over w_0.

Lemma 8 *$t_1 \to 0$ as $\pi \to 0$ and $t_1 \to T$ as $\pi \to 1$.*

If $\pi = 0$, by Lemma 5, a full disclosure contract is optimal, leading to $t_1 = 0$. This yields the first claim by virtue of continuity.

By Proposition 1 and continuity, it must be that $T \to \infty$ and, therefore, $w_t \to \infty$ for any $t < \tau$ as $\pi \to 1$. Incentivizing the agent not to fake failure and to truthfully disclose failure at some time $t < \infty$ requires $W_t \to \infty$, which cannot be optimal. This proves the second claim.

Step V

Lemma 9 *Stochastic (i.e., random) termination is not optimal.*

To simplify notation, we consider that the principal optimally sets $\gamma_t = 0$, which is strictly optimal when the contract does not incentivize disclosure of failure and weakly optimal when the contract incentivizes disclosure of failure. Consider that the principal randomly terminates the agent's contract at time t at endogenous rate $\delta_t \geq 0$ or with some atom of probability $\Theta_t \in [0, 1]$, in that the termination time T is (possibly) stochastic. Upon termination at time T, the agent receives zero payoff. Thus, the agent's payoff after privately observed failure at time $t < T$ reads

$$\begin{aligned} w_t &:= \max_{\tau^A \geq t} \mathbb{E}_t^A \left[(\tau^A \wedge T - t)\phi + \mathbb{P}(\tau^A < T)\beta_{\tau^A}\right] \\ &= \max_{\tau^A \geq t} \int_t^{\tau^A} e^{-\int_t^s \delta_u du} \prod_{t \leq u \leq s} (1 - \Theta_u)\phi ds + e^{-\int_t^{\tau^A} \delta_u du} \prod_{t \leq u \leq \tau^A} (1 - \Theta_u)\beta_{\tau^A}, \end{aligned} \tag{2.39}$$

where the second equality integrates out the random termination event. When the principal does not incentivize disclosure so that the agent chooses $\tau^A > t$, we use expression (2.39) to obtain

$$\dot{w}_t = \delta_t w_t - \phi, \quad \text{if } \Theta_t = 0 \tag{2.40}$$
$$w_t = \frac{\lim_{s\uparrow t} w_s}{1-\Theta_t}, \quad \text{if } \Theta_t > 0.$$

If the principal terminates the project with some atom of probability $\Theta_t > 0$, then w increases discretely from $w_{t^-} := \lim_{s\uparrow t} w_s$ to $\frac{w_{t^-}}{1-\Theta_t}$, if the project is not terminated (which happens with probability $1-\Theta_t$), and w drops to zero (i.e., $w_t = 0$), if the project is terminated at time t (which happens with probability Θ_t).[19] When the principal incentivizes disclosure of failure so that $\tau^A = t$ and $w_t = \beta_t$, the agent has no incentives to delay disclosure of failure only if

$$\frac{\partial w_t}{\partial \tau^A}\Big|_{\{\tau^A=t\}} \leq 0 \iff \dot{\beta}_t \leq \delta_t \beta_t - \phi, \quad \text{if } \Theta_t = 0 \tag{2.41}$$
$$\beta_t \leq \frac{\lim_{s\uparrow t} \beta_s}{1-\Theta_t}, \quad \text{if } \Theta_t > 0.$$

To get some intuition for (2.41), suppose that $\Theta_t = 0$ and the agent has privately observed failure at time $s \leq t$. If the agent discloses failure at time t, he receives β_t dollars. If the agent discloses failure at time $t + dt$, he derives benefits of shirking ϕdt. Then, the project is terminated with probability $\delta_t dt$ and the agent receives zero dollars at $t + dt$. Otherwise, with probability $1 - \delta_t dt$, there is no termination and the agent receives β_{t+dt} dollars at $t + dt$. Thus, the agent does not delay disclosure and reports failure at time t if $\beta_t \geq (1 - \delta_t dt)\beta_{t+dt} + \phi dt$. Taking $dt \to 0$ and rearranging yields the first line of (2.41).[20] In optimum, this incentive constraint (2.41) is tight and becomes (2.40), as $w_t = \beta_t$.

Similarly, the agent's continuation payoff (before time τ) is

$$W_t = \int_t^\infty e^{-\Lambda(s-t)-\int_t^s \delta_u du} \prod_{t\leq u\leq s} (1-\Theta_u)\Lambda\left(p\alpha_s + (1-p)(1-\pi)w_s\right) ds, \tag{2.42}$$

so that

$$\dot{W}_t = (\Lambda + \delta_t)W_t - \Lambda(p\alpha_t + (1-p)(1-\pi)w_t), \quad \text{if } \Theta_t = 0 \tag{2.43}$$
$$W_t = \frac{\lim_{s\uparrow t} W_s}{1-\Theta_t}, \qquad \text{if } \Theta_t > 0.$$

Like in the baseline, one can then calculate that $\dot{W}_t = \dot{w}_t$ with $W_t = w_t$ and $\Theta_t = 0$ if and only if $\alpha_t = (1 - \pi + \pi/p)w_t + \frac{\phi}{\Lambda p}$. In the following, we conjecture that $\Theta_t = 0$ is optimal whenever $w_t > 0$, in that $\Theta_t = 0$ and $\dot{w}_t = \delta_t w_t - \phi$ (for $w_t > 0$) while $\Theta_t = 1$ if $w_t = 0$. At the end of the proof, we verify that this is indeed the case.

In principle, there are three state variables: i) the agent's payoff after privately observed failure $w_t = w$, evolving according to (2.40), ii) the principal's belief of

[19] In more detail, consider the principal terminates financing at time t with an atom of probability Θ_t, and denote by w_t the value after this random event has realized. Likewise, $w_{t^-} = \lim_{s\uparrow t} w_s$ is the value at time t^-, just before the random termination event realizes. Having privately observed failure before time t, the agent's payoff at time t^- is $w_{t^-} = w_t(1-\Theta_t) + o(dt)$, or $w_t = w_{t^-}/(1-\Theta_t)$.

[20] Likewise, suppose that $\Theta_t > 0$. If the agent reports failure at time $t^- = \lim_{s\uparrow t} s$ just before the random termination event realizes, she receives β_{t^-} dollars with certainty. If the agent reports failure at t just after the random termination event realizes, the agent receives zero dollars in case of termination (with probability Θ_t) and β_t dollars otherwise. Thus, the agent does not delay disclosure if the second line of (2.41) holds.

whether the project has failed so far $q_t = q$, evolving according to $\dot{q}_t = (1-q_t)(1-p)(1-\pi)\Lambda > 0$ when the principal does not incentivize disclosure, and iii) the agent's continuation payoff W_t. We can express the principal's value function at time t, denoted f_t, and the agent's continuation value (before failure), denoted W_t, as functions of (w,q), in that $f_t = f(w_t, q_t)$ and $W_t = W(w_t, q_t)$.[21] To simplify notation, we omit time subscripts whenever no confusion is likely to arise.

Note that by $\dot{q}_t = (1-q_t)(1-p)(1-\pi)\Lambda > 0$, there exists a one-to-one mapping from time t to the belief q_t (when the principal does not incentivize disclosure). Likewise, without random termination, it holds that $\delta_t = 0$ and $\dot{w}_t = -\phi < 0$, so there exists a one-to-one mapping from time t to w_t and therefore also a one-to-one mapping from w_t to q_t and we can express q_t as function of w_t (i.e., $q_t = q(w_t)$). Crucially, the one-to-one mapping between time t and w_t and therefore the one-to-one mapping between w_t and q_t need not exist if there is the possibility of random termination (as $\delta_t > 0$ affects $\dot{w}_t$), so with random termination we generally have to explicitly keep track of both w_t and q_t (when the principal does not incentivize disclosure).

In what follows, we prove that random termination at rate $\delta > 0$ is not optimal. In the general formulation with random termination, the principal's value function at time t, denoted f_t, is a function of (w,q), in that $f_t = f(w_t, q_t)$. By the dynamic programming principle, the value function $f(w,q)$ solves the HJB equation

$$(1-q)\Lambda\big(p+(1-p)\pi\big)f(w,q) = \tag{2.44}$$

$$\max_{\delta\geq 0}\Big\{ -\kappa - f_w(w,q)\phi + \delta(f_w(w,q)w - f(w,q)) + f_q(w,q)\dot{q} + (1-q)\Lambda\Big[p(\mu-\alpha) - (1-p)(1-\pi)(w+f(w,q))\mathbb{I}\{(w,q)\in\mathcal{I}\}\Big], \tag{2.45}$$

where we use $\gamma = 0$, $\dot{w} = \delta w - \phi$, and α denotes the agent's pay for success along the optimal path. Here, $\mathbb{I}\{\cdot\}$ denotes the indicator function which is equal to one if $\{\cdot\}$ is true and equal to zero otherwise. Note that $(w,q)\in\mathcal{I}$ if and only if the contract incentivizes disclosure of failure in state (w,q) in which case $q = \dot{q} = 0$ and $\beta = w$. Conversely, $(w,q)\notin\mathcal{I}$ if and only if the contract does not incentivize disclosure of failure in state (w,q) in which case $q \geq 0$, $\dot{q} > 0$, and $\beta = 0$. In (2.44), a subscript denotes the partial derivative, in that $f_x(w,q) = \frac{\partial f(w,q)}{\partial x}$ for $x \in \{w,q\}$. When $\delta = 0$ and the principal does not incentivize disclosure of failure, then (2.44) is equivalent to (2.20), after expressing q as function of w (i.e., $q_t = q(w_t)$).[22] When $\delta = 0$ and the principal incentivizes disclosure of failure so that $\beta = w > q = \dot{q} = 0$, (2.44) simplifies to (2.14) (with $f(w,0)$ instead of $F(w)$ denoting the value function).

We study the principal's incentives to stochastically terminate the contract at rate δ. As the contract is terminated (with certainty) once $w = 0$, it suffices to consider

[21] Given optimal controls $\alpha_t = \alpha(w_t, q_t)$ and $\delta_t = \delta(w_t, q_t)$ (which are functions of (w_t, q_t)), we can express $W_t = W(w_t, q_t)$ as function of (w_t, q_t) using (2.43), leading to the PDE $\dot{W}_t = W_w(w_t, q_t)\dot{w}_t + W_q(w_t, q_t)\dot{q}_t = (\Lambda + \delta(w_t, q_t))W(w_t, q_t) - \Lambda((1-p)(1-\pi)w_t + p\alpha(w_t, q_t))$. Solving this PDE yields the function $W(w_t, q_t)$.

[22] To see this, conjecture that there exists a function $g(w)$ such that $g(w) = f(w,q)$, so that $\frac{dg(w)}{dt} = \frac{df(w,g)}{dt}$. As such,

$$\frac{dg(w)}{dt} = g'(w)\dot{w} = f_w(w,q)\dot{w} + f_q(w,q)\dot{q}. \tag{2.46}$$

Inserting $q = q(w)$, $g(w) = f(w,q)$, and (2.46) as well as $\delta = 0$ into (2.44) yields (2.20) (with $g(w)$ instead of $f(w)$ denoting the principal's value function).

values $w > 0$. Taking the partial derivative with respect to δ in (2.44) yields

$$\frac{\partial f(w,q)}{\partial \delta} \propto f_w(w,q)w - f(w,q),$$

where $\propto$ denotes proportionality. Hence, the principal's objective in (2.44) is linear in δ. Note that if it were $f_w(w,q)w > f(w,q)$, then — by the HJB equation (2.44) — setting $\delta \to \infty$ would be optimal and would yield unbounded payoff for the principal, which cannot be as the principal's payoff is bounded from above by the net present value $\mu p - \kappa/\Lambda$. Thus, it holds that

$$f_w(w,q)w - f(w,q) \leq 0, \tag{2.47}$$

with equality if $\delta > 0$ is optimal. That is, when $\delta > 0$ is optimal, the smooth pasting condition

$$f(w,q) = f_w(w,q)w \tag{2.48}$$

holds for $(w_t, q_t) = (w,q)$. Hence, $\delta(f_w(w,q)w - f(w,q)) = 0$, so that (2.44) simplifies to

$$\begin{aligned}(1-q)\Lambda\big(p + (1-p)\pi\big)f(w,q) = \Big\{ &-\kappa - f_w(w,q)\phi + f_q(w,q)\dot{q} \\ &+ (1-q)\Lambda\Big[p(\mu-\alpha) - (1-p)(1-\pi)(w + f(w,q))\mathbb{I}\{(w,q) \in \mathcal{I}\}\Big].\end{aligned} \tag{2.49}$$

In addition, the super contact condition

$$f_{ww}(w,q) = 0 \tag{2.50}$$

must hold for $(w_t, q_t) = (w,q)$ when $\delta > 0$ is optimal. The smooth pasting condition (2.48) can be interpreted as local optimality condition and the super contact condition (2.50) can be interpreted as global optimality condition; for a more detailed discussion, see Dumas (1991a).[23] In what follows, we show that $f(w,q) > f_w(w,q)w$ holds for all $w > 0$, which implies that stochastic termination is not optimal and $\delta = 0$.

First, we consider that the principal incentivizes disclosure of failure so that $\dot{q}_t = \dot{q} = 0$ and $q_t = q = 0$ as well as $w_t = \beta_t$, which is the case during the disclosure stage of the contract from Proposition 6 for times $t \geq t_1$. We conjecture that there is no stochastic termination when the contract incentivizes disclosure. Then, our previous results imply that the optimal contract incentivizes disclosure at all future times until termination, whereby $W_t = w_t = \beta_t$, $\dot{W}_t = \dot{w}_t$, and $\alpha_t = \alpha(w_t) = (1 - \pi + \pi/p)w_t + \frac{\phi}{\Lambda p}$ at all future times until termination. Next, we show that in this case, $f(w,0) > f_w(w,0)w$ for $w > 0$, which proves that stochastic termination at rate $\delta > 0$ is indeed not optimal when the contract incentivizes disclosure. Note that the HJB equation (2.49) simplifies to (2.14), and the principal's value function only depends on w in that $F(w) = f(w,0)$, whereby $f(w,0) = F(0) = 0$. The ODE (2.14) with $\alpha(w) = (1 - \pi + \pi/p)w + \frac{\phi}{\Lambda p}$, $\dot{w} = -\phi$, and $F(0) = f(0,0) = 0$ admits the

[23]To see that the super contact condition must hold, assume to the contrary that (2.48) holds (i.e., $f(w,q) = f_w(w,q)w$), but (2.50) does not hold (i.e., $f_{ww}(w,q) \neq 0$). Then, there exists $w' \neq w$ with $w' > 0$ such that $f_w(w',q)w' - f(w',q) > 0$ (which violates (2.47)).

closed form solution (2.16) (see Lemma 1), that is,

$$F(w) = f(w,0) = \left(\mu p - \frac{\kappa}{\Lambda}\right)\left(1 - \exp\left(\frac{-w\Lambda}{\phi}\right)\right) - w.$$

Thus, $F''(w) = f_{ww}(w,0) < 0$. Due to $f(0,0) = 0$ and $f_{ww}(w,0) < 0$, it follows that $f(w,0) > f_w(w,0)w$ for all $w > 0$. Consequently, when the principal incentivizes disclosure (i.e., during the disclosure stage), stochastic termination at rate $\delta > 0$ is not optimal.

Second, we consider that the principal does not incentivize disclosure of failure so that $\dot{q}_t > 0$, which is the case during the unconditional financing stage (i.e., when $t < t_1$) of the contract from Proposition 6. We show that in this case, $f(w,q) > f_w(w,q)w$ for all $w > 0$ so that stochastic termination at rate $\delta > 0$ is indeed not optimal. When the principal does not incentivize disclosure of failure, it is optimal to stipulate minimal rewards for success so that the incentive condition for effort is tight and $\alpha = \alpha(w) = (1-\pi)w + \frac{\phi}{\Lambda p}$ leading to $\alpha'(w) > 0$ (i.e., $\alpha(w)$ increases with w).

Importantly, at time $t \geq 0$ when $(w,q) = (w_t, q_t)$ and $\delta_s = \Theta_s = 0$ for $s < t$, the value function satisfies $f_w(w,q) \geq 0$ (in the interior of the region where the principal does not incentivize disclosure and the value function is differentiable), as otherwise the principal could improve her payoff at time zero by shortening the deadline and decreasing the initial value of w (i.e., w_0). To see this formally, suppose to the contrary that under the optimal contract $\mathcal{C}$ at time t with $q_t = q$ and $w_t = w$ as well as $\delta_s = 0$ for $s < t$, there exists $w' < w_t$ such that $f(w', q_t) > f(w_t, q_t)$. Then, the principal could improve her payoff at time 0 by writing a contract $\mathcal{C}'$ that sets the same rewards for success α_s and public failure γ_s up to time t, and mimicks the original contract $\mathcal{C}$ up to time t in all aspects except one: it changes the initial value of w from w_0 to w_0' via $w_0' = w_0 - (w_t - w') < w_0$, thereby inducing a different path of w denoted w'. Under both contracts the belief q_s is the same and w_s decreases with $\dot{w}_s = \dot{w}_s' = -\phi$ for $s \leq t$. Under the alternative contract $\mathcal{C}'$, w_s' reaches the value w' at time $s = t$ (if no termination before t). As the alternative contract stipulates lower w (i.e., $w_s' < w_s$ for all $s \leq t$) but the same rewards for success up to t, it incentivizes effort. If project success or failure is publicly observed before time t, the principal realizes the same payoffs under both contracts. If the principal incentivizes disclosure at some time $s < t$ and the agent discloses failure at time s, the principal realizes higher payoff under the contract $\mathcal{C}'$ as $w_s' < w_s$. If there is no termination before time t, the contract $\mathcal{C}'$ yields strictly higher payoff, as it yields strictly higher continuation payoff $f(w', q_t) > f(w_t, q_t)$ at time t. Thus, the contract $\mathcal{C}'$ yields strictly higher payoff for the principal at time zero than the contract $\mathcal{C}$, a contradiction. Therefore, $f_w(w,q) \geq 0$.

Next, suppose that there exists time t with $w_t = w > 0$ and $f_w(w,q)w = f(w,q)$. Consider the first time $t \geq 0$ at which $f_w(w,q)w = f(w,q)$ so that for all $s < t$: $f_w(w_s,q_s)w_s < f(w_s,q_s)$ and $\delta_s = 0$. It follows that $f_w(w,q) \geq 0$ and $f(w,q) \geq 0$. Then, it must be that

$$\frac{\partial}{\partial q}\left(f_w(w,q)w - f(w,q)\right) \leq 0 \iff f_q(w,q) \geq f_{wq}(w,q)w, \tag{2.51}$$

as otherwise there exists $q' > q$ (with $q' < 1$) so that $f_w(w,q')w > f(w,q')$, which would violate (2.47).[24] As clearly the value function decreases in the belief of whether the project has failed so far in that $f_q(w,q) \leq 0$, the left hand side of (3.28) is negative, implying $f_{wq}(w,q) \leq 0$.[25] Next, we differentiate both sides of (2.49) with respect to w (in the interior of the region where the principal does not incentivize disclosure and the value function is twice differentiable) to obtain

$$-f_{ww}(w,q)\phi = (1-q)\Lambda\big(p+(1-p)\pi\big)f_w(w,q) + (1-q)\Lambda p\alpha'(w) - f_{wq}(w,q)\dot{q}. \tag{2.53}$$

As $f_w(w,q) \geq 0$, $f_{wq}(w,q) \leq 0$, $\dot{q} > 0$, and $\alpha'(w) > 0$, the right hand side of (2.53) is strictly positive, implying $f_{ww}(w,q) < 0$. Because $f_{ww}(w,q) < 0$ and $f_w(w,q)w = f(w,q)$, there exists $w' < w$ such that $f_w(w',q)w' > f(w',q)$, which violates (2.47) and yields a contradiction.[26] Thus, it holds that $f(w,q) > f_w(w,q)w$ for all $w > 0$, so random termination at rate $\delta > 0$ is not optimal when the principal does not incentivize disclosure of failure (i.e., during the unconditional financing stage).[27]

Finally, we verify that stochastic termination with some atom of probability $\Theta \in (0,1)$ is not optimal when $w > 0$. By (2.40), the principal's (net) payoff upon termination with an atom of probability Θ is

$$\mathcal{L}(\Theta) := f\left(\tilde{w},q\right)(1-\Theta) - f(w,q), \tag{2.54}$$

with $\tilde{w} = \frac{w}{1-\Theta}$. That is, if the project is not terminated, w increases multiplicatively times $\frac{1}{1-\Theta}$. Note that

$$\mathcal{L}'(\Theta) = f_w(\tilde{w},q)\tilde{w} - f(\tilde{w},q). \tag{2.55}$$

For $\Theta \in (0,1)$ to be optimal, it must be that the first order condition $\mathcal{L}'(\Theta) = 0$ holds, i.e., $f_w(\tilde{w},q)\tilde{w} = f(\tilde{w},q)$. However, we have shown that $f(w,q) > f_w(w,q)w$ for all $w > 0$, implying that stochastic termination with some atom of probability $\Theta > 0$ is not optimal when $w > 0$. Overall, we conclude that stochastic termination — at some rate δ or with some atom of probability $\Theta \in (0,1)$ — is not optimal, and termination at deterministic time T is optimal.

[24] Note that being in state (w,q), any state (w,q') with $1 > q' > q$ is accessible. To reach state (w,q') (in finite time from (w,q)), the principal can set $\delta = \delta(w,\hat{q}) = \frac{\phi}{w}$ for all $\hat{q} \in [q,q']$, leading to $\dot{w} = 0$ and $\dot{q} > 0$ until q hits q'.

[25] To see this, note that by the law of iterated expectations, we can rewrite the principal's value function at time t with $(w,q) = (w_t,q_t)$ as

$$f(w,q) = (1-q)\mathbb{E}_t\left[\hat{f}_t|q=0\right] + q\mathbb{E}_t\left[\hat{f}_t|q=1\right],$$

where $\hat{f}_t$ is the principal's actual payoff from time t onwards, i.e.,

$$\hat{f}_t := \int_t^{T\wedge\tau^A} \left(\mu dS_s - \kappa ds - dc_s\right). \tag{2.52}$$

Note that $\mathbb{E}_t\left[\hat{f}_t|q=1\right] \leq 0$, as the the project has negative net present value after failure (and the agent is protected by limited liability). As $f(w,q) \geq 0$, $\mathbb{E}_t\left[\hat{f}_t|q=0\right] \geq 0$ and $f_q(w,q) = \mathbb{E}_t\left[\hat{f}_t\middle|q=1\right] - \mathbb{E}_t\left[\hat{f}_t\middle|q=0\right] \leq 0$.

[26] Note that being in state (w,q), any state (w',q) with $w' < w$ is accessible, as the principal can always decrease w to w' by paying the agent $w - w'$ dollars.

[27] Notably, we have shown that $f_w(w,q)w = f(w,q)$ implies $f_{ww}(w,q) < 0$, thus stochastic termination at rate $\delta > 0$ cannot be optimal because the smooth pasting condition (2.48) and the super contact condition (2.50) cannot jointly hold.

2.10.5 Proof of Proposition 4

Preliminaries

This Section characterizes the optimal contract with monitoring and inspections. Recall that an inspection at time t allows the principal to detect whether the project has failed up to time t. Importantly, the principal cannot fully commit to an inspection/monitoring policy. That is, inspection times are not contractible, but inspection outcomes are public knowledge and contractible. As such, the principal conducts an inspection only if she finds it privately optimal to do so. With F_t denoting the principal's payoff when there is no inspection at time t and $\hat{F}_t$ denoting the principal's payoff after an inspection at time t yields that the project has not failed yet, the principal finds it privately optimal to conduct an inspection at time t if and only if $\hat{F}_t(1-q_t) - K \geq F_t$. Note that after an inspection yields that the project has not failed yet, the principal faces the same decision problem as at time zero and her continuation payoff becomes F_0 which is the principal's payoff at time zero, while the agent's continuation payoff becomes W_0 and w is reset to w_0. That is, $\hat{F}_t = F_0$.

When w reaches zero, the principal must either terminate or inspect the project with certainty. As we show below in Corollary 4, the agent has then incentives to truthfully disclose failure just before inspection/termination. As a result, whenever the principal inspects the project at $w = 0$, the inspection yields that the project has not failed yet so that the principal's payoff after inspection becomes F_0. Therefore, an inspection is preferred over contract termination (when w reaches zero) if and only if $F_0 \geq K$. Otherwise, in the case $F_0 < K$, the principal prefers contract termination over an inspection (when w reaches zero). We consider that the principal inspects the project at $w = 0$ as long as she weakly prefers an inspection over termination, which breaks the ties in the knife-edge case $F_0 = K$. As a result, when $F_0 \geq K$, it follows that $T = \infty$, and financing is never terminated. Unless otherwise mentioned, we make the assumption that K is not too large in that $K \leq F_0$, so that an inspection occurs with non-trivial probability within the optimal contract. If there are no inspections at all, the optimal contract is described in Proposition 6.

We assume that the termination time T is deterministic.[28] As in the baseline, we consider that the principal incentivizes full effort $a_t = 1$ and that the principal incentivizes the agent not to fake failure. We argue that the principal does not and cannot use inspections to incentivize the agent not to fake failure. To see this, suppose to the contrary that the principal promises to inspect the project with probability $\Theta_t > 0$ when the agent reports failure at time t, so as to discourage the agent to fake failure. That is, if the agent fakes failure, she is detected with probability Θ_t in which case her payoff is zero, while she collects rewards for failure β_t otherwise (with probability $1 - \Theta_t$). When the agent anticipates an inspection with probability Θ_t and $\beta_t(1 - \Theta_t) \leq W_t$, the agent does not fake failure, and any failure report at time t is true. Then, an inspection after a failure report at time t yields that the project indeed has failed so that the principal's continuation payoff becomes zero. However, when the agent anticipates an inspection with probability Θ_t and has sufficient incentives not to fake failure so that any reported failure is true, the principal who cannot fully commit to inspections would profitably deviate by not inspecting the project after a failure report at time t to save the inspection costs K.

The structure of the model solution with monitoring is as follows. We first solve the *relaxed problem*, in which the agent *cannot* fake failure, leading to the constraint

[28] Likely, one can adapt the arguments presented in the proof of Lemma 9 to verify that stochastic termination is indeed not optimal.

in the strategy space $\tau^A \geq \tau$. We show that the optimal contract is a full disclosure contract, when the constraint $\tau^A \geq \tau$ is in place. Whenever we refer to the *relaxed problem*, we implicitly assume that the constraint $\tau^A \geq \tau$ is in place. Then, we solve the full problem in which the agent can fake failure so that $\tau^A < \tau$ is possible. In detail, the proof is split in six separate parts. Part I characterizes the agent's incentives with monitoring. Part II characterizes the principal's solution to the relaxed problem and shows that the optimal contract then always incentivizes disclosure. Part III considers the case that the solution to the relaxed problem also solves the full problem in which case the contract always incentivizes disclosure. Part IV considers the case that the solution to the relaxed problem does not solve the full problem in which case the contract does not always incentivize disclosure of failure. Part V characterizes the solution to the full problem. For parts II to V, we conjecture that inspections (i.e., monitoring) are not stochastic, and thus occur at deterministic dates. Importantly, Corollary 3 in Part I shows that an inspection at deterministic date occurs if and only if $w = 0$. That is, given the assumptions of deterministic inspection dates and $F_0 \geq K$, the principal inspects the project at all times w reaches zero (and at no other times). Part VI proves and verifies that deterministic monitoring and inspections are indeed optimal.

Part I — Agent's incentive compatibility

To characterize truth telling incentives, let $dM_t \in \{0, 1\}$ indicate whether the principal inspects the project at time t. We define the left limit of a process x_t as $x_{t^-} := \lim_{s\uparrow t} x_{t^-}$. If the principal inspects the project at time t, then time t^- is the time just before the inspection and time t is the time just after the inspection. Note that $\mathbb{P}(dM_t = 1) = m_t$ is the likelihood of an inspection. In general, we can write

$$m_t = \theta_t dt + \Theta_t, \tag{2.56}$$

with $\theta_t \geq 0$ and $\Theta_t \in [0, 1]$. If $\Theta_t = 0$, the principal inspects the firm with infinitesimal probability $\theta_t dt$, i.e., at rate θ_t. Otherwise, if $\Theta_t > 0$, the principal inspects the firm with an atom of probability and $\Theta_t = 1$ implies a deterministic inspection at time t. Recall that the principal cannot fully commit to a monitoring policy m_t, i.e., m_t is not contractible. Unless otherwise mentioned, we do not distinguish between the actual monitoring policy m_t and the monitoring policy $\hat{m}_t$ that is anticipated by the agent. Crucially, in optimum, these two policies coincide, so that $m_t = \hat{m}_t$.

The following Lemma derives for all times t with $dM_t = 0$ the truth telling incentive constraint so that the agent does not hide failure

$$d\beta_t \leq m_t \beta_t - \phi dt. \tag{2.57}$$

Notably, with $m_t = \theta_t dt + \Theta_t$ the constraint (2.57) is equivalent to

$$\begin{aligned} \dot{\beta}_t &\leq \theta_t \beta_t - \phi, \quad \text{if } \Theta_t = 0 \\ \beta_t &\leq \frac{\lim_{s\uparrow t} \beta_s}{1 - \Theta_t}, \quad \text{if } \Theta_t > 0.. \end{aligned} \tag{2.58}$$

Note that $\beta_{t^-} = \beta_t$, when β_t is continuous. Likewise,

$$\begin{aligned} \dot{w}_t &= \theta_t w_t - \phi, \quad \text{if } \Theta_t = 0 \\ w_t &= \frac{\lim_{s\uparrow t} w_s}{1-\Theta_t}, \quad \text{if } \Theta_t > 0. \end{aligned} \tag{2.59}$$

holds.

Lemma 10 *A contract $\mathcal{C}$ induces truthful disclosure of failure (i.e., $\tau^A = \tau$ with certainty) from time t' onwards if and only if* (2.58) *holds for all $t \in [t', T]$ with $dM_t = 0$ and $T \wedge \tau^M < \infty$ (almost surely), where $\tau^M = \inf\{s > t : dM_s = 1\}$. It induces full effort, $a_t = 1$ for all $t \in [0, T \wedge \tau]$, if and only if $\alpha_t \geq r_t + \phi/(\Lambda p)$.*

Without loss of generality, normalize for the proof $t' = 0$. First, consider any $t \geq \tau$ and that the project has failed. Define $\tau^M := \inf\{s \geq t : dM_s = 1\}$ as the next time the principal inspects the project, whereby — for convenience — the notation does not make the dependence of τ^M on t explicit. Once an inspection yields that the agent has hidden failure, it is optimal to terminate financing and to fire the agent (without any severance pay).

Then, given a contract deadline $T \geq t$ the agent's payoff becomes

$$\begin{aligned} w_t &:= \max_{\tau^A \in [t,T]} \mathbb{E}_t^A \left[(\tau^A \wedge \tau^M - t)\phi + \mathbb{P}(\tau^A < \tau^M)\beta_{\tau^A} \right] \\ &= \max_{\tau^A \in [t,T]} \int_t^{\tau^A} e^{-\int_t^s \theta_u du} \prod_{t \leq u \leq s} (1-\Theta_u)\phi ds + e^{-\int_t^{\tau^A} \theta_u du} \prod_{t \leq u \leq \tau^A} (1-\Theta_u)\beta_{\tau^A}, \end{aligned} \tag{2.60}$$

where we have integrated out the random inspection event. That is, when there is no inspection, the relation (2.59) holds. The above expression (2.60) for w_t is maximized for $\tau^A = t$ only if $\frac{\partial w_t}{\partial \tau^A} = \dot{\beta}_{\tau^A} - \theta_{\tau^A}\beta_{\tau^A} + \phi \leq 0$ for $\tau^A = t$, in case $\Theta_t = 0$, or if $\beta_t(1-\Theta_t) \leq \beta_{t^-}$, in case $\Theta_t > 0$. Notably, this is (2.58) and a necessary condition for truthful disclosure of failure. Since the project may complete at any time during $t \in [0, T]$, $\tau^A = \tau$ can be achieved with certainty only if (2.58) holds for all $t \in [0, T]$. Then, we obtain

$$\beta_t \geq \max_{\tau^A \in [t,T]} \int_t^{\tau^A} e^{-\int_t^s \theta_u du} \prod_{t \leq u \leq s} (1-\Theta_u)\phi ds + e^{-\int_t^{\tau^A} \theta_u du} \prod_{t \leq u \leq \tau^A} (1-\Theta_u)\beta_{\tau^A}, \tag{2.61}$$

for all $\tau^A \in [t, T]$. It must hold that $T \wedge \tau^M < \infty$ almost surely, as otherwise the agent would hide failure forever and derive infinite payoff from doing so (in which case (2.61) cannot hold for $\beta_t < \infty$). On the other hand, if (2.58) holds for any $t \in [0, \tau^M \wedge T]$ and $\mathbb{E}[T \wedge \tau^M] < \infty$, it follows that (2.61) holds and hence that $\beta_t \geq w_t$. As a result, w_t is maximized for $\tau^A = t$ so that $\tau^A \leq \tau$ is optimal. Due to the assumed constraint $\tau^A \geq \tau$, this already implies truthful disclosure of failure $\tau^A = \tau$.

Given truthful disclosure of failure, the agent chooses effort $\{a_s\}_{\{s \geq t\}}$ to maximize

$$W_t := \int_t^T e^{-\int_t^s \Lambda(s-t)} \Big(\Lambda\big(r_s + pa_s(\alpha_s - r_s)\big) + \phi_t(1-a_s) \Big) ds, \tag{2.62}$$

which boils down to maximize the integrand point-wise. This implies that the incentive condition

$$\alpha_t \geq r_t + \phi/(\Lambda p)$$

must hold to induce full effort. Next, we state the following Corollary which shows that a deterministic inspection occurs whenever w reaches zero.

Corollary 3 *Suppose that at time t, the principal conducts the next inspection at deterministic time τ^M. Then, it must be that $\lim_{s\uparrow\tau^M} w_s = 0$ and $\tau^M = \inf\{s \geq t : w_s = 0\}$.*

The principal conducts an inspection at time τ^M with probability one, so that $\Theta_{\tau^M} = 1$. By (2.60), it holds that $\lim_{s\uparrow\tau^M} w_s = 0$. On the other hand, when $w_s = 0$ for any $s \geq t$, the principal must either terminate financing or inspect the project, implying $\tau^M = \inf\{s \geq t : w_s = 0\}$. Another Corollary demonstrates that "just before" a deterministic inspection or termination date, the agent always discloses failure truthfully, so that any inspection at deterministic date yields that the project has not failed yet.

Corollary 4 *Suppose that the principal inspects or terminates the project at deterministic date τ^M. If the agent has privately observed failure before time τ^M, the agent truthfully discloses failure just before the inspection (termination) at time $\lim_{t\uparrow\tau^M} t$. Therefore, an inspection at time τ^M yields then that the project has not failed yet.*

Consider time $\lim_{t\uparrow\tau^M} t$, which is the time just before the inspection (termination) at deterministic date τ^M. Suppose the agent has privately observed failure before τ^M. If the agent discloses failure at time $\lim_{t\uparrow\tau^M} t$, the principal terminates financing and the agent's payoff is zero as $\lim_{t\uparrow\tau^M} w_t = 0$ by Corollary 3. If the agent does not disclose failure at time $\lim_{t\uparrow\tau^M} t$, the inspection reveals failure or the principal terminates financing, so the agent is fired and his payoff is zero. Thus, the agent finds it (weakly) optimal to disclose failure at time $\lim_{t\uparrow\tau^M} t$. Therefore, an inspection at time τ^M — if it occurs — yields that the project has not failed.

Part II — The principal's solution to relaxed problem

Lemma 11 *Restrict the agent's strategy to $\tau^A \geq \tau$, i.e., consider the relaxed problem. Then, the optimal contract is a full disclosure contract, i.e., induces $\tau^A = \tau$ with certainty. In addition, the optimal contract sets $w_t = \beta_t$, $\gamma_t = 0$, and $\alpha_t = w_t(1-\pi) + \frac{\phi}{\Lambda p}$. The principal inspects the project at time t if and only if $w_{t^-} = 0$ and $F_0 < K$. Otherwise, when $F_0 < K$, the principal terminates the project at time t if and only if $w_{t^-} = 0$.*

Take any value of the agent's payoff after privately observed failure at time zero $w_0 > 0$. Consider $F_0 \geq K$. Then, the principal prefers to inspect the project rather than to terminate the project. By Corollary 3, an inspection occurs if and only if w reaches zero. At time 0 and after an inspection, the next inspection occurs in at most w_0/ϕ units of time, as up to the next inspection $\dot{w}_t \leq -\phi$. After an inspection, w is reset to w_0, and the principal's continuation becomes F_0 because after an inspection at time τ^M the principal faces the same decision problem as at time zero. If $F_0 < K$, the principal prefers termination over an inspection, so financing is terminated when w reaches zero.

The principal's payoff at time zero is the difference between total surplus $\mathcal{S}_0$ and the agent's payoff at time zero, W_0, in that total surplus $\mathcal{S}_0$ at time zero is split between agent and principal and $F_0 = \mathcal{S}_0 - W_0$. Given w_0 and given that the principal inspects or terminates the project when $w = 0$, the principal minimizes the agent's continuation payoff (subject to incentive compatibility for effort) by setting $\gamma_t = 0$, $w_t = \beta_t$, and $\alpha_t = (1-\pi)w_t + \frac{\phi}{\Lambda}$ at all times $t \leq T$. This choice induces $\dot{w}_t = \dot{\beta}_t = -\phi$ for $t \leq T$ whenever there is no inspection, so that the agent has no incentives to delay disclosure of failure leading to $\tau^A \leq \tau$. Due to the constraint $\tau^A \geq \tau$, the agent discloses failure at the time it occurs, i.e., $\tau^A = \tau$. In addition, by setting $\dot{w}_t = -\phi$ for $t \leq T$, the principal maximizes the time to the next inspection date or termination, thereby minimizing monitoring costs or deadweight losses. As a result, the principal also maximizes total surplus. In other words, for any w_0, setting $w_t = \beta_t$, $\gamma_t = 0$, and $\alpha_t = (1-\pi)w_t + \frac{\phi}{\Lambda}$ and incentivizing disclosure of failure at all times $t \leq T$ maximizes total surplus, minimizes W_0, and therefore maximizes the principal's payoff F_0, implying that a full disclosure contract is optimal.

Part III — Relaxed problem solves full problem

Lemma 12 *Suppose that $\Delta^W(w_0) \geq 0 \iff W_0 \geq w_0$ in the relaxed problem (i.e., with the constraint $\tau^A \geq \tau$). Then, the solution to the relaxed problem solves the full problem in which the agent can fake failure through reporting $\tau^A < \tau$. As such, the optimal contract is a full disclosure contract and always incentivizes disclosure of failure. In addition, $W_t \geq w_t = \beta_t$, $\gamma_t = 0$, and $\alpha_t = (1-\pi)w_t + \phi/(\Lambda p)$ for all $t \leq T$.*

Let the agent's continuation utility under truthful reporting before time t be denoted W_t, given in (2.62). We must show that under the proposed solution $W_t \geq w_t$ for all t or equivalently $W(w) \geq w$ in the state space $[0, w_0]$, in which case the agent is not tempted to fake failure, i.e., is not tempted to set $\tau^A < \tau$. Define $\Delta^W(w) = W(w) - w$, i.e., $\Delta_t^W = W_t - w_t$.

Let $\tau^M = \inf\{t \geq 0 : w_t = 0\}$. If $F_0 < K$ (in the relaxed problem), then $\tau^M = T$ and the principal terminates financing once w reaches zero. Otherwise, when $F_0 > K$, there is no financing deadline so that $T = \infty$ and monitoring occurs whenever $w = 0$ in which case w is reset to $w_0 > 0$. Given that stages repeat when there is monitoring, it suffices to show that $W_t \geq w_t$ for all $t < \tau^M$. Note that for $t < \tau^M$ the agent's continuation value is given by

$$W_t = \int_t^T e^{-\Lambda(s-t)} \Lambda((1-p)(1-\pi)w_s + p\alpha_s)ds,$$

so that

$$\dot{\Delta}_t^W = \dot{W}_t - \dot{w}_t = \Lambda W_t - \Lambda((1-p)(1-\pi)w_t + p\alpha_t) + \phi = \Lambda \Delta_t^W + \Lambda \pi w_t, \quad (2.63)$$

where it was used that $\gamma_t = 0$ and $\alpha_t = (1-\pi)w_t + \phi/(\Lambda p)$ (see Lemma 11). That is, $\dot{\Delta}_t^W > 0$ if $\Delta_t^W > 0$. Hence, $\Delta_t^W \geq 0$ for all $t \geq 0$ if and only if $\Delta_0^W \geq 0$, which concludes the proof.

Part IV — Relaxed problem does not solve full problem

Let us consider that the relaxed problem does not solve the full problem. Let in the following w_0^r the initial value of w in the relaxed problem. The next Lemma proves that when $\Delta^W(w_0^r) < 0$ under the relaxed problem, then a full disclosure contract is not optimal.

Lemma 13 *Let w_0^r the initial value of w under the solution of the relaxed problem. Suppose that $\Delta^W(w_0^r) < 0$ in the relaxed problem. Then, a full disclosure contract is not optimal for the full problem.*

Note that when $\Delta^W(w_0^r) < 0$ in the relaxed problem, then the agent would find it optimal to fake failure when $w = \beta$ is close to w_0^r and $W < w$ (if we removed the constraint $\tau^A \geq \tau$), so the relaxed problem cannot solve the full problem (in which the agent can fake failure). Consider now the full problem (in which the agent can fake failure), and suppose to the contrary that a full disclosure contract is optimal.

As $\Delta^W(w_0^r) < 0$ under the solution of the relaxed problem and as a full disclosure contract is optimal, one of the following two scenarios holds. First, it holds that $\Delta^W(w_0) = 0$ with $w_0 < w_0^r$, whereby the full disclosure contract sets $w_t = \beta_t$, and $\alpha_t = \beta_t(1-\pi) + \phi/(\Lambda p)$ and $\gamma_t = 0$ for all $t \leq T$. This leads to $\dot{\beta}_t = -\phi$ whenever there is no inspection. Denote by $F(w|w_0)$ the principal's value function under the solution for the relaxed problem, given a (potentially sub-optimal) starting value w_0. As $w_0 < w_0^r$ is not optimal in the relaxed problem, it holds that $\frac{dF(w_0|w_0)}{dw_0} > 0$ as well as $\frac{dF(w_0^r|w_0^r)}{dw_0^r} = 0$. The principal's payoff under the solution to the full problem is then $F_0 = F(w_0|w_0)$. As $w_0 < w_0^r$ is not optimal in the relaxed problem, there exists $\varepsilon > 0$ such that the principal can improve her initial payoff (attained under the full problem) from $F_0 = F(w_0|w_0)$ to $F_0^\varepsilon > F_0$ by increasing the initial value of w from w_0 to $w_0 + \varepsilon$ and not incentivizing disclosure of failure (and setting $\beta = \gamma = 0$ and $\alpha = w(1-\pi) + \phi/(\Lambda p)$ so that $\dot{w} = -\phi$) whenever $w \in (w_0, w_0 + \varepsilon)$. The underlying reason is that the likelihood of failure over $w \in (w_0, w_0+\varepsilon)$ is of order ε while the costs of financing the project over $(w_0, w_0 + \varepsilon)$ are of order ε too. Hence, the costs of inefficiently financing the project after its failure over $(w_0, w_0+\varepsilon)$ are bounded from above by $\kappa\Lambda(1-p)(1-\pi)\varepsilon^2 + o(\varepsilon^2) = o(\varepsilon^2)$. As a result, the principal's payoff becomes $F_0^\varepsilon = F_0 + \frac{dF(w_0|w_0)}{dw_0}\varepsilon - o(\varepsilon^2) > F_0$ for $\varepsilon > 0$ sufficiently small, so that the proposed contract cannot be optimal.

Second, it holds that $\Delta^W(w) = 0$ on an interval $[w', w_0]$, and the principal optimally sets $\alpha + \pi\gamma > (1-\pi)w + \frac{\phi}{\Lambda p}$ such that $\dot{W} = \dot{w}$ and $W = w$ on $[w', w_0]$ (e.g., $\alpha = (1 - \pi + \pi/p)w + \phi/(\Lambda p)$ and $\gamma = 0$). Take $\varepsilon > 0$ with $\varepsilon < w_0 - w'$. And, consider the principal changes her strategy by not incentivizing disclosure of failure and setting $\gamma = 0$ and $\alpha = (1-\pi) + \phi/(\Lambda p)$ whenever $w \in (w_0 - \varepsilon, w_0]$, so that $\dot{W} - \dot{w} = \Lambda(W - w) + \Lambda\pi w$ and $W < w$ on this interval. As a result, there exists a constant $\mathcal{K} > 0$ such that $w_0 - W_0 = \mathcal{K}\varepsilon + o(\varepsilon^2) \iff W_0 = w_0 - \mathcal{K}\varepsilon + o(\varepsilon^2)$. The likelihood of failure over $(w_0 - \varepsilon, w_0]$ is of order ε while the costs of financing the project over $(w_0 - \varepsilon, w_0]$ are of order ε too. Hence, the costs of inefficiently financing the project after its failure over $(w_0 - \varepsilon, w_0]$ are bounded from above by $\kappa\Lambda(1-p)(1-\pi)\varepsilon^2 + o(\varepsilon^2) = o(\varepsilon^2)$. As such, the potential inefficiency of this strategy is of order ε^2 (i.e., the effect on total surplus at time zero $\mathcal{S}_0$) and therefore negligible for $\varepsilon > 0$ sufficiently small. At the same time, this alternative strategy reduces the agent's continuation payoff at time zero from w_0 to $w_0 - \mathcal{K}\varepsilon + o(\varepsilon^2)$. As a result, for $\varepsilon > 0$ sufficiently small, following the alternative strategy and not incentivizing disclosure over $(w_0 - \varepsilon, w_0]$ increases the principal's payoff $F_0 = \mathcal{S}_0 - W_0$ (which is total surplus $\mathcal{S}_0$ minus the agent's payoff W_0) by $\mathcal{K}\varepsilon + o(\varepsilon^2) > 0$, so that the proposed contract cannot be optimal. Thus, a full disclosure contract is not optimal.

Lemma 14 *Suppose that $\Delta_t^W \geq 0$ (i.e., $W_t \geq w_t$). Then, up to the next inspection or up to termination (whichever happens first), the optimal contract incentivizes disclosure of failure.*

Let $\tau_t^M = \inf\{s \geq t : w_s = 0\}$. At time τ_t^M, the principal either terminates the project (when $F_0 < K$) or inspects the project (when $F_0 \geq K$) in which case w is reset to w_0. The principal sets $\beta_s = w_s$, $\gamma_s = 0$, and $\alpha_s \geq w_s(1-\pi) + \phi/(\Lambda p)$, so that $\dot{w}_s = \dot{\beta}_s = -\phi$ until τ_t^M, thereby providing incentives for effort and incentives not to delay disclosure of failure. Also note that up to time τ_t^M, $\dot{w}_s = -\phi$, so that — given W_t and w_t — the principal maximizes τ_t^M (which is optimal to minimize monitoring costs or to minimize deadweight losses associated with termination). If $\Delta_s^W \geq 0$ holds up to time τ_t^M, then the agent has no incentives to fake failure and, therefore, has incentives to disclose failure at the time it occurs (since $\dot{\beta}_s = -\phi$). The agent's continuation value reads

$$W_t = \int_t^T e^{-\Lambda(s-t)}\Lambda((1-\pi)w_t + p(\alpha_s - (1-\pi)w_s))ds,$$

so that

$$\dot{\Delta}_t^W = \dot{W}_t - \dot{w}_t = \Lambda W_t - \Lambda((1-\pi)w_t + p(\alpha_t - (1-\pi)w_t)) + \phi \leq \Lambda\Delta_t^W + \Lambda\pi w_t, \tag{2.64}$$

where it was used that $\alpha_t \geq (1-\pi)w_t + \phi/(\Lambda p)$. That is, $\dot{\Delta}_t^W \geq 0$ can be achieved if $\Delta_t^W \geq 0$, while respecting effort incentive compatibility $\alpha_t \geq (1-\pi)w_t + \phi/(\Lambda p)$. Hence, it follows that pay for success $\{\alpha_s\}_{s\geq t}$ can be chosen such that $\Delta_s^W \geq 0$ for all $s \in [t, \tau_t^M]$, so the agent does not want to fake failure. That is, by providing incentives for effort up to time τ_t^M and stipultating $\beta_s = w_s$, the principal also incentivizes truthful disclosure of failure up to time τ_t^M and maximizes the time τ_t^M, thereby maximizing the total surplus $\mathcal{S}_t$ while delivering continuation payoff W_t to the agent. Because the principal's payoff at time t is $\mathcal{S}_t - W_t$ (i.e., total surplus minus the agent's payoff), it must be optimal for the principal to incentivize disclosure of failure up to time τ_t^M.

Part V — Solution to the full problem

Lemma 15 *Let w_0^r the initial value of w under the solution of the relaxed problem. The solution of the full problem is as follows. There exists $w^c \in [0, w_0]$ such that the optimal contract incentivizes disclosure of failure whenever $w \leq w^c$ and does not incentivize disclosure of failure whenever $w \in (w^c, w_0]$. Over time, w drifts down with $\dot{w} = -\phi$. Once w reaches zero, the principal terminates financing (when $F_0 < K$) or conducts an inspection (when $F_0 \geq K$) in which case w is reset to w_0 and the principal's continuation payoff becomes F_0.*

If $\Delta^W(w_0^r) \geq 0$, a full disclosure contract is optimal for the full problem so that $w^c = w_0$ and $W_0 \geq w_0$. If $\Delta^W(w_0^r) < 0$, a full disclosure contract is not optimal for the full problem so that $w^c < w_0$ and $W_0 < w_0$. When K is sufficiently small, the optimal contract is a full disclosure contract so that $w^c = w_0$. When K is sufficiently large, the optimal contract does not always incentivize disclosure of failure so that $w^c < w_0$.

First, Corollary 3 shows that whenever there is an inspection at deterministic date τ^M, then $\lim_{t\uparrow\tau^M} w_t = 0$. On the other hand, when w reaches zero, the principal must either inspect or terminate the project. Thus, the principal inspects or terminates the project whenever w reaches zero; and, there are no inspections when $w > 0$. By Corollary 4, the agent has incentives to disclose failure truthfully just before inspection/termination. Thus, if it occurs, an inspection (at $w = 0$) yields that the project has not failed yet and therefore the principal's continuation payoff after the inspection becomes F_0. Hence, the principal indeed prefers an inspection over termination if and only if $F_0 \geq K$. On the other hand, if $F_0 < K$, the principal terminates financing when w reaches zero and all the results follow from Proposition 6. Unless otherwise specified, it suffices to consider the case $F_0 \geq K$.

Second, Lemma 14 implies that once the principal incentivizes disclosure of failure at some time t, then the principal optimally incentivizes disclosure of failure at all times $s \in [t, \tau_t^M]$, whereby $\tau_t^M = \inf\{s \geq t : w_s = 0\}$ is the next time w reaches zero. And, $\dot{w}_s = -\phi$ for all $s \in [t, \tau_t^M]$. Also note that w is the state variable and summarizes all payoff relevant information. Thus, the set of values of w on which the contract incentivizes disclosure of failure is an interval that includes zero. As a result, there exists $w^c \in [0, w_0]$ such that the contract incentivizes disclosure of failure if and only if $w \leq w^c$. And, the contract does not incentivize disclosure of failure if and only if $w \in (w^c, w_0]$.

Third, Lemma 13 implies that when $\Delta^W(w_0^r) < 0$ in the relaxed problem, then the optimal contract in the full problem does not always incentivize disclosure of failure, so that $w^c < w_0$ and $W_0 < w_0$. By contrast, when $\Delta^W(w_0^r) \geq 0$, Lemma 12 implies that the optimal contract is a full disclosure contract, so that $w^c = w_0$ and $W_0 \geq w_0$. That is, the contract always incentivizes disclosure of failure if and only if $W_0 \geq w_0$.

Finally, when $K = 0$, the principal inspects the project costlessly at all times, so that $W_0 > w_0 = 0$. As $K > 0$ tends to zero, the solution converges to the solution with $K = 0$, whereby $W_0 > w_0 = 0$. By continuity, it holds that when $K > 0$ is sufficiently small, then $W_0 > w_0$, so that by virtue of our previous findings the optimal contract is a full disclosure contract with inspections whenever w reaches zero. Finally, when K is sufficiently large so that $F_0 < K$, there is no monitoring and the optimal contract is described in Proposition 6. According to Proposition 6, the optimal contract then does not always incentivize disclosure of failure.

Part VI — Deterministic monitoring is optimal

In the general formulation of the contracting problem, the principal's value function F_t depends on the state variables w_t and q_t, where q_t is the principal's belief of whether the project has failed so far. We can express the principal's value function as function of (w, q) in that $F_t = F(w_t, q_t)$. Note that when the principal deviates from the proposed strategy of Proposition 4 and inspects the project in state (w, q), then with probability $1 - q$ the project has not failed and her continuation payoff is F_0, while with probability q the project has failed and continuation payoff is zero. As a result, the net gain of deviating from the proposed strategy and inspecting the project in state (w, q) is $F_0(1 - q) - K - F(w, q)$. Note that if the principal deviates and inspects the project in state (w, q), this inspection is unexpected by the agent. Because inspecting the project is always an option, the principal's value function satisfies

$$F(w, q) \geq F_0(1 - q) - K \tag{2.65}$$

for $w \in (0, w_0)$. If a stochastic inspection at $w > 0$ is optimal, it must be that the value matching condition

$$F(w, q) = F_0(1 - q) - K \tag{2.66}$$

holds, in that (2.65) holds in equality, and the smooth pasting condition

$$F_w(w, q) = 0 \tag{2.67}$$

holds. Importantly, at all times $t > 0$ with $\theta_s = \Theta_s = 0$ for all $s < t$ at which $F(w, q)$ is differentiable and $0 < w_t < w_0$, the value function satisfies $F_w(w, q) > 0$, as otherwise the principal could improve her payoff at time zero by decreasing the initial value of w (i.e., w_0). To see this formally, suppose to the contrary that under the optimal contract $\mathcal{C}$ at time t with $q_t = q$ and $w_t = w < w_0$ as well as $\theta_s = \Theta_s = 0$ for $s < t$, there exists $w' < w_t$ such that $F(w', q_t) \geq F(w_t, q_t)$. Then, the principal could improve her payoff at time 0 by writing a contract $\mathcal{C}'$ that sets rewards for public failure $\gamma_s' = 0$ and rewards for success $\alpha_s' < \alpha_s$ up to time t, and it changes the initial value of w from w_0 to w_0' via $w_0' = w_0 - (w_t - w') < w_0$, thereby inducing a different path of w denoted w'. Under both contracts the belief q_s is the same and w_s decreases with $\dot{w}_s = \dot{w}'_s = -\phi$ for $s \leq t$. Under the alternative contract $\mathcal{C}'$, w'_s reaches the value w' at time $s = t$ (if no termination before). As the alternative contract stipulates strictly lower w (i.e., $w'_s < w_s$ for all $s \leq t$), it is possible to find rewards for success $\alpha'_s < \alpha_s$, while meeting incentive compatibility. If termination occurs before time t, the principal realizes on average strictly higher payoff under the contract $\mathcal{C}'$. At time t, the contract $\mathcal{C}'$ also yields higher continuation payoff $F(w', q_t) \geq F(w_t, q_t)$. Thus, the contract $\mathcal{C}'$ yields strictly higher payoff for the principal at time zero than the contract $\mathcal{C}$, a contradiction. Therefore, $F_w(w, q) > 0$. As w is reset to w_0 after an inspection, the result $F_w(w_t, q_t) > 0$ extends to all times t with $\theta_s = \Theta_s = 0$ for $s \in (t^M, t)$ where t^M is the last inspection date before time t or zero.

Suppose that there exists time $t > 0$ with $w_t > 0$, and a stochastic inspection at time t is optimal. Consider the first time $t > 0$ at which a stochastic inspection is optimal in that $\theta_s = \Theta_s = 0$ for $s \in (t^M, t)$, where t^M is the last inspection date before time t or zero. Our previous arguments imply $F_w(w_t, q_t) > 0$. However, when stochastic monitoring is optimal, the value matching condition (2.66) and smooth pasting condition (2.67), i.e., $F_w(w_t, q_t) = 0$, must hold, which cannot be as $F_w(w_t, q_t) > 0$. Thus, stochastic monitoring is not optimal.

2.10.6 Proof of Corollary 1

Suppose that $\pi = 0$. In the limit case $\pi = 0$, the optimal contract is a full disclosure contract with $t_1 = 0$ and closed form expressions (see (2.17)):

$$w_0 = \frac{\phi}{\Lambda} \ln\left(\frac{\Lambda\mu p - \kappa}{\phi}\right);$$
$$T = \frac{1}{\Lambda} \ln\left(\frac{\Lambda\mu p - \kappa}{\phi}\right)$$

Calculate

$$\frac{\partial T}{\partial \Lambda} = -\frac{T}{\Lambda} + \frac{\mu\phi}{\Lambda\mu p - \kappa}.$$

Recall that parameters satisfy $\Lambda\mu p - \kappa > \phi$. For Λ sufficiently small such that $\Lambda\mu p - \kappa > \phi$ (i.e., for $\Lambda \downarrow (\phi + \kappa)/(\mu p)$) it follows that $T \downarrow 0$ so that $\frac{\partial T}{\partial \Lambda} > 0$ for Λ sufficiently small or equivalently $1/\Lambda$ sufficiently large.

As it is well known that $\lim_{x\to\infty} \ln(x)/x = 0$, the limit $\Lambda \to \infty$ yields

$$\lim_{\Lambda\to\infty} T = 0.$$

Hence, T decreases in Λ for Λ sufficiently large or equivalently increases in $1/\Lambda$ for $1/\Lambda$ sufficiently small. Also note that T decreases in ϕ.

Because of $\beta_{t_1} = T\phi$, it follows that β_{t_1} increases (decreases) in $1/\Lambda$ for $1/\Lambda$ sufficiently small (large). As it is well known that the function $x\ln(1/x)$ is hump shaped in x (for positive x), it follows that $\beta_{t_1} = \frac{\phi}{\Lambda}\ln\left(\frac{\Lambda\mu p - \kappa}{\phi}\right)$ is hump-shaped in ϕ and hence β_{t_1} increases (decreases) in ϕ for ϕ sufficiently small (large). Moreover, it is easy to see that T decreases in ϕ. By continuity, all the established results hold true for π sufficiently small.

2.11 Solution details for Section 2.5

2.11.1 Preliminaries

We solve the model under the specification presented in Section 2.5; that is, a failed project can produce (terminal) payoff μ at intensity λ if it continues to be operated (at flow cost κ) after failure. We assume that parameters satisfy $\lambda < \Lambda p$. Thus, a failed project has NPV $\mu^f = \mu - \kappa/\lambda$, and financing a failed project until payoff μ realizes generates surplus $\mu^f + \phi/\lambda$, as the agent derives flow benefits ϕ when the principal finances a failed project. Observing payoff μ, the principal cannot detect whether the payoff stems from a failed project or from success. When payoff μ realizes, the agent receives then pay for success α (provided failure has not become known to the principal yet).

Unless otherwise specified and in particular in Sections 2.11.4 through 2.11.9 we assume that $\mu^f < -\frac{\phi}{\lambda} \iff \lambda\mu < \kappa - \phi$, so financing a failed project is inefficient. Section 2.11.10 studies the optimal contract when $\mu^f \geq -\frac{\phi}{\lambda} \iff \lambda\mu \geq \kappa - \phi$. Importantly, Sections 2.11.9 and 2.11.10 highlight that the agent is paid for publicly observed failure and $\gamma_t > 0$ is possible only if $\lambda\mu < \kappa - \phi$. Throughout, we assume that parameters are such that the principal optimally implements full effort and terminates financing once failure becomes known to her (e.g. via a report or via observed failure). In addition, we consider that the agent is only paid at time τ^A, in that (2.26) holds.

We also introduce that the agent can always generate publicly observed failure yielding payouts γ_t to the agent. Because the agent can always generate publicly observed failure and the agent can always report privately observed failure which is not verifiable, it must hold that $\gamma_t \leq W_t$ and $\gamma_t \leq \beta_t$. If $\gamma_t > W_t$, the agent would generate failure on purpose which is inefficient. If $\gamma_t > \beta_t$, the agent would never report privately observed failure (which yields pay β_t) but rather generate publicly observed failure which yields pay $\gamma_t > \beta_t$; in this case, the contract leads then to the same outcomes as if $\gamma_t = \beta_t$. In addition, to solve for the optimal contract, we consider that the termination time T is deterministic.[29]

2.11.2 Second-best benchmark

We start by considering the benchmark in which failure is publicly observable and contractible. Then, as in Section 2.3.1, the setting and the optimal contract are time-stationary, and the agent is paid for success but not for failure (i.e., $\gamma = \beta = 0$). The principal finances the project at least until its completion at time τ.

When the principal terminates financing the project upon failure, then the solution and the optimal contract are described in Proposition 6, leading to initial payoff

$$F_0 = p\mu - \frac{\kappa + \phi}{\Lambda} \tag{2.68}$$

for the principal.

Suppose now the principal finances the project after its failure until the payoff μ realizes at time τ^λ. Then, the agent derives private flow benefits from shirking until time τ^λ, when the project fails. That is, the agent realizes expected payoff $\hat{\gamma} = \frac{\phi}{\lambda}$

[29] Likely, one can adapt the arguments presented in the proof of Lemma 9 to verify that stochastic termination is indeed not optimal.

upon failure, derived from flow benefits of shirking over $[\tau, \tau^\lambda]$. Analogous to (2.4), to incentivize full effort $a_t = 1$, it must be that

$$\Lambda(p\alpha + (1-p)\hat{\gamma}) \geq \phi + \Lambda\hat{\gamma} \iff \Lambda p(\alpha - \hat{\gamma}) \geq \phi.$$

In optimum, the incentive constraint is tight, so that

$$\alpha = \frac{\phi}{\Lambda p} + \hat{\gamma}.$$

The principal's payoff at time zero is

$$\tilde{F}_0 = p(\mu - \alpha) + (1-p)\mu^f - \frac{\kappa}{\Lambda} = \mu - \frac{\phi + \kappa}{\Lambda} + (1-p)\mu^f - \frac{p\phi}{\lambda} = F_0 + (1-p)\mu^f - \frac{p\phi}{\lambda}.$$

The principal optimally terminates financing upon failure if and only if

$$\tilde{F}_0 < F_0 \iff \mu^f < \frac{p\phi}{\lambda(1-p)}.$$

Note that in the second-best benchmark, the principal may commit to terminate financing of a failed project in order to increase the agent's effort incentives and to reduce agency costs. This may be the case even if a failed project has positive net present value.

2.11.3 Incentives

We characterize the agent's incentives, when $\lambda > 0$. To start with, recall that the agent's continuation payoff for times $t < T \wedge \tau$ is defined as

$$W_t = \int_t^T e^{-\Lambda(s-t)} \Lambda\big((1-p)((1-\pi)w_s + \pi\gamma_s) + p\alpha_s\big) ds, \tag{2.69}$$

so that differentiation with respect to t yields

$$\dot{W}_t = \Lambda W_t - \Lambda\big((1-p)((1-\pi)w_t + \pi\gamma_t) + p\alpha_t\big). \tag{2.70}$$

As is already derived in Section 2.5, the agent does not delay disclosure of failure if and only if

$$\dot{\beta}_t \leq -\phi - \lambda(\alpha_t - \beta_t), \tag{2.71}$$

which implies in optimum

$$\dot{w}_t = -\phi - \lambda(\alpha_t - w_t). \tag{2.72}$$

Specifically, if the principal terminates financing upon failure, one can write

$$w_t := \max_{\tau^A \in [t,T]} \left[\int_t^{\tau^A} (\phi + \lambda(\alpha_s - w_s)) ds + \beta_{\tau^A} \right],$$

whereby $\tau^A = t = \tau$ under a full disclosure contract, i.e., whenever the principal incentivizes disclosure of failure.

Provided the principal terminates the project after publicly observed failure, the incentive compatibility condition for effort is given by

$$\alpha_t \geq w_t(1-\pi) + \gamma_t \pi + \frac{\phi}{\Lambda p}, \tag{2.73}$$

where γ_t is the agent's pay for publicly observed failure. To incentivize the agent not to fake failure, it must be that $W_t \geq \beta_t$, in that (2.7) is met like in the baseline version of the model.

As stipulating rewards for success α_t creates incentives for the agent to hide failure and therefore exacerbates moral hazard over disclosure of failure, the optimal full disclosure contract (when $\mu^f < -\frac{\phi}{\lambda}$) minimizes rewards for success and boosts the agent's continuation payoff (in order to satisfy $W_t \geq \beta_t$) by stipulating rewards for publicly observed failure $\gamma_t \in (0, \beta_t]$. Because $\gamma_t \leq \beta_t$, the principal stipulates rewards for publicly observed failure $\gamma_t = \beta_t$ within a full disclosure contract.

2.11.4 Solution: full disclosure contract with $\mu^f < -\frac{\phi}{\lambda}$

Note that when the project has failed at time t, then total continuation surplus is

$$\mathcal{S}_t = \int_t^{\hat{T}} e^{-\lambda(s-t)}(\lambda\mu - \kappa + \phi)ds,$$

if the principal finances the project until time $\hat{T} > t$. As $\mu^f < -\frac{\phi}{\lambda}$, it follows that $\lambda\mu < \kappa - \phi$ and $\mathcal{S}_t < 0$. Thus, when $\mu^f < -\frac{\phi}{\lambda}$, the principal optimally terminates financing when failure becomes known to her (e.g., via a report or via observed failure). Unless otherwise mentioned and particular in Sections 2.11.4 to 2.11.9, we consider that $\mu^f < -\frac{\phi}{\lambda}$. Section 2.11.10 studies the case $\mu^f \geq -\frac{\phi}{\lambda}$ in detail.

In what follows, we define *adjusted rewards for success* $\delta_t := \alpha_t - w_t$ and $\Delta_t^W := W_t - w_t$, whereby incentive compatibility for effort requires $\delta_t \geq \phi/(\Lambda p)$. We rewrite

$$\dot{w}_t = -\phi - \lambda\delta_t$$

and

$$\dot{W}_t = \Lambda(W_t - w_t) - \Lambda p \delta_t + \Lambda(1-p)\pi(w_t - \gamma_t),$$

so that

$$\dot{\Delta}_t^W = \Lambda\Delta_t^W - (\Lambda p - \lambda)\delta_t + \phi + \Lambda(1-p)\pi(w_t - \gamma_t).$$

We can integrate the above ODE over time to obtain

$$\Delta_t^W = \int_t^T e^{-\Lambda(s-t)}((\Lambda p - \lambda)\delta_s - \phi - \Lambda(1-p)\pi(w_s - \gamma_s))ds$$

or equivalently

$$\Delta_t^W = \int_t^T ((\Lambda p - \lambda)\delta_s - \phi - \Lambda\Delta_s^W - \Lambda(1-p)\pi(w_s - \gamma_s))ds.$$

A full disclosure contract with deadline T implies that $\beta_t = \gamma_t = w_t$, $w_T = \beta_T = 0$, and, optimally, $W_t = w_t$ for all t, so that $\Delta_t^W = 0$. We verify below that $W_t = w_t$ indeed is optimal.

Thus, we can use $\gamma_t = w_t$ and $\Delta_t^W = 0$ and solve $\dot{\Delta}_t^W = 0$ for

$$\delta_t = \alpha_t - w_t = \frac{\phi}{\Lambda p - \lambda} \iff \alpha_t = w_t + \frac{\phi}{\Lambda p - \lambda},$$

so the incentive condition with respect to effort is slack and

$$\dot{w}_t = -\phi - \lambda\delta_t = -\phi - \frac{\lambda\phi}{\Lambda p - \lambda} = -\phi\left(\frac{\Lambda p}{\Lambda p - \lambda}\right).$$

Note that implementing the full disclosure contract is expensive for the principal because it requires i) rewards for publicly observed failure $\gamma_t = \beta_t$ and ii) excessively high rewards for success α_t.

Finally, we provide an argument that within the optimal full disclosure contract, $\gamma_t = \beta_t = w_t$ is indeed optimal. To see this, suppose to the contrary that $\gamma_t < \beta_t$. Then, the principal could increase γ_t up to β_t so that $\gamma_t = \beta_t$ and decrease α_t while inducing full effort $a_t = 1$ and leaving the agent's continuation payoff unchanged, which relaxes the incentive constraint (2.25) (i.e., (2.71) or (2.72)), increases $\dot{w}$, and improves the principal's payoff as the value function $F(w)$ is upward sloping. Thus, $\gamma_t = \beta_t$ is optimal within a full disclosure contract.

2.11.5 Verification: optimal full disclosure contract with $\mu^f < -\frac{\phi}{\lambda}$

Consider any time t with with $W_t = w_t$ and suppose the principal implements a full disclosure contract from time t onwards. We verify that the full disclosure (continuation) contract that is described above and implements $\gamma_s = 0$, $W_s = w_s = \beta_s$ and $\delta_s = \frac{\phi}{\Lambda p - \lambda}$ for all $s \geq t$ is indeed optimal.

Note that at time t when $W_t = w_t$, the time to deadline is $T - t = \inf\{s \geq t : w_s = 0\} - t$. At time t, the principal chooses δ_s for $s \geq t$, in order to maximize total surplus

$$\mathcal{S}_t = \int_t^T e^{-\Lambda(s-t)}(\Lambda p\mu - \kappa)ds,$$

subject to the requirement of delivering continuation payoff W_t to the agent and to meet incentive compatibility with respect to effort (i.e., $\delta_s \geq \phi/(\Lambda p)$) and disclosure, i.e., $W_s \geq w_s$ for $s \geq t$ (and $\beta_s = w_s$). The total surplus is split between principal and agent, so the principal's payoff is $F_t = \mathcal{S}_t - W_t$. Observe that maximizing total surplus is equivalent to maximizing time to deadline $T - t$, given state variables w_t and W_t.

Next, note that $w_t = \int_t^T (\phi + \lambda\delta_s)ds = \phi(T-t) + \lambda\int_t^T \lambda_s ds$, so $\phi(T-t) = w_t - \lambda\int_t^T \delta_s ds$. Therefore, maximizing time to deadline $T - t$ is akin to minimizing $\int_t^T \delta_s ds$ subject to incentive compatibility with respect to effort and $W_s \geq w_s \iff \Delta_s^W \geq 0$ for all $s \geq t$ starting with $\Delta_t^W \geq 0$ (and subject to delivering continuation payoff W_t to the agent). As such, within a full disclosure contract from time t onwards, the principal solves the problem

$$\min_{\{\delta_s\}_{s\geq t}} \int_t^T \delta_s ds$$
$$\text{s.t. } \delta_s \geq \phi/(\Lambda p)$$
$$\text{s.t. } \Delta_s^W \geq 0$$
$$\text{s.t. } \Delta_t^W = \int_t^T e^{-\Lambda(s-t)}((\Lambda p - \lambda)\delta_s - \phi)ds = \int_t^T ((\Lambda p - \lambda)\delta_s - \phi - \Lambda\Delta_s^W)ds$$
$$\text{s.t. } W_t = \int_t^T e^{-\Lambda(s-t)}\Lambda(w_s + p\delta_s)ds$$

Note that Δ_t^W is taken as given when solving the optimization problem, determined by the state variables W_t and w_t. The penultimate constraint can be rewritten as

$$\int_t^T \delta_s ds = \frac{1}{\Lambda p - \lambda}\left(\Delta_t^W + \int_t^T (\phi + \Lambda\Delta_s^W)ds\right).$$

As a result, minimizing $\int_t^T \delta_s ds$ is akin to minimizing Δ_s^W for all $s \geq t$ (i.e., to minimizing $\int_t^T \Delta_s^W ds$) subject to $\Delta_s^W \geq 0$. Thus, the full disclosure contract described above is indeed optimal, because it achieves $\Delta_s^W = 0$ for all $s \geq t$, implements full effort, and delivers continuation payoff W_t to the agent.

2.11.6 Unconditional financing and deadline extensions when $W_t = w_t$

Recall that at time t with $W_t = w_t$, the principal maximizes surplus subject to all relevant incentive compatibility constraints and the requirement to deliver continuation payoff W_t to the agent.

Denote by $T^D - t$ the time to deadline that pertains if the principal follows from time t onwards the optimal full disclosure contract described above. Also denote the payoff for privately observed failure under the full disclosure contract by w^D. As a full disclosure contract stipulates adjusted rewards for success $\delta_s^D = \phi/(\Lambda p - \lambda)$ and therefore $\dot{w}_s^D = -\phi - \lambda\phi/(\Lambda p - \lambda) = -\phi\left(\frac{\Lambda p}{\Lambda p - \lambda}\right)$, it follows that that

$$T^D - t = \frac{w_t^D(\Lambda p - \lambda)}{\phi\Lambda p}$$

Next, consider that the principal decides not to incentivize disclosure of failure from time t onwards for $\Delta \geq 0$ units of time, and the principal follows the optimal full disclosure contract from time $t + \Delta$ onwards whereby $W_{t+\Delta} = w_{t+\Delta}$ and $\Delta \leq T^U$. Observe that at time t, the principal maximizes time to deadline and achieves time to deadline $T^U - t$ by following this alternative strategy, i.e., by not incentivizing disclosure over $[t, t+\Delta)$ and incentivizing disclosure of failure afterwards over $[t + \Delta, T^U]$. It follows that $T^U - t - \Delta = \frac{w_{t+\Delta}(\Lambda p - \lambda)}{\phi\Lambda p}$. Let the deadline extension that is achieved by the alternative strategy be denoted $\Delta' := T^U - T^D$. We calculate that

$$\Delta' = T^U - T^D = \frac{\Lambda p - \lambda}{\phi\Lambda p}(w_{t+\Delta} - w_{t+\Delta}^D), \tag{2.74}$$

where $w^D_{t+\Delta}$ is the agent's payoff for privately observed failure at time $t+\Delta$ under the full disclosure contract and $w_{t+\Delta}$ is the agent's payoff for privately observed failure at time $t+\Delta$ under the alternative strategy.

Over the time interval $[t, t+\Delta]$, it must hold that $\gamma_s \leq W_s$, as otherwise the agent would precipitate failure (which is clearly inefficient). The incentive constraint for effort becomes $\delta_s \geq \phi/(\Lambda p) - \pi(w_s - \gamma_s)$. The principal chooses $\delta_s \geq \phi/(\Lambda p) - \pi(w_s - \gamma_s)$ to maximize the deadline extension $\Delta' = T^U - T^D$ or equivalently the deadline T^U given w_t and W_t. As $(T^U - t - \Delta) \propto w_{t+\Delta}$, this is akin to maximizing $w_{t+\Delta}$ given w_t. Since $w_{t+\Delta} = w_t - \phi\Delta - \lambda \int_t^{t+\Delta} \delta_s ds$, the principal chooses $\delta_s \geq \phi/(\Lambda p) - \pi(w_s - \gamma_s)$ to minimize $\int_t^{t+\Delta} \delta_s ds$, subject to the constraint

$$\int_t^{t+\Delta} ((\Lambda p - \lambda)\delta_s - \phi - \Lambda\Delta_s^W - \Lambda(1-p)\pi(w_s - \gamma_s))ds = 0. \tag{2.75}$$

Note that the above constraint (2.75) arises due to $W_{t+\Delta} - w_{t+\Delta} = W_t - w_t = 0 \iff \Delta^W_{t+\Delta} = \Delta^W_t = 0$ and $\Delta^W_t - \Delta^W_{t+\Delta} = \int_t^{t+\Delta}((\Lambda p - \lambda)\delta_s - \phi - \Lambda\Delta^W_s - \Lambda(1-p)\pi(w_s - \gamma_s))ds$.

We can solve (2.75) for

$$\int_t^{t+\Delta} \delta_s = \frac{1}{\Lambda p - \lambda}\left(\int_t^{t+\Delta} (\phi + \Lambda\Delta_s^W + \Lambda(1-p)\pi(w_s - \gamma_s))ds\right)$$

By contrast, the full disclosure contract, stipulating adjusted rewards for success $\delta^D_s = \frac{\phi}{\Lambda p - \lambda}$, $\gamma_s = w^D_s$, and $\Delta^W_s = 0$, implements

$$\int_t^{t+\Delta} \delta^D_s ds = \frac{1}{\Lambda p - \lambda}\int_t^{t+\Delta} \phi ds = \frac{\phi\Delta}{\Lambda p - \lambda}.$$

In addition,

$$w_{t+\Delta} = w_t - \phi\Delta - \lambda\int_t^{T+\Delta} \delta_s ds \quad \text{and} \quad w^D_{t+\Delta} = w_t - \phi\Delta - \lambda\int_t^{T+\Delta} \delta^D_s ds.$$

As a result and due to (2.74), the deadline extension satisfies

$$\begin{aligned}\Delta' = \frac{\Lambda p - \lambda}{\phi\Lambda p}(w_{t+\Delta} - w^D_{t+\Delta}) &= \frac{\lambda(\Lambda p - \lambda)}{\phi\Lambda p}\left(\int_t^{t+\Delta} \delta^D_s ds - \int_t^{t+\Delta} \delta_s ds\right) \\ &= \frac{\lambda}{\phi\Lambda p}\left(-\int_t^{t+\Delta} \Lambda\left[\Delta^W_s + (1-p)\pi(w_s - \gamma_s)\right] ds\right) \\ &\leq \frac{\lambda(1-(1-p)\pi)}{\phi\Lambda p}\left(-\int_t^{t+\Delta} \Lambda\Delta^W_s ds\right),\end{aligned}$$

where the inequality uses that $\gamma_s \leq W_s$ so that $\Delta_s^W + (1-p)\pi(w_s - \gamma_s) \geq \Delta_s^W(1-(1-p)\pi)$. Due to $\Delta_{t+\Delta}^W = 0$, it holds for $s \in [t, t+\Delta]$

$$
\begin{aligned}
\Delta_s^W &= \int_s^{t+\Delta} e^{-\Lambda(u-s)}((\Lambda p - \lambda)\delta_u - \phi - \Lambda(1-p)\pi(w_u - \gamma_u))du \\
&\geq -\int_s^{t+\Delta} e^{-\Lambda(u-s)} \left(\frac{\lambda\phi}{\Lambda p} + [\Lambda(1-p) + (\Lambda p - \lambda)]\pi(w_u - \gamma_u) \right) du \\
&\geq -\int_t^{t+\Delta} e^{-\Lambda(u-t)} \left(\frac{\lambda\phi}{\Lambda p} + \Lambda\pi \left(p\mu - \frac{\kappa}{\Lambda} \right) \right) du \\
&= -\frac{\mathcal{Y} \cdot (\phi\Lambda p)}{\Lambda(1-(1-p)\pi)}(1 - e^{-\Lambda\Delta}),
\end{aligned}
$$

where the first inequality uses incentive compatibility for effort, $\delta_u \geq \phi/(\Lambda p) - \pi(w_u - \gamma_u)$, and the second inequality uses that $w_u - \gamma_u \leq w_u \leq w_t = W_t \leq F_t + W_t \leq p\mu - \kappa/\Lambda$ as well as $\Lambda(1-p) + \Lambda p - \lambda \leq \Lambda$. And, we define

$$
\mathcal{Y} := \frac{(1-(1-p)\pi)}{\phi\Lambda p} \cdot \left(\frac{\lambda\phi}{\Lambda p} + \Lambda\pi \left(p\mu - \frac{\kappa}{\Lambda} \right) \right) > 0. \tag{2.76}
$$

As a result,

$$
\Delta' \leq \frac{\lambda(1-(1-p)\pi)}{\phi\Lambda p} \left(-\int_t^{t+\Delta} \Lambda\Delta_s^W ds \right) \leq \lambda\mathcal{Y}\Delta \left(1 - e^{-\Lambda\Delta} \right) =: \overline{\Delta}. \tag{2.77}
$$

2.11.7 Full disclosure continuation contract optimal when $W_t = w_t$

Taking stock, at time t, not incentivizing disclosure over a period Δ brings a deadline extension Δ' bounded from above by $\overline{\Delta}$ in (2.77). Let us analyze the effect of this alternative strategy on total surplus $\mathcal{S}_t$, whereby the principal's payoff is $F_t = \mathcal{S}_t - W_t$. Note that when the net effect of this alternative strategy on total surplus at time t is negative, the full disclosure continuation contract is preferred over the alternative strategy, as at time t the principal maximizes total surplus subject to all relevant incentive constraints and subject to delivering continuation payoff W_t to the agent.

The only potential benefit of the alternative strategy (relative to the full disclosure contract from time t onward) realizes when the project does not complete until deadline T^D (with probability $e^{-\Lambda(T^D-t)}$), but succeeds within the extension period Δ' (with probability $(1-e^{-\Lambda\Delta'})p$), leading to benefits bounded from above by

$$
B^+ = e^{-\Lambda(T^D-t)}(1-e^{-\Lambda\Delta'})p\mu \leq (1-e^{-\Lambda\Delta'})p\mu \leq (1-e^{-\Lambda\overline{\Delta}})p\mu.
$$

A cost of the alternative strategy is that the project may hiddenly fail over $[t, t+\Delta]$ (so that failure is privately observed by the agent) and the financing costs are expended inefficiently after failure until time $t+\Delta$, so the costs of the alternative strategy are bounded from below by

$$
C^- = (1-e^{-\Lambda\Delta})(1-p)(1-\pi)((\Delta - \bar{t})(\kappa - \phi) - (1-e^{-\lambda\Delta})\mu).
$$

Here, $\Delta - \bar{t}$ is the average amount of time that a failed project is (inefficiently) continued, and recall that the failed project yields payoff μ over Δ with probability of at most $1 - e^{-\lambda\Delta}$. Note that

$$
\bar{t} = \mathbb{E}_t\left[\tau - t | \text{Project hiddenly fails over } [t, t+\Delta]\right] = (1-p)(1-\pi)\mathbb{E}_t\left[\tau - t | \tau < t+\Delta\right],
$$

where the second equality reflects that the project hiddenly fails with probability $(1-p)(1-\pi)$ upon completion at time τ. To derive $\bar{t}$, consider without loss of generality $t = 0$ and establish a more general result regarding an exponentially distributed time τ with intensity Λ.

Note that $\mathbb{E}_t\tau = 1/\lambda$ (absent frictions) and the memoryless property of τ implies $\mathbb{E}_t[\tau|\tau > t+\Delta] = 1/\lambda + \Delta$ (absent frictions). Thus,

$$\frac{1}{\Lambda} = \mathbb{E}_t[\tau] = \mathbb{P}_t(\tau < t+\Delta)\mathbb{E}_t[\tau|\tau < t+\Delta] + \mathbb{P}_t(\tau > t+\Delta)\mathbb{E}_t[\tau|\tau > t+\Delta]$$
$$= (1-e^{-\Lambda\Delta})\mathbb{E}_t[\tau|\tau < t+\Delta] + e^{-\Lambda\Delta}\left(\frac{1}{\Lambda}+\Delta\right).$$

Thus,

$$(1-e^{-\Lambda\Delta})\mathbb{E}_t[\tau|\tau < t+\Delta] = \frac{1-e^{-\Lambda\Delta}}{\Lambda} - e^{-\Lambda\Delta}\Delta.$$

Using the above properties accounting that the project hiddenly fails with probability $(1-p)(1-\pi)$ upon completion at time τ, we calculate

$$(1-e^{-\Lambda\Delta})(1-p)(1-\pi)\bar{t} = (1-p)(1-\pi)\left(\frac{1-e^{-\Lambda\Delta}}{\Lambda} - e^{-\Lambda\Delta}\Delta\right),$$

and

$$(1-e^{-\Lambda\Delta})(1-p)(1-\pi)(\Delta-\bar{t}) = (1-p)(1-\pi)\left(\Delta - \frac{1-e^{-\Lambda\Delta}}{\Lambda}\right)$$

The alternative strategy is strictly worse than following a full disclosure contract from t onwards if $B^+ \leq C^-$, i.e., if

$$(1-e^{-\Lambda\overline{\Delta}})p\mu \leq (\kappa-\phi)(1-p)(1-\pi)\left[\left(\Delta - \frac{1-e^{-\Lambda\Delta}}{\Lambda}\right) - (1-e^{-\Lambda\Delta})(1-e^{-\lambda\Delta})\frac{\mu}{\kappa-\phi}\right], \tag{2.78}$$

where $\overline{\Delta}$ is defined in (2.77). That is, when (2.78) holds for all $\Delta \geq 0$, a full disclosure (continuation) contract from time t onwards is optimal, so that $\beta_s = w_s = W_s$, $\delta_s = \frac{\phi}{\Lambda p - \lambda}$ for $s \geq t$. Also note that (2.78) holds in equality for $\Delta = 0$.

In the following, we demonstrate that (2.78) holds for all $\Delta \geq 0$ when $\lambda > 0$ is sufficiently small. That is, we show there exists $\overline{\lambda} > 0$, such that a full disclosure contract from time t with $w_t = W_t$ is optimal when $\lambda < \overline{\lambda}$. As $\overline{\Delta}$ is proportional to λ (so that $\lim_{\lambda\to 0}\overline{\Delta} = 0$) and as $\lim_{\lambda\to 0}(1-e^{-\Lambda\Delta})(1-e^{-\lambda\Delta}) = 0$ and as $\Delta - \frac{1-e^{-\Lambda\Delta}}{\Lambda} > 0$ for $\Delta > 0$, it follows that (2.78) holds for $\Delta > 0$ when $\lambda > 0$ is sufficiently small. Thus, for any $\Delta > 0$, we can find $\lambda^\Delta > 0$ such that (2.78) holds provided $\lambda \in [0, \lambda^\Delta)$. It remains to show that we can pick the sequence λ^Δ such that $\lim_{\Delta\to 0}\lambda^\Delta > 0$.

To do so, we characterize the limiting behavior of (2.78) as $\Delta \to 0$. For this sake, we define

$$\mathcal{G}(\Delta) := (\kappa-\phi)(1-p)(1-\pi)\left(\Delta - \frac{1-e^{-\Lambda\Delta}}{\Lambda} - (1-e^{-\Lambda\Delta})(1-e^{-\lambda\Delta})\frac{\mu}{\kappa-\phi}\right)$$
$$-(1-e^{-\Lambda\overline{\Delta}})p\mu,$$

which is the difference between the right hand side and left hand side of (2.78). That is, (2.78) holds for Δ if and only if $\mathcal{G}(\Delta) \geq 0$. One can calculate that

$$\mathcal{G}'(0) = \mathcal{G}(0) = 0 \quad \text{and} \quad \mathcal{G}''(0) = \Lambda(1-p)(1-\pi)\big(\kappa - \phi - 2\lambda\mu\big) - 2\lambda\Lambda^2\mu p\mathcal{Y}.$$

Clearly, $\lim_{\lambda\to 0}\mathcal{G}''(0) > 0$. When $\mathcal{G}''(0) > 0 = \mathcal{G}'(0) = \mathcal{G}(0)$, (2.78) holds for values $\Delta > 0$ that lie in a neighbourhood of zero. Thus, we can pick the sequence λ^Δ such that $\lim_{\Delta\to 0}\lambda^\Delta > 0$, which concludes the proof. To summarize, we have shown there exists $\overline{\lambda} > 0$ such that a full disclosure contract from time t with $w_t = W_t$ is optimal when $\lambda < \overline{\lambda}$.

2.11.8 Full disclosure contract is not optimal if and only if $\pi > 0$ or $\lambda > 0$

Suppose to the contrary that a (optimal) full disclosure contract $\mathcal{C}^0$ with payoff F^0 is optimal. As shown before, this full disclosure contract (optimally) sets $w_t = \gamma_t = W_t = \beta_t$, $T = \inf\{t \geq 0 : \beta_t = 0\}$ and $\alpha_t = \beta_t + \frac{\phi}{\Lambda p - \lambda}$, leading to $\dot{w}_t = -\phi\left(\frac{\Lambda p}{\Lambda p - \lambda}\right)$.

Take some Δ satisfying $T > \Delta > 0$ and define a contract $\tilde{\mathcal{C}}$ that does not incentivize disclosure of failure for $t \in [0, \Delta)$ as follows. $\tilde{\mathcal{C}}$ sets rewards for success $\tilde{\alpha}_t = w_t(1-\pi) + \phi/(\Lambda p)$, and stipulates no rewards for (publicly and privately observed) failure for $t < \Delta$. If the project is not completed by time Δ, it switches at time Δ to the (optimal) full disclosure contract in state $w_\Delta = w_0 - \Delta\phi\left(\frac{\Lambda p}{\Lambda p - \lambda}\right)$. Note that the agent optimally reports failure at time Δ, if the project fails before time Δ and failure is privately observed.

The likelihood of privately observed failure over $[0, \Delta]$ equals $\Lambda(1-p)(1-\pi)\Delta + o(\Delta^2)$, the likelihood of public failure over $[0, \Delta]$ equals $\Lambda(1-p)\pi\Delta + o(\Delta^2)$, and the likelihood that the project succeeds over $[0, \Delta]$ equals $\Lambda p\Delta + o(\Delta^2)$. In addition, for all $t \in [0, \Delta]$

$$\alpha_t = \alpha_s + \dot{\alpha}_s(t-s) + o(\Delta^2) = \alpha_s + o(\Delta) \quad \text{for any } s \in [0, \Delta),$$

as $|s - t| = o(\Delta)$.

It follows that

$$\bar{\alpha}_\Delta := \mathbb{E}[\alpha_\tau | \text{Success over } [0, \Delta]] = \alpha_s + o(\Delta) \text{ for any } s \in [0, \Delta),$$

which is the expected compensation in case of success over $[0, \Delta]$. Likewise, one calculates that the expected financing costs over $[0, \Delta]$ equal

$$\bar{\kappa}_\Delta = \kappa(\Delta - \mathbb{P}(\text{Success or failure over } [0, \Delta])\mathbb{E}[\Delta - \tau | \text{Success or failure over } [0, \Delta]]).$$

Hence, $\bar{\kappa}_\Delta = \kappa\Delta - o(\Delta^2)$. Also note that

$$\bar{\beta}_\Delta := \mathbb{E}[\beta_\tau | \text{Failure over } [0, \Delta]] = \beta_s + o(\Delta) \text{ for any } s \in [0, \Delta).$$

Here, $\bar{\beta}_\Delta$ is the expected compensation for any type of failure under the full disclosure contract.

Take arbitrary $s \in [0, \Delta)$. We can write the payoff under the full disclosure contract as

$$\begin{aligned} F^0 &= (1 - e^{-\Lambda\Delta}) \left(p(\mu - \bar{\alpha}_\Delta) - (1-p)\bar{\beta}_\Delta \right) - \bar{\kappa}_\Delta + e^{-\Lambda\Delta} F(w_\Delta) + o(\Delta^2) \\ &= \Lambda\Delta(p(\mu - \alpha_s) - (1-p)w_s) - \kappa\Delta + (1 - \Lambda\Delta)F(w_\Delta) + o(\Delta^2) \\ &= \Lambda\Delta \left(p\mu - w_\Delta - \frac{\phi p}{\Lambda p - \lambda} \right) - \kappa\Delta + (1 - \Lambda\Delta)F(w_\Delta) + o(\Delta^2), \end{aligned}$$

where it was used in the last inequality that $\alpha_s = \frac{\phi}{\Lambda p - \lambda} + w_\Delta + o(\Delta)$ and $\beta_s = \gamma_s = w_s$ for $s \in [0, \Delta)$ under the full disclosure contract.

Likewise, the contract $\tilde{\mathcal{C}}$ implements the same deadline T as $\mathcal{C}^0$ and, by arguments similar to the ones above, yields payoff at time zero

$$\begin{aligned} F^1 &= \Lambda\Delta(p(\mu - \tilde{\alpha}_s) - (1-p)(1-\pi)w_\Delta) - \kappa\Delta + (1 - \Lambda\Delta)F(w_\Delta) + o(\Delta^2) \\ &= \Lambda\Delta \left(p\mu - w_\Delta(1-\pi) - \frac{\phi}{\Lambda} \right) - \kappa\Delta + (1 - \Lambda\Delta)F(w_\Delta) + o(\Delta^2), \end{aligned}$$

where it was used in the last inequality that $\tilde{\alpha}_s = \phi/(\Lambda p) + w_\Delta(1-\pi) + o(\Delta)$ for $s \in [0, \Delta)$ under the contract $\tilde{\mathcal{C}}$.

Hence:

$$F^1 - F^0 = \Lambda\Delta w_\Delta \pi + \Lambda\Delta \left(\frac{\phi p}{\Lambda p - \lambda} - \frac{\phi}{\Lambda} \right) + o(\Delta^2),$$

which exceeds zero (when $\lambda > 0$ or $\pi > 0$) for Δ sufficiently small, contradicting the optimality of $\mathcal{C}^0$. Note that $w_\Delta > 0$ as $T > \Delta$. Thus, a full disclosure contract is not optimal, if $\pi > 0$ or $\lambda > 0$. On the other hand, if $\pi = \lambda = 0$, then a full disclosure contract is optimal by means of Lemma 5.

2.11.9 Optimal contract when $\mu^f < -\frac{\phi}{\lambda}$

Note that $\mu^f < -\frac{\phi}{\lambda}$ is equivalent to $\lambda\mu < \kappa - \phi$. Suppose that λ is sufficiently small so that (2.78) holds for all $\Delta \geq 0$, and suppose $\pi > 0$ or $\lambda > 0$. Also recall that Sections 2.11.3 through 2.11.8 consider $\mu^f < -\frac{\phi}{\lambda}$, in that a failed project is inefficient to continue financing and the principal terminates financing once failure becomes known to her. Under these circumstances, we have shown in Section 2.11.8 that i) a full disclosure contract is not optimal from time 0 and in Sections 2.11.6 and 2.11.7 that ii) a full disclosure contract becomes optimal from time t onwards when $W_t = w_t$. In addition, W_t and w_t are at the latest equal when $t = T$.

As such, the optimal contract features an unconditional financing stage $[0, t_1)$, during which the agent is not incentivized to disclose failure. As the principal minimizes agency costs during the unconditional financing stage, there are no rewards for publicly observed failure and there are minimal rewards for success. That is, $\alpha_t = w_t(1-\pi) + \frac{\phi}{\Lambda p}$, $\beta_t = \gamma_t = 0$, and $W_t < w_t$.

The unconditional financing stage is followed by a disclosure stage $[t_1, T]$ during which the agent is not incentivized to disclose failure. During the disclosure stage, it holds that $w_t = \beta_t = \gamma_t$ and $\alpha_t = w_t + \frac{\phi}{\Lambda p - \lambda}$, leading to $W_t = w_t$ (see Sections 2.11.4 and 2.11.5).

2.11.10 Optimal contract when $\mu^f \geq -\frac{\phi}{\lambda}$

Note that $\mu^f \geq -\frac{\phi}{\lambda}$ is equivalent to $\lambda\mu \geq \kappa - \phi$. Recall that we assume the principal optimally terminates financing once failure becomes known to her, and full effort is optimal. When $\mu^f \geq -\frac{\phi}{\lambda}$, the optimal contract takes a relatively simple form. The principal grants unconditional financing until time $T \leq \infty$, i.e., $t_1 = T$. During the unconditional financing stage, the agent is not paid for failure (i.e., $\beta_t = \gamma_t = 0$) and receives minimal rewards for success (i.e., $\alpha_t = (1-\pi)w_t + \frac{\phi}{\Lambda p}$). The disclosure stage becomes degenerate in that the principal incentivizes disclosure of failure only at time T where the agent's contract is terminated anyways and the principal terminates financing the project.

To prove that the optimal contract does not incentivize disclosure of failure, it suffices to show that the principal does not improve her payoff by incentivizing disclosure of failure. Suppose that the project has failed at time t and the agent has privately observed failure, leading to payoff w_t for the agent, and the principal finances the project until time T'. Note that $T' < T$ is possible if and only if the principal incentivizes disclosure of failure at some time $s \in (t, T)$. As the principal terminates financing once failure becomes known to her, $T' \leq T$. Given T', total surplus (after privately observed failure at time t) becomes

$$\hat{\mathcal{S}}_t = \int_t^{T'} e^{-\lambda(s-t)}(\lambda\mu - \kappa + \phi)ds,$$

which is split between principal and agent. Thus, the principal's payoff (after privately observed failure at time t) is $\hat{F}_t = \hat{\mathcal{S}}_t - w_t$. As $\lambda\mu \geq \kappa - \phi$, the principal maximizes total surplus $\hat{\mathcal{S}}_t$ (after privately observed failure at time t) by financing the project until time T, so that in optimum $T' = T$. As such, by not incentivizing disclosure at all times $s \in [t, T]$ and therefore by financing the project until time $T' = T$ (after failure is privately observed), the principal maximizes her payoff $\hat{F}_t = \hat{\mathcal{S}}_t - w_t$, given the agent's payoff after privately observed failure w_t. This shows that incentivizing disclosure of failure does not improve the principal's payoff (and in fact is strictly suboptimal when $\mu^f > -\frac{\phi}{\lambda}$). To minimize agency costs, the principal then sets no rewards for failure and minimal rewards for success, leading to $\beta_t = \gamma_t = 0$ and $\alpha_t = (1-\pi)w_t + \phi/(\Lambda p)$ for all $t \in [0, T]$.

2.11.11 Proof of Corollary 2

Suppose that $\mu^f < -\frac{\phi}{\lambda}$, i.e., $\lambda\mu < \kappa - \phi$ and λ is sufficiently small. We start by considering a full disclosure contract, so the principal terminates financing immediately upon failure. Then, $w = \beta$ is the only state variable. By the dynamic programming principle, the value function $F(w)$ solves

$$\Lambda F(w) = \Lambda p\mu - \kappa - \Lambda((1-p)w + p\alpha) - F'(w)(\phi + \lambda(\alpha - w)), \tag{2.79}$$

where in optimum (see Section 2.11.4):

$$\alpha = w + \frac{\phi}{\Lambda p - \lambda} \quad \text{and} \quad \gamma = w$$

so that $\frac{\partial \alpha}{\partial \lambda} > 0$. On $[0, w_0]$, the value function is increasing, in that $F'(w) \geq 0$; otherwise, the principal could improve her initial payoff $F(w_0)$ by decreasing w_0. Thus, due to $\frac{\partial \alpha}{\partial \lambda} > 0$ and $F'(w) \geq 0$, it follows that

$$\frac{\partial F(w)}{\partial \lambda} < 0$$

for any $w \in [0, w_0]$. As a result, the principal's payoff at time zero $F_0 = F(w_0)$ decreases with λ. In particular, since $\frac{\partial F(w)}{\partial \mu} > 0$ for any $w \in [0, w_0]$, the principal's payoff at time zero $F_0 = F(w_0)$ decreases when a mean preserving spread increases λ but decreases μ. Note that a full disclosure contract is optimal when $\pi = \lambda = 0$ (see Section 2.11.8 or Lemma 5). By continuity, when $\pi > 0$ and $\lambda > 0$ are sufficiently small, the principal's payoff at time zero F_0 decreases when a mean preserving spread increases λ but decreases μ.

2.12 Additional results

2.12.1 Solution when agent affects completion timing

We present an alternative formulation of the moral hazard problem similar to Mason and Välimäki (2015), Green and Taylor (2016), or Varas (2017). In this alternative model, the agent controls project completion, while the project is subject to failure risk during its development phase. The agent affects the project completion timing $\tau = \inf\{t \geq 0 : dN_t = 1\}$ with his effort $a_t \in \{0, 1\}$, whereas project failure at time $\tau^\delta := \inf\{t \geq 0 : dN_t^\delta = 1\}$ occurs at for simplicity exogenous rate $\delta > 0$. Specifically, assume that N and N^δ are jump processes with $\mathbb{E}dN_t = \Lambda a_t dt \cdot \mathbf{1}_{\{t<\tau^\delta\}}$ and $\mathbb{E}dN_t^\delta = \delta dt$. The agent derives private benefits $\phi(1 - a_t)$ as long as the project receives sufficient financing, i.e., before time $T_0 = T \wedge \tau^A$. Completion at time τ always results into success and yields terminal payoff μ to the principal. Failure (during project development) occurs at time τ^δ, in which case the project becomes worthless and produces zero payoffs for all times $t \geq \tau^\delta$. In particular, due to $\mathbb{E}dN_t = \Lambda a_t dt \cdot \mathbf{1}_{\{t<\tau^\delta\}}$, the project cannot be (successfully) completed anymore after failure has occurred during the project development phase. Also note that by exerting effort to complete the project, the agent accelerates completion and hence reduces the risk of failure during project development, in that the agent effectively controls project failure (risk) too. Like in the baseline model, failure is publicly observed with probability $\pi \in [0, 1]$. Otherwise, failure is privately observed by the agent and reported failure is not verifiable.

Incentive Compatibility. The agent is not paid for observed failure but obtains payoff w_t upon privately observed failure. We take the agent's continuation value for $t < T \wedge \tau \wedge \tau^\delta$:

$$W_t = \int_t^T e^{-(\Lambda+\delta)(s-t)}(\Lambda \alpha_s + \delta(1 - \pi)w_s)ds,$$

so that

$$\dot{W}_t = (\Lambda + \delta)W_t - \delta(1 - \pi)w_t - \Lambda \alpha_t.$$

Like in Green and Taylor (2016), the incentive condition w.r.t. effort a_t becomes

$$\alpha_t - W_t \geq \frac{\phi}{\Lambda}. \tag{2.80}$$

The intuition is as follows. If the agent exerts effort a_t, the project succeeds with probability Λdt in which case the agent receives a reward for success α_t but looses his continuation payoff W_t. On the other hand, shirking over a time interval of length dt yields private benefits ϕdt but then the project does not complete for sure.

Like in our baseline model, the incentive condition w.r.t. disclosure of failure is

$$\dot{\beta}_t \leq -\phi,$$

which in optimum leads to $\dot{w}_t = -\phi$. In addition, the agent must not find it optimal to fake bad outcomes, which requires

$$W_t \geq \beta_t.$$

Full disclosure contract. In a (optimal) full disclosure contract, it holds that $W_t = w_t = \beta_t$ and $\dot{w}_t = \dot{\beta}_t = -\phi$ for all t. Due to $W_T = w_T = 0$, this requires

$$0 = \dot{W}_t - \dot{w}_t = \Lambda(W_t - \alpha_t) + \delta\pi w_t + \phi.$$

Hence:

$$\alpha_t = W_t + \frac{\phi + \delta\pi w_t}{\Lambda},$$

so that the incentive condition w.r.t. effort (2.80) is slack.

Optimal contract. In order to reduce rewards for success α_t and hence the agent's stake W_t and agency costs, the optimal contract features an unconditional financing stage $[0, t_1)$, during which $\alpha_t = W_t + \phi/\Lambda$ and $\beta_t = 0$. During that stage, the optimal contract does not incentivize disclosure of failure. Thereafter, the optimal contract is a full disclosure contract with some deadline T, stipulating $\beta_t = w_t$ and $\alpha_t - W_t = \frac{\phi+\delta\pi w_t}{\Lambda}$. The further solution steps look like those presented in Section 4.3 and are therefore omitted.

2.12.2 Solution when success is unobservable

Suppose the principal observes success — just like failure — only with probability π. Otherwise, success is privately observed by the agent. Crucially, we assume that the agent can hide success from the principal (i.e., delay disclosure of success), but the agent cannot fake success (i.e., misreport success before it occurs, which reflects that it is arguably more difficult to fake good outcomes rather than bad outcomes.

We start with some notation. Denote the pay for (publicly) observed success by ω_t and denote the pay for reported success by α_t. Then, the agent discloses success truthfully if and only if $\dot{\alpha}_t \leq -\phi$. Likewise, the agent discloses failure truthfully if and only if $\dot{\beta}_t \leq -\phi$.

Full disclosure contract. Let us start by looking at the disclosure stage or — equivalently — at a full disclosure contract (recall that the disclosure stage is a full disclosure continuation contract). We illustrate that it is not consequential whether success is observable or not during the disclosure stage (i.e., during a full disclosure contract). To start with, note that the full disclosure contract from Proposition 2 stipulates

$$\alpha_t = w_t\left(1 - \pi + \frac{\pi}{p}\right) + \frac{\phi}{\Lambda p} \implies \dot{\alpha}_t < \dot{w}_t = -\phi.$$

Hence, within the optimal full disclosure contract from Proposition 2, the agent always strictly prefers to disclose success truthfully at the time it occurs. That is, for this contract the observability of success does not matter (one can set $\alpha_t = \omega_t$).

In the following, we take the optimal full disclosure contract (see Proposition 5) that sets $\gamma_t = w_t$ and $\alpha_t = w_t + \frac{\phi}{\Lambda p}$, thereby "minimizing" rewards for success. In this contract, $\beta_t = w_t$ and

$$\dot{\beta}_t = \dot{\alpha}_t = -\phi,$$

so that the agent possesses sufficient incentives to disclose failure and success truthfully. Rewards for observed and reported success (failure) are equal, in that $\omega_t = \alpha_t$. In this full disclosure contract, the agent's stake, capturing agency costs, is given by $W_t = w_t$.

Optimal contract. We construct the optimal contract using heuristic arguments. A formal proof conjectures the shape of the contract — which is discussed here — and then provides a (formal) verification argument — which is omitted here. The optimal contract features an unconditional financing stage $[0, t_1)$ during which it does not incentivize disclosure of failure and success. Thereafter, after t_1, the contract becomes a full disclosure contract with time to deadline $T - t_1$, stipulating $\beta_t = \gamma_t = w_t$ and $\alpha_t = \omega_t = w_t + \phi/(\Lambda p)$. It therefore holds — by construction — that $W_{t_1} = w_{t_1}$.

If the project fails (succeeds) at time $t < t_1$ and failure (success) is privately observed by the agent , the agent reports failure (success) at time t_1. Hence, privately observed project completion at time t yields payoff $\beta_{t_1} + \phi(t_1 - t)$, in case of failure, and payoff $\alpha_{t_1} + \phi(t_1 - t) = \beta_{t_1} + \phi(t_1 - t) + \phi/(\Lambda p)$ in case of success. In addition, the agent is not paid for observed failure, i.e., $\gamma_t = 0$ for $t < t_1$, and receives pay $\omega_t = \phi/(\Lambda p)$ for observed success, when $t < t_1$.

Let us look at the agent's incentives to exert effort over $[t, t + dt)$ with $t < t_1$. Shirking entails benefits ϕdt and, if the project completes, the project fails. Failure is observable with probability π, in which case the agent receives zero payoff, and otherwise with probability $1 - \pi$ the agent's payoff is $\beta_{t_1} + \phi(t_1 - t)$. If the agent works and the project succeeds, the agent receives pay $\phi/(\Lambda p)$, when success is observed (with probability π), and otherwise with probability $1 - \pi$ pay $\alpha_{t_1} + \phi(t_1 - t) = \beta_{t_1} + \phi(t_1 - t) + \phi/(\Lambda p)$. Therefore, the agent prefers to work (i.e., to exert full effort $a_t = 1$) if and only if

$$\begin{aligned}&\Lambda dt \cdot \Big(p[(1-\pi)(\alpha_{t_1} + \phi(t_1 - t)) + \pi\phi/(\Lambda p)] + (1-p)(1-\pi)[\beta_{t_1} + \phi(t_1 - t)]\Big)\\ &\geq \phi dt + \Lambda dt \cdot (1-\pi)[\beta_{t_1} + \phi(t_1 - t)].\end{aligned}$$

The first line depicts the agent's (expected) pay upon exerting effort and the second line depicts the agent's pay upon shirking. Using simple algebra, it can be verified that the above condition is — by construction of the proposed contract — satisfied. More straightforwardly, the proposed contract motivates effort because the agent's pay for success in each state of the world exceeds his pay for failure by $\phi/(\Lambda p)$, which outweighs the disutility of effort.

Note that the agent's continuation utility during the unconditional financing stage is given by

$$W_t = \int_t^T e^{-\Lambda(s-t)} \Big[\Lambda\Big((1-p)(1-\pi)w_s + p[\pi\omega_s + (1-\pi)(w_s + \phi/(\Lambda p))]\Big)\Big] ds,$$

so that

$$\begin{aligned}\dot{W}_t &= \Lambda W_t - \Lambda\Big((1-p)(1-\pi)w_t + p[\pi\omega_t + (1-\pi)(w_t + \phi/(\Lambda p))]\Big)\\ &= \Lambda(W_t - w_t) - \phi + \pi w_t.\end{aligned}$$

The second equality follows after plugging in the previously derived expressions for ω_t and α_t. Hence, due to $\dot{w}_t = -\phi$:

$$\dot{W}_t - \dot{w}_t = \Lambda(W_t - w_t) + \pi w_t,$$

and because of $W_{t_1} = w_{t_1}$ it follows that $W_t < w_t$ and $0 > \dot{W}_t > \dot{w}_t = -\phi$ for $t < t_1$. Hence, the provision unconditional financing over some period $[0, t_1)$ reduces the agent's stake W_t and hence agency costs relative to a full disclosure contract. Thus, the provision of unconditional financing over $[0, t_1)$ is optimal and the optimal contract (likely) takes the conjectured shape. A rigorous proof can be constructed along the lines of the proof of Proposition 6.

2.12.3 A model variant with convex costs of effort

Setup

We consider a model variant, in which the agent incurs convex and private non-pecuniary flow costs of exerting effort $\frac{\theta a_t^2}{2}$, where $\theta > 0$. The likelihood of success is pa_t with $a_t \in [0,1]$ and $p \in (0,1)$. Throughout, we assume that θ is sufficiently large, so that optimal effort a_t is interior and $a_t \in (0,1)$.

In addition, the agent derives private benefits ϕ from operating the project which pertain as long as the principal finances the project, that is, until time $T \wedge \tau^A$. The private benefits of operating the project become necessary in the formulation with private costs of effort (rather than private benefits of shirking) to obtain the tension that the agent would like to continue the project after its failure to derive flow benefits ϕ while the principal incurring flow costs $\kappa > \phi$ would like to terminate the project. All other model elements remain unchanged relative to the baseline.

In this model variant, the agent's continuation payoff (for times $t < \tau$) is

$$W_t = \int_t^T e^{-\Lambda(s-t)} \left(\Lambda\big[pa_s\alpha_s + (1-pa_s)((1-\pi)w_s + \pi\gamma_s)\big] - \frac{\theta a_s^2}{2} + \phi\right) ds, \quad (2.81)$$

so that

$$\dot{W}_t = \Lambda W_t + \frac{\theta a_t^2}{2} - \phi - \Lambda\big[pa_t\alpha_t + (1-pa_t)((1-\pi)w_t + \pi\gamma_t)\big]. \quad (2.82)$$

Incentives

As in the baseline, the agent finds it optimal not to delay disclosure of failure if and only if

$$\dot{\beta}_t \le -\phi.$$

In optimum, the incentive constraint is tight whenever the contract incentivizes disclosure of failure, implying that

$$\dot{w}_t = -\phi.$$

As a result, one obtains that at time t, the time to deadline is

$$T - t = \frac{w_t}{\phi}.$$

In addition, $W_t \geq \beta_t$ must hold to preclude that the agent fakes failure, i.e., misreports failure before it occurs.

The agent chooses effort $a_t \in [0,1]$ to maximize

$$\Lambda\Big[pa_t\alpha_t + (1-pa_t)((1-\pi)w_t + \pi\gamma_t)\Big] - \frac{\theta a_t^2}{2}.$$

Provided $a_t \in (0,1)$, we can take the first-order condition with respect to a_t and obtain

$$a_t = \frac{\Lambda p}{\theta}\left(\alpha_t - (1-\pi)w_t - \pi\gamma_t\right). \tag{2.83}$$

Inserting (2.83) into (2.82) and simplifying yields

$$\begin{aligned} \dot{W}_t &= \Lambda W_t + \frac{\theta a_t^2}{2} - \phi - \Lambda\big[(1-\pi)w_t + \pi\gamma_t\big] - \theta a_t^2 \\ &= \Lambda W_t - \frac{\theta a_t^2}{2} - \phi - \Lambda\big[(1-\pi)w_t + \pi\gamma_t\big]. \end{aligned} \tag{2.84}$$

Integrating (2.84) over time yields

$$W_t = \int_t^T e^{-\Lambda(s-t)}\left(\Lambda\big[(1-\pi)w_s + \pi\gamma_s\big] + \phi + \frac{\theta a_s^2}{2}\right) ds. \tag{2.85}$$

First-best and second-Best benchmarks

Note that first-best effort (i.e., the effort level absent any frictions and observable and contractible effort) is determined by maximizing

$$\Lambda p a_t \mu - \frac{\theta a_t^2}{2},$$

so that first-best effort — provided it is interior — is given by

$$a^{FB} = \frac{\Lambda p \mu}{\theta}.$$

We assume that $\Lambda p\mu < \theta$, so that indeed $a^{FB} < 1$.

We also consider the second-best benchmark in which effort is unobservable but failure and success are observable and contractible, which is analogous to the problem in Section 2.3.1. As in the baseline with publicly observed failure (i.e., as in Section 2.3.1), the agent is paid for success but not for failure, so that $\beta_t = \gamma_t = 0$ and (2.83) becomes

$$a_t = \frac{\Lambda p \alpha_t}{\theta}. \tag{2.86}$$

The agent is not punished for completion delays, there is no financing deadline, and the principal finances the project until completion at time τ, leading to $T = \infty$. At any time $t < \tau$, the principal maximizes

$$\Lambda p a_t(\mu - \alpha_t) = \Lambda p a_t \mu - \theta a_t^2,$$

so that second-best effort is given by

$$a^{SB} = \frac{\Lambda p \mu}{2\theta} = \frac{a^{FB}}{2}.$$

Thus, ideally, the principal would like to implement effort a^{SB} in the presence of agency conflicts over effort. However, as the following exposition below demonstrates, incentivizing disclosure of failure might make it necessary to stipulate higher effort beyond a^{SB}.

Full disclosure contract

We start by analyzing a full disclosure contract running from time t until the financing deadline $T = \inf\{s \geq t : w_s = 0\}$ with information rents w_t and continuation payoff W_t. The agent's rewards for reported failure are $\beta_t = w_t$ and it must hold that $W_t \geq w_t$ to preclude the agent fakes failure. The agent discloses failure at the time it occurs in that $\tau^A = \tau$, and the principal finances the project until time $T \wedge \tau$. As $\dot{w}_s = -\phi$ for $s \geq t$, it holds that $w_t = \phi(T - t)$. In other words, as w_t is a state variable for the dynamic optimization, we study a full disclosure contract in state w_t which — due to $\dot{w}_s = -\phi$ — implies time to deadline $T - t = \frac{w_t}{\phi}$. In addition to w_t, the problem also has the state variable W_t.

The total surplus — divided by principal and agent — from time t onwards is then given by

$$\mathcal{S}_t = \int_t^T e^{-\Lambda(s-t)} \left(\Lambda p a_s \mu - \frac{\theta a_s^2}{2} \right) ds - \int_t^T e^{-\Lambda(s-t)} (\kappa - \phi) ds. \tag{2.87}$$

As such, the principal's payoff from time t is then $F_t = \mathcal{S}_t - W_t$ where W_t is the agent's payoff from time t. The principal chooses the agent's contract terms to maximize her payoff F_t subject to the requirement of delivering continuation payoff W_t to the agent (i.e., constraint (2.85)) and that $W_s \geq w_s$ at all times $s \in [t, T]$. Since the second integral in (2.87) does not depend on a_s, the principal maximizes the objective

$$\int_t^T e^{-\Lambda(s-t)} \left(\Lambda p a_s \mu - \frac{\theta a_s^2}{2} \right) ds \tag{2.88}$$

subject to the requirement of delivering continuation payoff W_t to the agent (i.e., constraint (2.85)) and that $W_s \geq w_s$ at all times $s \in [t, T]$, whereby $W_T = w_T = 0$. In what follows, we solve for the optimal full disclosure contract from time t, given values W_t and w_t, in several steps.

Step 1 Consider the relaxed problem in which the principal maximizes the objective

$$\int_t^T e^{-\Lambda(s-t)} \left(\Lambda p a_s \mu - \frac{\theta a_s^2}{2} \right) ds$$

over $a_s \in [0, 1]$ and $\gamma_s \geq 0$ subject to (2.85). The Lagrangian for this problem reads

$$\mathcal{L} = \int_t^T e^{-\Lambda(s-t)} \left(\Lambda p a_s \mu - \frac{\theta a_s^2}{2} \right) ds \tag{2.89}$$

$$+ \ell \left(W_t - \int_t^T e^{-\Lambda(s-t)} \left(\Lambda \big[(1-\pi) w_s + \pi \gamma_s \big] + \phi + \frac{\theta a_s^2}{2} \right) ds \right), \tag{2.90}$$

with Lagrange multiplier ℓ. We can maximize the Lagrangian pointwise with respect to $a_s \in [0,1]$ and $\gamma_s \geq 0$. Provided $a_s \in [0,1]$ is interior, the first order condition

$$\frac{\partial \mathcal{L}}{\partial a_s} = e^{-\Lambda(s-t)}(\Lambda p\mu - (1+\ell)\theta a_s)ds = 0 \quad \Longleftrightarrow \quad a_s = \frac{\Lambda p\mu}{\theta(1+\ell)} =: a^\ell \tag{2.91}$$

holds, so that $a^\ell < a^{FB}$ if and only if $\ell > 0$. Note that $a_s = a^\ell$ is constant over time. In addition:

$$\ell = \frac{\Lambda p\mu}{\theta a^\ell} - 1. \tag{2.92}$$

The derivative with respect to γ_s implies that

$$\frac{\partial \mathcal{L}}{\partial \gamma_s} = -\ell e^{-\Lambda(s-t)}\Lambda\pi\gamma_s ds$$

so that $\gamma_s = 0$ whenever $\ell > 0$.

Next, we define

$$W_t^x = \int_t^T e^{-\Lambda(s-t)}\left(\Lambda(1-\pi)w_s + \phi + \frac{\theta(a^x)^2}{2}\right)ds \tag{2.93}$$

$$= \int_t^T e^{-\Lambda(s-t)}\left(\Lambda(1-\pi)\phi(T-s) + \phi + \frac{\theta(a^x)^2}{2}\right)ds, \tag{2.94}$$

for $x \in \{\ell, FB, SB\}$, using $T - s = \frac{w_s}{\phi} \iff w_s = \phi(T-s)$. As $a^\ell < a^{FB}$ if and only if $\ell > 0$, it follows that $W_t^\ell < W^{FB}$ if and only if $\ell > 0$. The above integral expression admits the following closed-form solutions:

$$W_t^x = \frac{1}{\Lambda}\left((1-e^{-\Lambda(T-t)})\left(\frac{\theta(a^x)^2}{2} + \pi\phi\right) + \Lambda(1-\pi)\phi(T-t)\right) \tag{2.95}$$

$$= \frac{1}{\Lambda}\left((1-e^{-\Lambda w_t/\phi})\left(\frac{\theta(a^x)^2}{2} + \pi\phi\right) + \Lambda(1-\pi)w_t\right), \tag{2.96}$$

for $x \in \{\ell, FB, SB\}$.

Step 2 As a next step, we conjecture that $\ell > 0$, so that $\gamma_s = 0$ for all $s \in [t, T]$. We can now rearrange (2.95) with $W_t^\ell = W_t$ to obtain

$$(a^\ell)^2 = \frac{2}{\theta}\left(\frac{1}{(1-e^{-\Lambda w_t/\phi})}(\Lambda(W_t - (1-\pi)w_t)) - \pi\phi\right) = \mathcal{R}(W_t, w_t)$$

$$\Longleftrightarrow \quad a^\ell = \sqrt{\mathcal{R}(W_t, w_t)}.$$

Using (2.92), the candidate solution for the multiplier becomes

$$\ell^* = \frac{\Lambda p\mu}{\theta\sqrt{\mathcal{R}(W_t, w_t)}} - 1.$$

With the candidate solution for the multiplier in hand, the following cases can prevail:

1. $\ell^* > 0$: In this case, $\ell^* = \ell$ and ℓ is indeed strictly positive. Then, it holds for all $s \in [t, T]$ that
$$a_s = a^\ell = \frac{\Lambda p\mu}{\theta(1+\ell^*)} \quad \text{and} \quad \gamma_s = 0.$$

2. $\ell^* \le 0$: In this case, $\ell = 0$, $W_t \ge W_t^{FB}$ and $\gamma_s \ge 0$ can be chosen constant over time. Thus, it holds for all $s \in [t, T]$ that $a_s = a^{FB}$. To derive the value of γ_s, we solve
$$W_t = \int_t^T e^{-\Lambda(s-t)} \left(\Lambda\big[(1-\pi)w_s + \pi\gamma_s\big] + \phi + \frac{\theta a_s^2}{2} \right) ds = W_t^{FB} + \int_t^T e^{-\Lambda(s-t)} \Lambda\pi\gamma_s ds$$
for γ_s, with $a_s = a^{FB}$ and γ_s constant over time. This yields
$$\gamma_s = \frac{W_t - W_t^{FB}}{\pi\big(1 - e^{-\Lambda w_t/\phi}\big)}.$$

Thus, in case (1) the principal would like to maximize incentives for effort, which makes it optimal to stipulate $\gamma = 0$.

Step 3 Finally, we must verify that in the solution to the relaxed problem, the constraint $W_s \ge w_s$ holds for all $s \in [t, T]$. To do so, take the law of motion of W_t — that is, (2.84) — under the optimally constant effort $a_t \equiv a$ and constant pay for publicly observed failure $\gamma_s\gamma$ for all $s \in [t, T]$ and calculate for $\Delta_t^W = W_t - w_t$:
$$\dot\Delta_s^W = \dot W_s - \dot w_s = \Lambda W_s - \frac{\theta a^2}{2} - \Lambda\big[(1-\pi)w_s + \pi\gamma\big] = \Lambda\Delta_s^W - \frac{\theta a^2}{2} + \Lambda\pi(w_s - \gamma), \tag{2.97}$$
Also note that $W_t \ge w_t \iff \Delta_t^W \ge 0$ and $W_T = w_T \iff \Delta_T^W = 0$. In addition, we can derive the second derivate with respect to time (i.e., s) by differentiating (2.97) with respect to time
$$\ddot\Delta_s^W = \Lambda\dot\Delta_s^W - \Lambda\pi\phi, \tag{2.98}$$
using that $\dot w_s = -\phi$.

At time $s = T$, $\Delta_T^W = w_T = 0$, so that $\dot\Delta_s^W < 0$ for s in a neighbourhood of T and therefore $\Delta_s^W > 0$ for s in a neighbourhood of T. Suppose to the contrary that there exists $s \in (t, T)$ with $\Delta_s^W < 0$. Due to $\dot\Delta_s^W < 0$ and $\Delta_s^W > 0$ for s in a neighbourhood of T and due to $\Delta_t^W \ge 0 = \Delta_T^W$, it must be that there exist $s' \in (t, T)$ such that $\Delta_{s'}^W$ is the local minimum of Δ_s^W on $[t, T]$, so that $\dot\Delta_{s'}^W = 0 < \ddot\Delta_{s'}^W$. However, inserting $\dot\Delta_s^W = 0$ into (2.98) implies that $\ddot\Delta_s^W < 0$, a contradiction. It follows that $\Delta_s^W \ge 0$ at all times $s \in [t, T]$, which was to show.

Step 4 We now determine the optimal starting values of $W_t \ge w_t$ at time $t = 0$. By the dynamic programming principle, we can express the principal's value function as function of (W, w) and $F(W, w)$ solves the HJB equation
$$\Lambda F(W,w) = \max_{\{a\in[0,1],\gamma\ge 0\}} \Big\{ \Lambda pa(\mu - \alpha) - \kappa - \Lambda(1-pa)((1-\pi)w + \pi\gamma) + F_W(W,w)\dot W + F_w(W,w)\dot w \Big\}.$$

Inserting (2.83), $\dot{w} = -\phi$, and (2.84), we obtain

$$\Lambda F(W,w) = \max_{\{a\in[0,1],\gamma\geq 0\}} \Bigg\{ \Lambda p a \mu - \kappa - \Lambda((1-\pi)w + \pi\gamma) - \theta a^2 - F_w(W,w)\phi \\ + F_W(W,w)\left(\Lambda W - \frac{\theta a^2}{2} - \phi - \Lambda\big[(1-\pi)w + \pi\gamma\big]\right)\Bigg\}.$$

The first-order condition with respect to a implies that

$$a(W,w) = \frac{\Lambda p \mu}{\theta(2 + F_W(W,w))}.$$

As we have shown that $a_s = a(W_s, w_s)$ is constant in optimum for all times $s \in [t,T]$, it follows that $F_W(W,w)$ is constant in optimum for all times $s \in [t,T]$ too, in a sense that

$$F_W(W_{s'}, w_{s'}) = F_W(W_{s''}, w_{s''}) \ \ \forall\ s', s'' \in [t,T].$$

In addition, $1+\ell = 2 + F_W(W,w)$ so that $F_W(W,w) = \ell - 1$. At time $t=0$, it must be that $F_W(W_0, w_0) \leq 0$ because otherwise the principal could increase W_0 to $W_0 + \varepsilon$, thereby improving her initial payoff $F_0 = F(W_0, w_0)$ as well as the agent's incentives not to fake failure (as captured by W). Thus, $F_W(W_s, w_s) = F_W(W_0, w_0) \leq 0$, which implies that $a(W_s, w_s) \geq a^{SB}$. The first order condition with respect to γ implies that $\gamma = 0$ if $F_W(W,w) > -1$ and $F_W(W,w) = -1$ if $\gamma > 0$ in which case $a(W,w) = a^{FB}$. It also holds that $F_W(W,w) \geq -1$. Because $F_W(W_0, w_0) \leq 0$, the initial value of W_0 is then set to

$$W_0 = \min\{w_0, W^{SB}(w_0)\},$$

where $W^{SB}(w_0)$ is defined in (2.93) for $t=0$ and $T = w_0/\phi$. Also note that $F_W(W_0, w_0) = 0$ if and only if $W_0 = W^{SB}$. That is, $a > a^{SB}$ if and only if $F_W(W_0, w_0) < 0$.

Unconditional Financing

The interpretation of $a > a^{SB}$ — i.e., when $W_0 = w_0 > W_0^{SB}$ — is that incentivizing disclosure of failure requires to stipulate high effort (compared to the benchmark with observable failure), which is costly. In other words, a full disclosure contract requires excessively high effort relative to the benchmark with observable failure. As a result, there is a potential improvement by stipulating unconditional financing and not incentivizing disclosure of failure over an initial period $[0,\Delta)$, which allows the principal to implement lower effort and increases the principal's payoff by an amount of order Δ. As the likelihood of failure over a short interval with length Δ is of order Δ and the financing costs over $[0,\Delta)$ are of order Δ too, it follows that any inefficiency of not being able to terminate financing upon failure is of order Δ^2 and therefore negligible for sufficiently small Δ. In other words, the principal may be able to improve upon a full disclosure contract by not incentivizing disclosure over an interval $[0,\Delta)$ with sufficiently small $\Delta > 0$, because the benefits from doing — which are of order Δ — outweigh the costs — which are of order Δ^2 — when Δ is small. That is, a full disclosure contract may not be optimal. As such, the optimal contract may feature an unconditional financing stage $[0, t_1)$ followed by a disclosure stage $[t_1, T]$. The solution during the disclosure stage was characterized above.

During the unconditional financing stage, the principal's value function $f(w,q)$ depends on w and the belief of whether the project has failed so far q. The solution during the unconditional financing stage is characterized by the HJB equation:

$$0 = \max_a \Big\{ -(1-q)\Lambda\big(pa + (1-pa)\pi\big) f(w,q) \tag{2.99}$$

$$+(1-q)\Lambda pa(\mu - \alpha) - \kappa + f_w(w,q)\dot{w} + f_q(w,q)\dot{q} \Big\}, \tag{2.100}$$

whereby $\gamma = 0$, and

$$\dot{q} = (1-q)(1-pa)(1-\pi)\Lambda. \tag{2.101}$$

The incentive condition for effort is

$$a = \frac{\Lambda p}{\theta}(\alpha - (1-\pi)w) \iff \alpha = \frac{\theta a}{\Lambda p} + (1-\pi)w.$$

Optimal effort is then determined via the first-order condition in the HJB equation (2.99), which is

$$(1-q)(\Lambda p\mu - 2\theta a) - f_q(w,q)\Lambda(1-q)(1-\pi)p - (1-q)\Lambda p(1-\pi))(f(w,q)+w) = 0.$$

We can solve

$$a = a(w,q) = \frac{\Lambda p\mu - \Lambda(1-\pi)p(f(w,q) + w + f_q(w,q))}{2\theta}.$$

Finally, note that unlike in the baseline, we have to solve a partial differential equation during the unconditional financing stage, as the choice of effort determines the law of motion of q.

2.12.4 Calculating the average financing horizon

Take $\mathcal{E} = \mathbb{E}[T \wedge \tau^A]$ the average length of the financing period. Define

$$\mathcal{E}_t := \mathbb{E}_t[T \wedge \tau^A | t < T \wedge \tau^A]$$

and note that $\mathcal{E} = \mathcal{E}_0$ and $\mathcal{E}_T = T$, by definition.

On $[t_1, T]$ for $t < T \wedge \tau^A$, we can write

$$\mathcal{E}_t = \mathbb{E}_t[T \wedge \tau^A] = \mathbb{E}_t[T \wedge \tau] = \mathbb{E}_t\left[\int_t^{T\wedge\tau} sds + \mathbf{1}_{\{T \geq \tau\}} T\right] = \int_t^T e^{-\Lambda s}\Lambda sds + e^{-\Lambda(T-t)}T,$$

where the second equality uses truthful disclosure of failure, $\tau = \tau^A$ and the third equality uses integration by parts. Thus, differentiating the above expression w.r.t. t implies that $\mathcal{E}_t$ solves on $[t_1, T]$ the time ODE

$$\Lambda\mathcal{E}_t - \dot{\mathcal{E}}_t = \Lambda t, \tag{2.102}$$

subject to $\mathcal{E}_T = T$. We obtain the closed-form solution

$$\mathcal{E}_t = t + \frac{1 - e^{-\Lambda(T-t)}}{\Lambda}.$$

Taking this solution, we proceed and solve backwards in time. Recall that

$$q_t = 1 - e^{-\Lambda(1-p)(1-\pi)t}.$$

At $t = t_1$, the contract elicits a progress report. If the project has failed already, which is with probability $\lim_{t\uparrow t_1} q_t$, the financing deadline is $T \wedge \tau^A = \tau^A = t_1$, since the agent reports failure at t_1. Otherwise, with probability $1 - \lim_{t\uparrow t_1} q_t$, the project is not complete at time t_1, in which case the expected financing date is given by $\mathcal{E}_{t_1}$. This leads to the value matching condition

$$\lim_{t\uparrow t_1} \mathcal{E}_t = \mathcal{E}_{t_1}(1 - \lim_{t\uparrow t_1} q_t) + t_1 \lim_{t\uparrow t_1} q_t. \tag{2.103}$$

On $[0, t_1)$, the financing is only terminated if the project succeeds, which happens at rate $\Lambda(1 - q_t)p$. Thus, for $t < T \wedge \tau^A$ we can write

$$\begin{aligned}\mathcal{E}_t = \mathbb{E}_t\left[\int_t^{T\wedge\tau^A} sds + \mathbf{1}_{\{T\geq\tau\}}T\right] &= \int_t^{t_1} e^{-\Lambda(p+(1-p)\pi)\int_t^s(1-q_u)du}\Lambda(p+(1-p)\pi)sds \\ &+ e^{-\Lambda(p+(1-p)\pi)\int_t^{t_1}(1-q_u)du}\lim_{t\uparrow t_1}\mathcal{E}_t.\end{aligned}$$

Differentiating w.r.t. t implies that $\mathcal{E}_t$ solves the time ODE (subject to (2.103)):

$$\Lambda(p + (1-p)\pi)(1 - q_t)\mathcal{E}_t - \dot{\mathcal{E}}_t = \Lambda(p + (1-p)\pi)(1 - q_t)t. \tag{2.104}$$

Chapter 3

Agency Conflics and Short- vs. Long-termism in Corporate Policies[1]

3.1 Introduction

Should firms target short-term objectives or long-term performance? The question of the optimal horizon of corporate policies has received considerable attention in recent years, with much of the discussion focusing on whether short-termism destroys value. The worry often expressed in this literature is that short-termism—induced, for example, by stock market pressure—may lead firms to invest too little (see Asker, Farre-Mensa, and Ljunqvist, 2015; Bernstein, 2015; Gutierrez and Philippon, 2017, for empirical evidence). Another line of argument recognizes, however, that while firms must invest in their future if they are to have one, they must also produce earnings today to pay for doing so. In line with this view, Giannetti and Yu (2018) find that firms with more short-term institutional investors suffer smaller drops in investment and have better long-term performance than similar firms following shocks that change an industry's economic environment.

While empirical evidence relating short- or long-termism to firm performance is accumulating at a fast pace, financial theory has made little headway in developing models that characterize the optimal horizon of corporate policies or the relation between firm characteristics and this horizon. In this paper, we attempt to provide an answer to these questions through the lens of agency theory. To do so, we develop a dynamic agency model in which the agent controls both current earnings and firm growth (i.e., future earnings) through unobservable investment. In this multitasking model, the principal optimally balances the costs and benefits of incentivizing the manager over the short or long term. As shown in the paper, this can lead to optimal short- or long-termism, depending on the severity of agency conflicts and firm characteristics. Additionally, we show that the same firm can find it optimal at times to be short-termist (i.e., favor current earnings) and at other times to be long-termist (i.e., favor growth). Our findings are generally consistent with the views expressed in The Economist[2] that "long-termism and short-termism both have their virtues and vices—and these depend on context."

We start our analysis by formulating a dynamic agency model in which an investor (the principal) hires a manager (the agent) to operate a firm. In this model, agency problems arise because the manager can take hidden actions that affect both

[1]This Chapter is based on Gryglewicz et al. (2020).

[2]See "The tyranny of the long-term,"The Economist, November 22, 2014.

earnings and firm growth. As in He (2009b) and Bolton, Wang, and Yang (2017), earnings are proportional to firm size, which is stochastic and governed by a (controlled) geometric Brownian motion (i.e., subject to permanent growth shocks). In contrast with these models, earnings are also subject to moral hazard and short-term shocks that do not necessarily affect (or correlate with) long-term prospects (i.e., shocks to firm size). The agent controls the drifts of the earnings and firm size processes through unobservable investment. Notably, the agent can stimulate current earnings via short-term investment and firm growth via long-term investment.

Investment is costly and the manager can divert part of the funds allocated to investment, which requires the compensation contract to provide sufficient incentives to the agent. Under the optimal contract, the manager is thus punished (rewarded) if either cash flow or firm growth is worse (better) than expected. Because the manager has limited liability, penalties accumulate until the termination of the contract, which occurs once the manager's stake in the firm falls to zero. Since termination generates deadweight costs, maintaining incentive compatibility is costly. Based on these trade-offs, the paper derives an incentive compatible contract that maximizes the value that the principal derives from owning the firm. It then analytically demonstrates that the optimal contract can generate short- or long-termism in corporate policies, defined as short- or long-term investment levels above first-best levels.

Our theory of short- and long-termism differs from existing contributions in two important respects. First, while most dynamic agency models focus either on short- or long-term agency conflicts, we consider a multitasking framework with both long- and short-term agency conflicts. We show that agency conflicts over different horizons interact, which can generate short- and long-termism in corporate policies. Second, unlike most models on short-termism, we do not assume that focusing either on the short or the long term is optimal. In our model, the optimal corporate horizon is determined endogenously and reflects both agency conflicts and firm characteristics. These unique features allow us to generate a rich set of testable predictions about firms' optimal investment rates and the horizon of corporate policies.

A first result of the paper is to show that short- or long-termism can only arise when the firm is exposed to a dual moral hazard problem. To understand why this condition is necessary, first consider long-termism. In our model, positive growth shocks lead to a permanent increase in earnings (and risk) and to a greater misalignment between shareholders' interests and management's incentives by diluting the manager's stake in the firm. To offset these adverse dilution effects and reduce agency costs, the manager's promised wealth must increase sufficiently in response to positive growth shocks. When the firm is exposed to both long- and short-term moral hazard, the contract optimally grants the manager a larger stake in the firm, which increases potential dilution effects. The principal then counteracts these dilution effects by tying the agent's compensation more to long-term performance (i.e., long-term shocks), which leads to higher-powered long-term incentives.[3] The incentive compatibility condition with respect to long-term investment, which associates higher-powered incentives to higher levels of investments, in turn implies that the firm must also increase long-term investment, possibly beyond first-best levels. Our

[3]Throughout, we refer to the manager's exposure to earnings shocks as short-term incentives and the exposure to growth as long-term incentives. However, we stress that long-term incentives are understood as incentives for a task that concerns the long term, rather than incentive pay contingent on outcomes that realize in the distant future.

analysis demonstrates that long-termism is more likely to arise when cash flows are more volatile or when the investment technology is less efficient.

A second result of the paper is to show that short-termism can only arise if the firm is exposed to a dual moral hazard problem and there are direct externalities between short- and long-term investment. Notably, we show that a necessary condition for short-termism is that shocks to firm size and shocks to cash flows are correlated. When this correlation is negative—an assumption supported in the data (see, e.g., Gryglewicz, Mancini, Morellec, Schroth, and Valta, 2020)—we additionally show that short-termism occurs when the agent's stake in the firm is low and the risk of termination and agency costs are high. Indeed, in such instances, the benefits of long-term growth are limited. By contrast, stimulating short-term investment increases earnings and reduces the risk of termination and agency costs. Interestingly, a recent study by Barton, Manyika, and Williamson (2017) finds using a data set of 615 large- and mid-cap US publicly listed companies from 2001 to 2015 that "the long-term focused companies surpassed their short-term focused peers on several important financial measures." While our model does indeed predict that firm performance should be positively related to the corporate horizon, it in fact suggests the reverse causality.[4]

Third, we show that while Tobin's q unambiguously increases with financial slack (and in particular after positive cash flow realizations), long-term investment in general does not. This happens because long-term incentives are provided to implement optimal long-term investment and to insure the agent against dilution of her stake. Thus, long-term investment need not fully reflect the fundamental value of capital within the firm, as given by Tobin's q.

Incentives are provided in the optimal contract by making the agent's compensation contingent on firm performance via exposure to the firm's stock price and earnings. In previous dynamic contracting models, the optimal contract generates just enough incentives to the agent (i.e., incentive compatibility constraints are tight) because incentive provision comes with the threat of termination and is therefore costly to implement. A distinctive feature of our model is that the optimal contract introduces exposure to permanent shocks that is not needed to incentivize investment. In particular, the agent is provided minimal long-run incentives when the firm is close to financial distress and higher-powered long-run incentives after positive past performance when sufficient slack has been accumulated. In this region, incentives have option-like features and increase after positive performance.

To understand this result, note that when the manager's stake is large, and therefore subject to substantial dilution risk upon unexpected firm growth, it becomes optimal to mitigate these adverse dilution effect through high-powered incentive pay. This generates the distinct prediction that extra pay-for-performance is introduced when the manager's stake in the firm and dilution risk are large enough. We show indeed that in such instances the principal can eliminate dilution risk by fully exposing the manager's wealth to permanent shocks while maintaining incentive compatibility. When this is the case, long-run incentives are effectively costless and the manager is exposed to permanent, growth shocks beyond the level needed to incentivize long-term investment. In other words, positive permanent shocks lead to additional

[4]Interestingly, this causality issue is already discussed in The Economist, Schumpeter's article "Corporate short-termism is a frustratingly slippery idea," who writes: "Do short-term firms become weak or do weak firms rationally adopt strategies that might be judged short term?" Similarly, Barton, Manyika, and Williamson (2017) write in their own study "one caveat: we've uncovered a correlation between managing for the long term and better financial performance; we haven't shown that such management caused that superior performance."

pay-for-performance, and negative permanent shocks eventually eliminate this extra sensitivity to performance implied by the optimal contract. Our model therefore provides a rationale for the asymmetry of pay-for-performance observed in the data (see, e.g., Garvey and Milbourn, 2006; Francis, Iftekhar, Kose, and Zenu, 2013).

Our paper relates to the growing literature on short-termism. Influential contributions in this literature include Stein (1989), Bolton, Scheinkman, and Xiong (2006), and Aghion and Stein (2008) in which stock market pressure leads managers to boost short-term earnings at the expense of long-term value. In related work, Thakor (2018) builds a model in which short-termism is efficient, as it limits managerial rent extraction and leads to a better allocation of managers to projects. Narayanan (1985) develops a model in which short-term projects privately benefit managers by enhancing reputation and increasing wages. Von Thadden (1995) studies a dynamic model of financial contracting in which the fear of early project termination by outsiders leads to short-term biases of investment. Marinovic and Varas (2019b) and Varas (2017) develop dynamic contracting models in which the manager can undertake inefficient actions to boost short-run performance at the expense of the long run. Likewise, Zhu (2018) develops a model of persistent moral hazard in which the agent can choose between a short- and long-term action and characterizes the contract that implements the long-term action. In contrast with these models, we do not assume that focusing either on the short or the long term is optimal, and there is no intrinsic conflict between short- and long-termism in our setup. Hoffmann and Pfeil (2021) build a model in which the agent privately observe cash flows that he can divert and/or invest to increase the likelihood of adoption of future technologies.[5] Their model does not address the issue of short- versus long-termism in corporate policies.

Our modeling of cash flows with permanent and transitory shocks is similar to that in Décamps, Gryglewicz, Morellec, and Villeneuve (2017) and Hackbarth, Rivera, and Wong (2018). The model of Décamps et al. (2017) does not feature agency conflicts. The model of Hackbarth, Rivera, and Wong (2018) shows that debt financing may render short-termism optimal for shareholders. Their dynamic agency model differs from ours in that it considers different managerial preferences and focuses on the agency-induced cost of debt (overhang). Consequently, the mechanism generating short-termism is distinct from ours. Notably, short-termism only arises because debt overhang reduces the benefits of long-term investment to shareholders that, in the presence of a resource constraint, leaves more resources for short-term investment. Unlike our model, their model does not feature long-termism or asymmetric pay-for-performance.

Our paper is more generally related to the growing literature on dynamic contracting. Most contributions in this literature study agency conflicts over the short run, using a stationary environment characterized by identically and independently distributed cash flow shocks; see, for example, DeMarzo and Sannikov (2006), Biais et al. (2007), Sannikov (2008), Zhu (2012), Miao and Rivera (2016), Malenko (2019b), and Szydlowski (2015). Likewise, Biais et al. (2010) and DeMarzo et al. (2012) study dynamic contracting models with time-varying firm size in which cash flow shocks are short lived. In these models, the manager can affect current, but not directly future, firm performance. In contrast, He (2009b) and He (2011) focus on agency

[5]While Hoffmann and Pfeil (2021) find that overinvestment is more likely for firms with a superior investment technology, our model implies that overinvestment rather arises when the investment technology is inefficient. In Hoffmann and Pfeil (2021) firm size remains constant over time, thereby ruling out potential dilution of the managerial stake, so the mechanism leading to overinvestment differs from ours.

conflicts over the long run by considering a framework in which the manager can affect firm growth.[6] In these last two models, instantaneous earnings are not subject to short-term moral hazard. Our model combines both strands of the literature in a unified framework in which the optimal horizon of corporate policies arises endogenously. Our framework is also related to Ai and Li (2015) and Bolton, Wang, and Yang (2017), who study optimal investment under limited commitment. These models do not feature moral hazard. Ai and Li (2015) demonstrate that shareholders' limited commitment can lead to overinvestment in a model in which firms are subject to permanent shocks. In contrast, we assume full commitment of shareholders (the principal) and identify agency frictions as a potential driver of overinvestment.

Section 2 presents the model and its solution. Section 3 analyzes the implications of the model for optimal investment. Section 4 derives predictions on the horizon of corporate policies. Section 5 shows how the optimal contract can be implemented by exposing the manager to the firm's stock price and earnings. Section 6 focuses on asymmetric pay-for-performance. Section 7 shows the robustness of our results to alternative model specifications. Section 8 concludes.

3.2 The model

3.2.1 Assumptions

Throughout the paper, time is continuous and uncertainty is modeled by a probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathcal{P})$ with the filtration $\mathbb{F} = \{\mathcal{F}_t : t \geq 0\}$, satisfying the usual conditions. We consider a principal-agent model in which the risk-neutral owner of a firm (the principal) hires a risk-neutral manager (the agent) to operate the firm's assets. In the model, firm performance depends on investment, which can be targeted toward the short- or long-run and entails a monetary cost. Agency problems arise because investment decisions are delegated to the manager, who can take divert part of the resources allocated to investment.

The firm employs capital to produce output, whose price is normalized to one. At any time $t \geq 0$, earnings are proportional to the capital stock K_t (i.e., the firm employs an "AK" technology) and subject to permanent (long-term) and transitory (short-term) shocks. Permanent shocks change the long-term prospects of the firm and influence cash flows permanently by affecting firm size. Following He (2009b), DeMarzo et al. (2012), and Bolton, Chen, and Wang (2011), we consider that the firm's capital stock (firm size) $\{K\} = \{K_t\}_{t\geq 0}$ evolves according to the controlled geometric Brownian motion process[7]:

$$dK_t = (\ell_t \mu - \delta)\, K_t dt + \sigma_K K_t dZ_t^K, \tag{3.1}$$

where $\mu > 0$ is a constant, $\delta > 0$ is the rate of depreciation, $\sigma_K > 0$ is a constant volatility parameter, $\{Z^K\} = \{Z_t^K\}_{t\geq 0}$ is a standard Brownian motion, and ℓ_t is the firm's long-term investment choice. For the problem to be well defined, we consider that $\ell_t \in [0, \ell_{\max}]$ with $\ell_{\max} < \frac{r+\delta}{\mu}$ where $r \geq 0$ is the constant discount rate of the

[6] In a similar setting, Gryglewicz and Hartman-Glaser (2019) show that agency conflicts over the long run can lead to the early exercise of real options.

[7] This specification for capital accumulation and revenue in which capital dynamics are governed by a controlled geometric Brownian motion has been used productively in asset pricing (e.g., Cox, Ingersoll, and Ross, 1985; Kogan, 2004), corporate finance (e.g., Abel and Eberly, 2011; Bolton, Wang, and Yang, 2017), or macroeconomics (e.g., Gertler and Kiyotaki, 2010; Brunnermeier and Sannikov, 2014).

firm owner. In addition to these permanent shocks, cash flows are subject to short-term shocks that do not necessarily affect long-term prospects. Specifically, cash flows dX_t are proportional to K_t but uncertain and governed by

$$dX_t = K_t dA_t = K_t \left(s_t \alpha dt + \sigma_X dZ_t^X \right), \tag{3.2}$$

where α and σ_X are strictly positive constants, $s_t \in [0, s_{\max}]$ is the firm's short-term investment choice, and $\{Z^X\} = \{Z_t^X\}_{t \geq 0}$ is a standard Brownian motion. In the following, $\{Z^X\}$ is allowed to be correlated with $\{Z^K\}$ with correlation coefficient ρ in that[8]

$$\mathbb{E}[dZ_t^K dZ_t^X] = \rho dt, \text{ with } \rho \in (-1, 1). \tag{3.3}$$

Investment entails costs $\mathcal{I}(K_t, s_t, \ell_t)$. We assume that the investment cost is homogeneous of degree one in capital K_t, as in DeMarzo et al. (2012) and Bolton, Chen, and Wang (2011). That is, we have that $\mathcal{I}(K_t, s_t, \ell_t) \equiv K_t \mathcal{C}(s_t, \ell_t)$, where we assume $\mathcal{C}$ is increasing and convex in its arguments. Unless otherwise mentioned, we consider quadratic costs of investment

$$\mathcal{C}(s_t, \ell_t) = \frac{1}{2} \left(\lambda_s s_t^2 \alpha + \lambda_\ell \ell_t^2 \mu \right), \tag{3.4}$$

where $s_{\max}, \ell_{\max}$ are large enough to ensure that investment is interior at all times.[9]

The manager is protected by limited liability, does not accept negative payments from the principal dC_t, and cannot be asked to cover the investment cost $\mathcal{I}(K_t, s_t, \ell_t) = K_t \mathcal{C}(s_t, \ell_t)$ out of her own pocket. More specifically, the principal has to allocate funds to the manager before she can carry out the investment decisions s_t, ℓ_t. As a result, over $[t, t + dt]$ the agent is paid $dC_t + K_t \mathcal{C}(s_t, \ell_t) dt$ and wage payments net of investment cost dC_t must be positive (i.e., $dC_t \geq 0$). At any time t, the manager has full discretion over investment s_t, ℓ_t and can divert from the funds $K_t \mathcal{C}(s_t, \ell_t)$ she is handed over from the principal. In particular, the manager can change recommended short-run (respectively long-run) investment s_t (respectively ℓ_t) by any amount ε^s (respectively ε^ℓ) and keep the difference between actual investment cost

[8] In general, the correlation coefficient ρ between short-term and permanent cash flow shocks can be positive or negative. Considering, for example, the automobile industry, there is a general tendency for buyers of moving away from diesel cars toward electric cars. In the case of Volkswagen, this negative permanent demand shock on diesel cars has been compounded by the diesel gate, implying a positive correlation between short- and long-run cash flow shocks. In the case of Tesla Motors, the positive long-run demand shock on electric cars has been dampened by negative shocks on the supply chain (notably for Model 3), implying a negative correlation between short-run and long-run cash flow shocks. Additional examples of a negative correlation include decisions to invest in research and development or to sell assets. When the firm sells assets today, it experiences a positive cash flow shock. However, it also decreases permanently future cash flows. Examples of positive correlation include price changes due to the exhaustion of the existing supply of a commodity or improving technology for the production and discovery of a commodity. Using the same type of cash flow model as adopted in this paper, Gryglewicz et al. (2020) estimate that for Compustat firms between 1975 and 2014, the correlation between short-term and permanent cash flow shocks is, on average, negative.

[9] The assumption of quadratic investment cost is made merely for analytical parsimony in that all our results in Sections 1-4 hold true for any other cost function that is strictly convex in s, ℓ. This includes cost functions where short- and long-run investment are substitutes or complements, which occurs when $\frac{\partial^2 \mathcal{C}(s,\ell)}{\partial s \partial \ell} \neq 0$. We purposefully refrain from such a specification because interactions between short- and long-run investment arise endogeneously in our model and we attribute these interactions entirely to the presence of moral hazard over different time horizons. The upper bounds on the investment levels can be related to the maximum time the manager can spend on the job. The upper bound on long-term investment (i.e., $\ell_{\max} < \frac{r+\delta}{\mu}$) also naturally arises in our model as a necessary condition to obtain finite firm values.

and allocated funds, i.e.,

$$K_t\left[\mathcal{C}(s_t, \ell_t) - \mathcal{C}(s_t - \varepsilon^s, \ell_t - \varepsilon^\ell)\right]$$

for herself. Because $\{X\}$ and $\{K\}$ are subject to Brownian shocks—as long as $\sigma_X > 0$ and $\sigma_K > 0$—there is moral hazard over short- and long-term investment decision. For simplicity, we assume that diversion does not entail efficiency losses.

In the baseline version of our model, we assume the agent has sufficient private funds so that she can, in principle, also boost firm investment (i.e., implement investment $\hat{s}_t > s_t$ or $\hat{\ell}_t > \ell_t$). While this assumption does not drive our main results, it offers several advantages. First, it considerably simplifies the analysis. Second, and most importantly, it allows us to connect more easily to the existing models of He (2009) and DeMarzo et al. (2012) and to clearly demonstrate how the combination of short- and long-run moral hazard induces short- and long-termism. We analyze the case of limited private wealth in Section 5 and show that our results on short- and long-termism hold in this alternative setting.

As in DeMarzo and Sannikov (2006), Biais et al. (2007), and DeMarzo et al. (2012), the agent is more impatient than the principal and has a discount rate $\gamma > r$. As a result, the principal cannot indefinitely postpone payments to the agent. The agent possesses an outside option normalized to zero and maximizes the present value of her expected payoffs.[10] Because the agent is protected by limited liability, her continuation value can never fall below her outside option, in which case she would profit from leaving the firm. Her employment starts at time $t = 0$ and is terminated at an endogenous stopping time τ, at which point the firm is liquidated. At the time of liquidation, the principal recovers a fraction $R > 0$ of assets, valued at RK_τ. Liquidation is inefficient and generates deadweight losses.[11]

Before proceeding, note that when $\sigma_K = 0$, we obtain the environment of the dynamic agency model of DeMarzo et al. (2012) or the financing frictions model of Bolton, Chen, and Wang (2011). Since there is no noise to hide the long-term investment choice, the long-term agency conflict is irrelevant in that case. By contrast, when $\sigma_X = 0$, we obtain the cash flow environment used in the dynamic capital structure (Leland, 1994; Strebulaev, 2007) and real options literature (Carlson, Fisher, and Giammarino, 2006; Morellec and Schürhoff, 2011) as well as in the dynamic agency models of He (2009b, 2011). Since there is no noise to hide the short-term investment choice, the short-term agency conflict is irrelevant in that case.

3.2.2 The contracting problem

To maximize firm value, the investor chooses short- and long-term investment $\{s\}, \{\ell\}$ and offers a full-commitment contract to the agent at time $t = 0$, which specifies wage payments $\{C\}$, recommended investment $\{s\}, \{\ell\}$, and a termination time τ. Because the agent cannot be paid any negative amount net of investment cost, the process $\{C\}$ is nondecreasing in that $dC_t \geq 0$ for all $t \geq 0$. Moreover, the contract

[10] As in Albuquerque and Hopenhayn (2004) and Rampini and Viswanathan (2013), we could assume that the manager can appropriate a fraction of firm value so that the manager has reservation value θK_t, where $\theta \geq 0$ is a constant parameter. The entire analysis can be conducted by replacing 0 with θ.

[11] We could equally assume that the firm replaces the manager instead of liquidating when w falls to zero. The model results would remain unchanged as long as finding a new manager (i.e., replacement) is costly for the firm. For instance, one could assume some replacement cost kK_τ, which could be microfounded by a costly labor market search.

cannot request the agent to finance investment so that she is handed over the investment cost $\mathcal{I}(K_t, s_t, \ell_t)$ at time t from the principal. We let $\Pi \equiv (\{C\}, \{s\}, \{\ell\}, \tau)$ represent the contract, where all elements are progressively measurable with respect to $\mathbb{F}$. With the agent's actual investment choice $\{\hat{s}\}, \{\hat{\ell}\}$, we call a contract incentive compatible if $s_t = \hat{s}_t$ and $\ell_t = \hat{\ell}_t$ for all $t \geq 0$ and focus throughout the paper on incentive compatible contracts, where we denote the set of these contracts by $\mathbb{IC}$. Since we only consider contracts of the set $\mathbb{IC}$, we will not formally distinguish between recommended and actual investment.

For an incentive compatible contract Π, let us define the agent's expected payoff at time $t \geq 0$ (i.e., her continuation value) as

$$W_t = W_t(\Pi) \equiv \mathbb{E}_t \left[\int_t^\tau e^{-\gamma(u-t)} dC_u \right].$$

$W_t = W_t(\Pi)$ equals the promised value the agent gets if she follows the recommended path from time $t \geq 0$ onwards. $W_0 = W_0(\Pi)$ is the agent's expected payoff at inception.

The principal receives the firm cash flows net of investment cost and pays the compensation to the manager. As a result, given the contract Π, the principal's expected payoff can be written as

$$\hat{P}(W, K) \equiv \mathbb{E} \left[\int_0^\tau e^{-rt} (dX_t - K_t \mathcal{C}(s_t, \ell_t) dt - dC_t) + e^{-r\tau} R K_\tau \middle| W_0 = W, K_0 = K \right]. \tag{3.5}$$

The objective of the principal is to maximize the present value of the firm cash flows plus termination value net of the agent's compensation, where we make the usual assumption that the principal possesses full bargaining power. Denote the set of incentive compatible contracts by $\mathbb{IC}$. The investor's optimization problem reads

$$P(W, K) \equiv \max_{\Pi \in \mathbb{IC}} \hat{P}(W, K) \text{ s.t. } W_t \geq 0 \text{ and } dC_t \geq 0 \text{ for all } t \geq 0. \tag{3.6}$$

With slight abuse of notation, we denote by $\Pi \equiv (\{C\}, \{s\}, \{\ell\}, \tau)$ the solution to this optimization problem.

3.2.3 First-best short- and long-term investment

We start by deriving the value of the firm and the optimal investment levels absent agency conflicts (i.e., when there is no noise to hide the agent's action in that $\sigma_X = \sigma_K = 0$). Throughout the paper, we refer to this case as the first-best (FB) outcome.

Given the stationarity of the firm's optimization problem, the choice of s and ℓ is time invariant absent agency conflicts, and the first-best firm value reads

$$P^{FB}(K) = \max_{(s,\ell) \in [0, s_{\max}] \times [0, \ell_{\max}]} \frac{K}{r + \delta - \mu\ell} \left[\alpha s - \frac{1}{2} \left(\lambda_s \alpha s^2 + \lambda_\ell \mu \ell^2 \right) \right] \equiv K p^{FB},$$

where the short- and long-term investment choice $\{s^{FB}, \ell^{FB}\}$ maximize firm value. We denote the scaled firm value absent moral hazard by p^{FB}. Simple algebraic derivations lead to the following result:

Proposition 7 (First-best firm value and investment choices) *Assume the bounds $i_{\max}$ for $i \in \{s, \ell\}$ are such that the first-best solution is interior. Then the following holds:*

i) First-best short-term investment satisfies: $s^{FB} = \frac{1}{\lambda_s}$.

ii) First-best long-term investment satisfies: $\ell^{FB} = \frac{1}{\mu}\left[r+\delta-\sqrt{(r+\delta)^2-\frac{\mu\alpha}{\lambda_s\lambda_\ell}}\right] = \frac{p^{FB}}{\lambda_\ell}$.

3.2.4 Model solution

We now solve the model with agency conflicts over the short and long term (i.e., assuming $\sigma_K > 0$ and $\sigma_X > 0$). Recall that the contract specifies the firm's investment policy $\{s\}, \{\ell\}$, payments to the agent C, and a termination date τ all as functions of the firm's profit history. Given an incentive compatible contract and the history up to time t, the discounted expected value of the agent's future compensation is given by W_t. As in DeMarzo and Sannikov (2006) or DeMarzo et al. (2012), we can use the martingale representation theorem to show that the continuation payoff of the agent solves

$$dW_t = \gamma W_t dt - dC_t + \beta_t^s(dX_t - \alpha s_t K_t dt) + \beta_t^\ell(dK_t - (\mu\ell_t - \delta)K_t dt). \tag{3.7}$$

This equation shows that the agent's continuation value must grow at rate γ to compensate for her time preference. In addition, compensation must be sufficiently sensitive to firm performance, as captured by the processes $\beta_t^s = dW_t/dX_t$ and $\beta_t^\ell = dW_t/dK_t$, to maintain incentive compatibility. To understand why such a compensation scheme may align incentives, suppose that the agent decides to deviate from the recommended choice and chooses investment $\hat{s}_t = s_t - \varepsilon$ during an instant $[t, t+dt]$. By doing so, she keeps the amount of investment cost saved

$$K_t\big(\mathcal{C}(s_t,\ell_t) - \mathcal{C}(s_t-\varepsilon,\hat{\ell}_t)\big)dt \simeq K_t\mathcal{C}_s(s_t,\ell_t)\varepsilon dt = K_t\alpha\lambda_s s_t\varepsilon dt.$$

At the same time, however, she lowers mean cash flow by $K_t\alpha\varepsilon dt$ so that her overall compensation is reduced by $\alpha K_t\beta_t^s\varepsilon dt$. Therefore, the agent does not deviate from the prescribed short-run investment if $\beta_t^s = \lambda_s s_t$. Similarly, the agent does not deviate from the prescribed long-run investment if $\beta_t^\ell = \lambda_\ell \ell_t$. Both incentive compatibility constraints require that the agent has enough skin in the game, as reflected by sufficient exposure to firm performance.

The investor's value function in an optimal contract, given by $P(W,K)$, is the highest expected payoff the investor may obtain given K and W. While there are two state variables in our model, the scale invariance of the firm's environment allows us to write $P(W,K) = Kp(w)$ and reduce the problem to a single state variable: $w \equiv \frac{W}{K}$, the scaled promised payments to the agent as in He (2009b) and DeMarzo et al. (2012).

To characterize the optimal compensation policy and its effects on the investor's (scaled) value function $p(w)$, note that it is always possible to compensate the agent with cash so that it costs at most \$1 to increase w by \$1 and $p'(w) \geq -1$. In addition, as shown by Eq. (3.7), deferring compensation increases the growth rate of W (and of w) and thus lowers the risk of liquidation but is costly due to the agent's impatience, $\gamma > r$. As a result, the optimal contract sets $dc \equiv \frac{dC}{K}$ to zero for low values of w and only stipulates payments to the manager once the firm has accumulated sufficient slack. That is, there exists a threshold $\overline{w}$ with

$$p'(\overline{w}) = -1 \text{ and } dc = \max\{0, w - \overline{w}\}, \tag{3.8}$$

where the optimal payout boundary is determined by the super-contact condition:

$$p''(\overline{w}) = 0. \tag{3.9}$$

When w falls to zero, the contract is terminated and the firm is liquidated so that

$$p(0) = R. \tag{3.10}$$

When $w \in [0, \overline{w}]$, the agent's compensation is deferred and $dc = 0$. The Hamilton-Jacobi-Bellman equation for the principal's problem is then given by (see Appendix 3.9.2):

$$\begin{aligned}(r+\delta)p(w) = \max_{s,\ell,\beta^s,\beta^\ell} \Bigg\{ & \alpha s - \mathcal{C}(s,\ell) + p'(w)w(\gamma+\delta-\mu\ell) + \mu\ell p(w) \\ & + \frac{p''(w)}{2}\Big[(\beta^s\sigma_X)^2 + \sigma_K^2(\beta^\ell - w)^2 + 2\rho\sigma_X\sigma_K\beta^s(\beta^\ell - w)\Big]\Bigg\},\end{aligned} \tag{3.11}$$

subject to the incentive compatibility constraints on β^s and β^ℓ.

Due to the scale invariance (i.e., $P(W, K_0) = p(w)K_0$) the investor's maximization problem at $t = 0$ can now be rewritten as

$$\max_{w_0 \in [0,\overline{w}]} p(w_0)K_0,$$

with unique solution $w_0 = w^*$ satisfying

$$p'(w^*) = 0. \tag{3.12}$$

As a consequence, the principal initially promises the agent utility w^*K_0 and expects a payoff $P(K_0w^*, K_0) = p(w^*)K_0$. For convenience, we normalize K_0 to unity in the following and refer to $p(w^*)$ as expected payoff instead of scaled expected payoff. The following proposition summarizes our results about the optimal contract. Its proof is deferred to Appendix 3.9.2.

Proposition 8 (Firm value and optimal compensation under agency)
Let $\Pi \equiv (\{C\}, \{s\}, \{\ell\}, \tau)$ denote the optimal contract solving problem (3.6). *The following holds true:*

1. *There exist $\mathbb{F}$-progressive processes $\{\beta^\ell\}$ and $\{\beta^s\}$ such that the agent's continuation utility W_t evolves according to* (3.7). *The optimal contract is incentive compatible in that $\beta^s = \lambda_s s$ and $\beta^\ell = \lambda_\ell \ell$, where $\{s\}, \{\ell\}$ are the firm's optimal investment decisions.*

2. *Firm value is proportional to firm size in that $P(W, K) = Kp(w)$. The scaled firm value $p(w)$ is the unique solution to Eq.* (3.11) *subject to* (3.8), (3.9), *and* (3.10) *on $[0, \overline{w}]$. For $w > \overline{w}$, the scaled value function satisfies $p(w) = p(\overline{w}) - (w - \overline{w})$. Scaled cash payments $dc = \frac{dC}{K}$ reflect w back to $\overline{w}$.*

3. *The function $p(w)$ is strictly concave on $[0, \overline{w})$.*

Before proceeding, note that w serves as a proxy for the firm's financial slack in our model so that states where w is close to zero—and the firm close to liquidation—correspond to financial distress. Since the firm has to undergo inefficient liquidation after a series of adverse shocks drive w down to zero, the principal becomes effectively risk averse with respect to the volatility of w so that the value function is strictly concave (i.e., $p''(w) < 0$ for $w < \overline{w}$). Put differently, the concavity of p implies that the principal would like to minimize the volatility of w while maintaining incentive compatibility.

Note also that overall value, $P(W,K) + W$, is split between the principal and the manager, where the manager obtains a fraction,

$$\mathcal{S}(w) = \frac{W}{P(W,K) + W} = \frac{w}{p(w) + w},$$

of overall value. Because of $\mathcal{S}'(w) > 0$ for all $w \in (0, \overline{w})$, the scaled continuation value w corresponds (monotonically) to the fraction of overall firm value that goes to the manager.[12] Therefore, we also refer to w as the agent's or manager's stake in the firm. When the manager's stake w falls down to zero, she has no more incentives to stay and accordingly leaves the firm. In this case, deadweight losses are incurred due to contract termination.

3.3 Short- versus long-run incentives

This section examines the implications of agency conflicts for long- and short-term investment choices. For clarity of exposition, we assume that the correlation between short- and long-run shocks ρ is zero and the parameters are such that investment levels s and ℓ are interior. Section 3.4.2 analyzes the effects of nonzero correlation.

3.3.1 Short-term investment and incentives

Optimal short-term investment $s = s(w)$ is obtained by taking the first-order condition in Eq. (3.11) after using the incentive compatibility condition $\beta^s = \lambda_s s$. This yields the following result:

Proposition 9 (Optimal short-term investment) *Optimal short-term investment is given by*

$$s(w) = \frac{\overbrace{\alpha}^{\substack{\text{Direct benefit}\\ \text{of investment}}}}{\underbrace{\lambda_s \alpha}_{\substack{\text{Direct cost}\\ \text{of investment}}} \underbrace{- p''(w)(\lambda_s \sigma_X)^2}_{\substack{\text{Agency cost}\\ \text{of investment}}}}. \tag{3.13}$$

Short-term investment is strictly lower than under first-best except at the boundary in that $s(w) < s^{SB}$ for $w < \overline{w}$ and $s(\overline{w}) = s^{FB}$. If $\gamma - r$ and σ_K are sufficiently small, then $s(w)$ increases in w (i.e., $\frac{\partial s(w)}{\partial w} > 0$).

An important implication of Proposition 9 is that agency conflicts lead to underinvestment for the short run (i.e., $s(w) < s^{FB}$ when $\rho = 0$). Upon increasing the investment rate s, the firm does not only incur direct, monetary cost of investment but also agency costs because higher s requires higher incentives β^s. Consequently, the agent's stake becomes more volatile, which raises the risk of costly liquidation

[12] In fact, $\mathcal{S}'(w) \propto p(w) - p'(w)w$, which is always positive since $\frac{\partial}{\partial w}(p(w) - wp'(w)) = -wp''(w) > 0$ due to concavity of the value function. Since $\mathcal{S}'(0) = R \geq 0$, it follows that $\mathcal{S}'(w) > 0$ for all $w \in (0, \overline{w}]$.

and therefore leads to endogeneous agency costs or incentive costs of investment. These agency costs decrease in the level of financial slack w and vanish at the payout boundary $\overline{w}$ where $p''(\overline{w}) = 0$, at which point the firm's short-run investment reaches first best, $s(\overline{w}) = s^{FB}$.

3.3.2 Long-term incentives and investment

Next, we characterize the firm's optimal long-term investment ℓ and the agent's long-term incentives β^{ℓ}. Using the HJB Eq. (3.11) and the incentive compatibility condition $\beta^{\ell} = \lambda_{\ell}\ell$, we get the following result:

Proposition 10 (Optimal long-term investment) *Optimal long-term investment is given by*

$$\ell(w) = \frac{\overbrace{\mu(p(w) - p'(w)w)}^{\substack{\text{Direct benefit}\\ \text{of investment}}} \overbrace{-p''(w)w\lambda_{\ell}\sigma_K^2}^{\text{Reducing dilution risk}}}{\underbrace{\lambda_{\ell}\mu}_{\substack{\text{Direct cost}\\ \text{of investment}}} \underbrace{-p''(w)(\lambda_{\ell}\sigma_K)^2}_{\substack{\text{Agency cost}\\ \text{of investment}}}}. \tag{3.14}$$

The firm always underinvests for the long-term close to the boundary in that there exists $\varepsilon > 0$ *such that* $\ell(w) < \ell^{FB}$ *for* $w \in [\overline{w} - \varepsilon, \overline{w}]$.

To get some intuition for the results in Proposition 10, let us consider the costs and benefits from marginally increasing long-term investment ℓ:

$$\frac{\partial p(w)}{\partial \ell} \propto \underbrace{\mu(p(w) - p'(w)w)}_{\text{Direct benefit}} \underbrace{-\lambda_{\ell}\mu\ell}_{\substack{\text{Direct}\\ \text{cost}}} + \underbrace{p''(w)\ell(\lambda_{\ell}\sigma_K)^2}_{\text{Agency cost}} \underbrace{-p''(w)w\lambda_{\ell}\sigma_K^2}_{\substack{\text{Reducing dilution}\\ \text{risk}}}. \tag{3.15}$$

Consider first the costs of raising long-term investment. The above expression shows that, in addition to the direct cost of investment, the firm incurs an agency cost. This agency cost arises because increasing long-run investment requires higher long-run incentives β^{ℓ} and therefore makes w more volatile. The agency cost of investment depends on the principal's effective risk aversion $-p''(w)$ and decreases optimal investment $\ell(w)$.

Consider next the benefits of raising long-term investment. The first difference between optimal short- and long-term investment is that the direct benefit of long-term investment is time varying and given by $p(w) - p'(w)w$. This value represents Tobin's marginal q in our model, as it equals the marginal value of capital $P_K(W, K)$. Note that long-term investment expenditures today lead to a higher average cash flow rate in the future. However, due to the possibility of firm liquidation owing to the moral hazard problem, the firm cannot perpetually enjoy this increase in the cash flow rate, so the benefit of long-term investment $p(w) - p'(w)w$ is strictly lower than p^{FB}. Ceteris paribus, this lowers the firm's investment rate $\ell(w)$. Remarkably, in contrast to the case of short-term investment, long-term investment is below the first-best level for high w close to $\overline{w}$. The reason is that while the agent becomes a residual claimant on cash flows at $\overline{w}$, she is not a residual claimant on the benefits of long-term growth at the first-best level because of the agency-induced firm liquidation in the future. It holds, however, that long-term investment is more profitable when the firm has more financial slack and the distance to liquidation is far (i.e., $p(w) - wp'(w)$ increases in w).

A second difference is that investment in $\ell(w)$ offers an additional benefit compared to investment in $s(w)$: it mitigates the dilution of the agent's stake w. Since $p''(w) \leq 0$, this effect unambiguously increases long-term investment. To understand the source of this effect, first note that by Ito's lemma, the dynamics of the agent's stake are given by[13]

$$dw = (\gamma + \delta - \mu\ell)wdt + \beta^s \sigma_X dZ^X + (\beta^\ell - w)\sigma_K dZ^K, \tag{3.16}$$

so the instantaneous variance of dw satisfies

$$\Sigma(w) \equiv \frac{\mathbb{V}(dw)}{dt} = (\beta^s \sigma_X)^2 + (\beta^\ell - w)^2 \sigma_K^2. \tag{3.17}$$

From Eq. (3.16), we see that a positive permanent shock $dZ^K > 0$ has two opposing effects on the manager's incentives. First, the agent is rewarded for strong performance via the sensitivity β^ℓ and is promised higher future payments W. This increases $w = \frac{W}{K}$ (via its numerator) by $\beta^\ell \sigma_K dZ^K$, which equals $\lambda_\ell \ell(w)\sigma_K dZ^K$. Second, firm size K grows more than expected, thereby reducing the agent's stake $w = \frac{W}{K}$ (via its denominator) by $-w\sigma_K dZ^K$. We refer to the reduction of the agent's stake upon a positive shock $dZ^K > 0$ as dilution and the volatility generated by this effect (i.e., $-w\sigma_K$), as dilution risk. Altogether, we have that $dw/dZ^K = (\beta^\ell - w)\sigma_K$. Because performance-based compensation and dilution move w in opposite directions, long-run incentives β^ℓ mitigate the dilution effect that, ceteris paribus, lowers risk (see Eq. (3.17)) and is thus beneficial. This makes contracting for long-term investment cheaper and increases $\ell(w)$.

More generally, our model suggests that the manager's compensation should increase with firm size. Indeed, an increase in firm size (due to a positive permanent shock $dZ^K > 0$) raises both the firm's future cash flow rate and the magnitude of future cash flow shocks. As a result, the firm becomes not only more profitable but also more risky (in absolute terms). Both effects call for an increase in the manager's continuation value, which better aligns the manager's and the principal's interests and facilitates contracting for long-term investment.

It is illustrative to look at this effect from the perspective of agency costs. As long as $\beta^\ell < w$, raising β^ℓ lowers the volatility and instantaneous variance $\Sigma(w)$ of w, and therefore the risk of liquidation, so that the effective (marginal) agency cost of long-run investment is pinned down by the net change in risk, that is, by

$$\underbrace{-p''(w)\ell(\lambda_\ell \sigma_K)^2}_{\text{Agency cost }(>0)} + \underbrace{p''(w)w\lambda_\ell \sigma_K^2}_{\substack{\text{Reducing dilution}\\ \text{risk }(<0)}} = \underbrace{-p''(w)\sigma_K^2 \lambda_\ell(\lambda_\ell \ell - w)}_{\text{Effective agency cost }(0)}. \tag{3.18}$$

As is the case with the agency cost of investment, the benefits of mitigating dilution risk depend on how much volatility in w matters for the investor's value function (i.e., on principal's effective risk aversion $-p''(w)$). Therefore, it is most beneficial to alleviate dilution via long-run incentives β^ℓ when the concavity of the scaled value function is the largest. The effect disappears at $w = \overline{w}$ where $p''(\overline{w}) = 0$. When w is close to $\overline{w}$, and therefore $p''(w) \simeq 0$ and $p'(w) \simeq -1$, the firm always underinvests because direct benefits of investment $p(w) - wp'(w) \simeq p(w) + w < p^{FB}$

[13]As discussed in Appendix 3.9.2, the dynamics in Eq. (3.16) are under an auxiliary measure $\tilde{\mathcal{P}}$ rather than under the physical measure $\mathcal{P}$. The choice of the probability measure does not matter since w has the same volatility under both measures and volatility is the only quantity we study in in the following discussion.

are reduced by the presence of moral hazard and agency-induced firm liquidation, which implies $\ell(w) = (p(w) + w)/\lambda_\ell < p^{FB}/\lambda_\ell = \ell^{FB}$.

3.4 Short- and long-termism in corporate policies

Because the manager's ability to divert funds decreases the benefits of investment, each moral hazard problem working in isolation leads to underinvestment relative to the first-best levels. The novel insight of our model is that a simultaneous moral hazard problem over both the short and long run can generate overinvestment. We call overinvestment for the long run (i.e., $\ell > \ell^{FB}$), long-termism and overinvestment for short-run (i.e., $s > s^{FB}$), short-termism. Below we analyze and contrast the circumstances that lead to long-termism and short-termism. We find that long-termism can arise irrespective of whether the different sources of cash flow risk are correlated, while short-termism requires $\rho \neq 0$.

3.4.1 Long-termism

Proposition 10 and Eq. (3.14) reveal that moral hazard decreases long-run investment via the direct benefit channel and the agency cost channel. The firm can potentially overinvest to reduce dilution risk. In the next proposition, we show that the last effect can dominate the former two effects and present sufficient conditions for overinvestment to arise.

Proposition 11 (Long-termism) *The following holds true:*

i) Long-termism (i.e., $\ell(w) > \ell^{FB}$) arises only if $\sigma_X > 0$ and $\sigma_K > 0$.
ii) Assume $\sigma_X > 0$ and $\sigma_K > 0$. Then, there exist w^L and w^H with $0 < w^L < w^H < \overline{w}$ such that $\ell(w) > \ell^{FB}$ for $w \in (w^L, w^H)$, provided that μ and $\gamma - r$ are sufficiently low. The firm underinvests (i.e., $\ell(w) < \ell^{FB}$) when $w < w^L$ or $w > w^H$ (i.e., when w is close to zero or close to $\overline{w}$).
iii) Higher volatility $\sigma_X > 0$ or $\sigma_K > 0$ favors long-termism: if μ is sufficiently low and parameters are such that $\sup\{\ell(w) : 0 \leq w \leq \overline{w}\} = \ell^{FB}$, then there exists $\varepsilon > 0$ such that $\sup\{\ell(w) : 0 \leq w \leq \overline{w}\} > \ell^{FB}$ if σ_X or σ_K increases by ε.

The first part of Proposition 11 states that long-termism can only arise when firm cash flows are subject to both transitory and permanent shocks (i.e., when $\sigma_X > 0$ and $\sigma_K > 0$), and the firm is exposed to a simultaneous moral hazard problem over both the short and long run. When permanent cash flow shocks are removed from the model (i.e., $\sigma_K = 0$), long-term investment ℓ is observable and contractible. In addition, there is no risk of dilution of the agent's stake, as all shocks are purely transitory in nature. Under these circumstances, long-term investment satisfies

$$\ell(w) = \frac{p(w) - wp'(w)}{\lambda_\ell} < \frac{p^{FB}}{\lambda_\ell} = \ell^{FB}.$$

Because short-run agency lowers the direct benefits of long-run investment, the firm always underinvests for the long term.

To see why transitory shocks, or equivalently moral hazard over the short term, are essential for long-termism, we start with the following observation. Since the direct benefit of long-term investment under moral hazard is below the first-best level, it follows from equation (3.15) that a necessary condition for overinvestment in $\ell(w)$ is that the dilution effect exceeds the agency cost effect. Using Eq. (3.18), this is equivalent to requiring that the effective (marginal) agency cost is negative. Thus, overinvestment in ℓ or long-termism arises only if

$$-p''(w)(\lambda_\ell \ell - w)\lambda_\ell \sigma_K^2 < 0 \Longleftrightarrow w > \lambda_\ell \ell = \beta^\ell,$$

that is, if the manager's stake is large relative to her long-term incentives. When $\sigma_X = 0$, the firm faces no transitory cash flow risk and therefore optimally grants the manager a relatively low stake, which puts a limit on potential dilution effects. More specifically, if it were that $w > \beta^\ell = \lambda_\ell \ell$, it follows from Eq. (3.17) that the firm would profit from decreasing w by making infinitesimal payouts $dc > 0$ and thus reducing the risk in w by

$$\Sigma(w) - \Sigma(w - dc) \simeq (w - \lambda_\ell \ell)dc > 0.$$

This strategy would reduce the risk the manager is exposed to and still provide sufficient incentives. Consequently, $\sigma_X = 0$ implies that $w \leq \lambda_\ell \ell$ for all w, the effective agency cost of long-term investment is positive, and the firm underinvests in ℓ.[14]

When both σ_X and σ_K are strictly positive, the above argument does not work, as the firm also needs to account for short-run risk and incentives. To decrease termination risk, it can then be optimal for the firm to delay payments to the manager further, even if $w \simeq \beta^\ell$ and the manager's stake is barely exposed to permanent cash flow risk. This can lead to w exceeding β^ℓ, that is, to a negative effective agency cost and to overinvestment in ℓ. The mechanism is as follows. When the agent holds a large stake w, the risk of dilution identified above generates additional termination risk, which diminishes the risk reduction induced by postponing payouts. The principal can mitigate these adverse dilution effects by tying the agent's compensation more to long-term performance, which leads to higher long-run incentives β^ℓ. The incentive compatibility condition $\beta^\ell = \lambda_\ell \ell$ then implies that the firm must also increase long-term investment.

The second part of Proposition 11 shows that long-termism arises when the asset growth rate μ is low (i.e., when long-run investment is sufficiently inefficient). Proposition 11 therefore offers a potential explanation for the puzzling empirical evidence that in recent years capital is not allocated to the industries with the best growth opportunities (as recently shown by Lee, Shin, and Stulz, 2018). Additionally, long-termism arises when cash flow is sufficiently volatile in either time-horizon (i.e., $\sigma_X > 0$ and $\sigma_K > 0$ are large), and when the agent is sufficiently patient (i.e., $\gamma - r > 0$ is low).

The intuition for these findings is as follows. As explained above, long-termism requires the dilution effect to exceed the agency cost effect, which happens when the manager's stake is large relative to her long-term incentives, $w > \beta^\ell$. When this is the case, the effective agency cost is negative. Both higher cash flow risk (σ_X and σ_K) and lower cost of delaying payouts ($\gamma - r$) increase the value of deferred compensation so that $\overline{w}$ rises, leading to an (average) increase in the manager's stake within the firm. On the other hand, low asset growth rate decreases contracted long-term investment ℓ and accordingly long-term incentives β^ℓ.

[14] In fact, the inequality is strict: $w < \beta^\ell = \lambda_\ell \ell(w)$. To get some intuition, note that in case $\beta^\ell = w$, the firm becomes riskless. The benefits of reducing w by an infinitesimal amount $dc > 0$ are proportional to $(\gamma - r)o(dc)$ and therefore of order $o(dc)$, while the cost—stemming from the additional risk of liquidation—are of order $o((dc)^2)$. Consequently, the firm would never set $\beta^\ell = w$ so that $\beta^\ell = \lambda_\ell \ell(w) > w$ for all $w \in [0, \overline{w}]$. This result is established in He (2009).

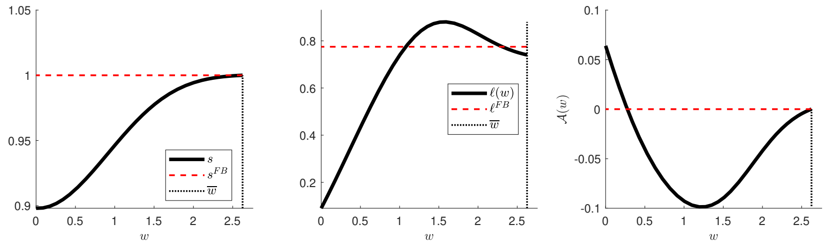

FIGURE 3.1: Numerical example of long-termism. The first two panels depict optimal investment as functions of w. The third panel at the right displays effective agency cost $\mathcal{A}(w) = -p''(w)(\lambda_\ell \ell(w) - w)$. The parameters are $\alpha = 0.25$, $\sigma_K = 0.25$, $\sigma_X = 0.2$, $\rho = 0$, $\mu = 0.025$, $r = 0.046$, $\gamma = 0.048$, $\delta = 0.125$, $\lambda_s = \lambda_\ell = 1$, and $R = 0.2$.

To generate long-termism, the agency-cost-based motives for overinvestment must also exceed the preference for underinvestment that arises because of the diminished direct benefit of investment. Recall that the marginal direct benefit of long-run investment under moral hazard equals $\mu(p(w) - p'(w)w)$ and is below its first-best counterpart, while the marginal direct cost $\lambda_\ell \mu$ is at the first-best level. Since both the direct benefit and cost are proportional to μ, this motive to underinvest is quantitatively low when μ is low and can then be overcome by the agency-cost-based preference for overinvestment.

Last, we note that while marginal q, $p(w) - wp'(w)$, unambiguously increases in w, long-term investment in general does not. This is because the marginal cost of capital—which includes the direct cost and the effective agency cost—is time varying. In particular, as explained above, the effective agency cost can increase in w. Thus long-term investment and marginal q can move in opposite directions. In general, the investment sensitivity with respect to marginal q possesses an ambiguous sign.[15]

Fig. 3.1 presents a quantitative example illustrating long-termism. The parameters satisfy the conditions set in Proposition 11 and are as follows. We set the discount rate parameters to $r = 4.6\%$ and $\gamma = 4.8\%$ and the depreciation rate to $\delta = 12.5\%$, similar to DeMarzo et al. (2012). The volatility parameter of the long-term shock is set to $\sigma_K = 20\%$, in line with Kogan (2004), while the volatility parameter of the short-term shock is set to $\sigma_X = 25\%$, in line with DeMarzo et al. (2012). The drift parameter for the profitability process is set to $\alpha = 25\%$. The left plot shows that the firm underinvests in the short run for all w. The middle plot shows that the firm overinvests in the long run for intermediate values of w. This is when the dilution effect, whose magnitude is proportional to $p''(w)w\sigma_K^2$, is the strongest. The right plot also shows that long-termism is related to a negative effective agency cost. Conversely, according to Proposition 11, long-termism never arises in financial distress (i.e., when w is close to 0), or when the firm is expected to make direct payments to the manager (i.e., when w is close to $\overline{w}$).

[15]Bolton, Chen, and Wang (2011) also show that investment can be decreasing in marginal q when the marginal source of financing is credit line.

3.4.2 Correlated cash flow shocks and short-termism

As shown in Proposition 9, short-termism cannot occur in our baseline model with independent shocks (i.e., when $\rho = 0$). When permanent and transitory cash flow shocks are correlated, direct externalities between short- and long-term investment and incentives arise. These externalities can lead to corporate short-termism (i.e., to $s > s^{FB}$), as we demonstrate below.

To start with, note that when shocks are correlated, optimal short- and long-term investment are given by

$$s(w) = \frac{\alpha + \overbrace{p''(w)\rho\sigma_X\sigma_K\lambda_s\left(\lambda_\ell \ell(w) - w\right)}^{\text{Externality}}}{\lambda_s\alpha - p''(w)(\lambda_s\sigma_X)^2} \tag{3.19}$$

and

$$\ell(w) = \frac{\mu\left(p(w) - p'(w)w\right) + \overbrace{p''(w)\rho\sigma_X\sigma_K\lambda_\ell\lambda_s s(w)}^{\text{Externality}} - p''(w)w\lambda_\ell\sigma_K^2}{\lambda_\ell\mu - p''(w)(\lambda_\ell\sigma_K)^2}. \tag{3.20}$$

Compared to Eq. (3.13) and (3.14), new terms appear that affect optimal investment levels and incentives. Since $s(w)$ depends on $\ell(w)$ and vice versa, there are direct externalities between investment levels and incentives. Intuitively, when the two sources of risk are positively correlated, exposing the manager's continuation payoff to both transitory and permanent shocks creates additional volatility and is therefore costly. Conversely, when the correlation is negative, exposure to both shocks partially cancels out, thereby reducing the volatility of the manager's continuation payoff w.

From Eq. (3.20), the externality of $s(w)$ on $\ell(w)$ is negative (positive) if $\rho > 0$ ($\rho < 0$). The magnitude of the externality scales with the curvature of the value function $p''(w)$ (i.e., the principal's effective risk aversion) and is therefore relatively weaker once w is sufficiently large and the risk of termination is sufficiently remote.

Likewise, Eq. (3.19) demonstrates that the choice of long-term investment $\ell(w)$ also feeds back into the choice of short-term investment $s(w)$. However, the externality effect in the numerator of $s(w)$ in Eq. (3.19) has two separate components:

$$p''(w)\rho\sigma_X\sigma_K\lambda_s\left(\lambda_\ell\ell(w) - w\right) = \underbrace{p''(w)\rho\sigma_X\sigma_K\lambda_s\lambda_\ell\ell(w)}_{\text{Agency cost}}\underbrace{-p''(w)\rho\sigma_X\sigma_K\lambda_s w}_{\text{Reducing dilution risk}}. \tag{3.21}$$

This decomposition shows that when the correlation between shocks is nonzero, incentives for the short run are also used to counteract the dilution in the manager's stake arising upon positive permanent shocks $dZ^K > 0$. As discussed in Section 3.4.1, with no correlation, the principal counteracts this dilution effect by tying the manager's compensation to permanent shocks and increasing long-term incentives. When the two sources of cash flow risk are positively (negatively) correlated, it is possible to reduce dilution risk also by means of higher (lower) short-term incentives.

Notably, when $\rho < 0$ and w is low, positive risk externalities of short- and long-term incentives emerge and may dominate dilution effects of short-term incentives. In this case, short-termism, $s(w) > s^{FB}$, can become optimal.

Proposition 12 (Short-termism under distress with $\rho < 0$) *The following holds true:*

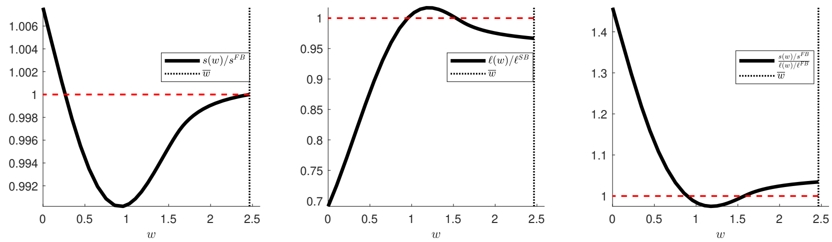

FIGURE 3.2: Numerical example of short-termism. The parameters are $\alpha = 0.25$, $\sigma_X = 0.15$, $\sigma_K = 0.5$, $\rho = -0.75$ $\mu = 0.025, r = 0.046$, $\gamma = 0.047$, $\delta = 0.125$, $\lambda_s = 1.15, \lambda_\ell = 0.25$, and $R = 0.75$.

i) *Short-termism arises only if $\sigma_X > 0$, $\sigma_K > 0$, and $\rho \neq 0$. Conversely, if either $\sigma_X = 0$, $\sigma_K = 0$, or $\rho = 0$, short-termism cannot arise and $s(w) \leq s^{FB}$ for all w.*
ii) *Assume $\sigma_X > 0$, $\sigma_K > 0$, and $\rho < 0$. Then, there exist $w^L < w^H$ with $s(w) > s^{FB}$ for $w \in (w^L, w^H)$, provided σ_X is sufficiently small. When, in addition, λ_ℓ and $\gamma - r$ are sufficiently small, the set $\{w \in [0, \overline{w}] : s(w) > s^{FB}\}$ is convex and contains zero and $s(w)$ decreases on this set.*

While long-termism occurs mainly for large values of the manager's stake w with the objective to alleviate the excessive dilution risk via long-run incentives β^ℓ, short-termism is more likely to occur for low values of w when the correlation between shocks is negative. When the agent's stake w is small, dilution risk is negligible and positive externalities between short- and long-term incentives induce more short-term investment. In addition, short-termism can arise when cash flow risk σ_X is small so that short-run agency cost is sufficiently low and does not dominate the externality effect.

Fig. 3.2 provides an example of short-termism when the correlation between long- and short-term shocks is negative. Consistently with Proposition 12, the firm overinvests in the short-run when in distress and w is close to zero. Fig. 3.2 further illustrates that both short- and long-termism may, but need not happen, within the same firm, depending on the level of financial slack as measured by w. In distress, the firm overinvests in generating (short-term) profits, while after a strong performance, the firm overinvests in (long-term) growth. While the effects of absolute short-termism appear to be quantitatively small, the effects of relative short-termism $\frac{s(w)/s^{FB}}{\ell(w)/\ell^{FB}}$, which determines whether investment is distorted toward the short-term compared to first-best, can be quantitatively large. Absent agency fictions, this ratio equals, by construction, one, and a value above (below) one indicates an investment distortion toward the (long) short run. The right-hand side plot of Fig. 3.2 presents the relative short-termism ratio, which, for our parameter values, is a nonmonotonic U-shaped function of w. The relative short-termism for large w close to $\overline{w}$ arises for all parameters (compare Proposition 9 and 10). The ratio is below one for intermediate w whenever substantial absolute long-termism arises. Relative short-termism again dominates for low w, and this region exists due to negative ρ and relatively low cost of short-term investment (low σ_X and λ_ℓ; cf. Proposition 12).

Our focus on the case of negative correlation is due to the findings in Chang, Dasgupta, Wong, and Yao (2014) and Gryglewicz et al. (2020) that the correlation coefficient between permanent and short-term cash flow shocks ρ is, on average, negative. When this is the case, our model predicts that firms with a high risk of liquidation (i.e., firms that perform worse and have little financial slack) should find it optimal to focus on the short term (i.e., current earnings) while firms with a low risk of liquidation (i.e., cash-rich firms that perform well) should find it optimal to focus on the long term (i.e., asset growth). Interestingly, a recent study by Barton, Manyika, and Williamson (2017) find, using a data set of 615 large- and mid-cap US publicly listed companies from 2001 to 2015, "the long-term focused companies surpassed their short-term focused peers on several important financial measures." While our model does indeed predict that firm performance should be positively related to the corporate horizon, it in fact suggests the reverse causality.

For completeness, we also investigate optimal investment when the correlation between cash flow shocks is positive. In this case, the firm can overinvest in both short- and long-term investment at the same time. This happens when the agent's stake in the firm is large, thereby exposing the manager to a high risk of dilution. To reduce this dilution risk, the principal provides high-powered incentives to the manager. Importantly, when correlation is positive, unexpected asset growth $dZ^K > 0$ triggers, on average, unexpected cash flow ρdZ^K, which leads a reward $(\beta^\ell + \rho\beta^s)dZ^K$ for the agent.[16] Consequently, both short- and long-run incentives counteract the adverse dilution in the agent's stake so that the desire to mitigate dilution risk translates into high-powered incentives and, accordingly, to overinvestment for both time horizons. The next proposition characterizes this outcome.

Proposition 13 (Short-termism with $\rho > 0$) *Assume $\sigma_X > 0$, $\sigma_K > 0$ and $\rho > 0$. Then, there exist $w^L < w^H$ with $s(w) > s^{FB}$ for $w \in (w^L, w^H)$, provided $\sigma_X > 0$ and $\gamma - r$ is sufficiently small. When in addition μ is sufficiently small, the set $\{w \in [0, \overline{w}] : s(w) > s^{FB}\}$ is convex with $\inf\{w \in [0, \overline{w}] : s(w) > s^{FB}\} > 0$ and $\sup\{w \in [0, \overline{w}] : s(w) > s^{FB}\} = \overline{w}$.*

3.5 Incentive contracts contingent on stock prices

The optimal contract provides short- and long-run incentives, β_t^s and β_t^ℓ, by conditioning the agent's compensation on earnings and asset size. In practice, executive compensation is commonly linked to stock prices (via stock and option grants) and to accounting results (via performance-vesting provisions of these grants and via performance-based bonuses). The use of both stock prices and accounting measures in designing executive compensation has been increasing over time. Stock and option grants constitute a majority of CEO compensation (Edmans, Gabaix, and Jenter, 2017). A majority of equity grants have accounting-based performance-vesting provisions with earnings being the most common metric (Bettis, Bizjak, Coles, and Kalpathy, 2018). We now show that the optimal contract implied by our model is broadly consistent with these patterns and can be implemented by exposing the manager to stock prices and earnings.

We start with writing the dynamics of earnings and stock prices. The firm's (instantaneous) earnings net of investment cost are given by

$$dE_t = (\alpha s_t - \mathcal{C}(s_t, \ell_t))K_t dt + K_t \sigma_X dZ_t^X,$$

[16] To see this, one can decompose $dZ_t^X = \rho dZ_t^K + \sqrt{1-\rho^2} dZ_t^T$, where $\{Z^T\}$ is a standard Brownian motion, independent of $\{Z^X\}$. Hence, $\mathbb{E}(Z_t^X | Z_t^K) = \rho Z_t^K$ or in differential form $\mathbb{E}(dZ_t^X | dZ_t^K) = \rho dZ_t^K$.

while the stock price (which, with full equity financing and the total share supply normalized to one, is equivalent to firm value) evolves according to[17]

$$\frac{dP_t}{P_t} = \mu_t^P dt + \Sigma_t^X dZ_t^X + \Sigma_t^K dZ_t^K,$$

where the expressions for μ_t^P, Σ_t^X, and Σ_t^K are given in the Online Appendix. The principal provides the incentives to the manager by choosing the manager's exposures to earnings and stock price changes, respectively defined by

$$\beta_t^E = \frac{dW_t}{dE_t} \text{ and } \beta_t^P = \frac{dW_t}{dP_t}.$$

The exposures β_t^E and β_t^P are set so as to generate the required short- and long-run incentives under the optimal contract. The Online Appendix derives the following expressions for the exposures implied by the optimal contract

$$\beta_t^P = \lambda_\ell \ell_t \left(\frac{1}{p(w_t) + p'(w_t)(\lambda_\ell \ell_t - w_t)} \right)$$

and

$$\beta_t^E = \lambda_s s_t \left(\frac{p(w_t) - p'(w_t) w_t}{p(w_t) + p'(w_t)(\lambda_\ell \ell_t - w_t)} \right).$$

An appropriate exposure to the firm's stock price (which takes into account the nonlinear relation between stock price and asset size) provides the right amount of long-run incentives. It additionally provides some short-run incentives, as the stock price is also subject to short-run shocks. The exposure to earnings is set to provide the required residual exposure to short-run shocks. This characterization of the optimal contract highlights an important implication of our model: while stock prices account for both short- and long-run shocks to firm value, exposing the manager solely to the firm's stock price cannot in general provide a right mix of short- and long-run incentives. To achieve optimal incentives, the manager also needs to be exposed to short-run accounting performance metrics such as earnings.

Last, we analyze how the manager's exposure to the firm's stock price relative to her exposure to earnings changes over time. To do so, we analyze the ratio:

$$\frac{\beta^P}{\beta^E} = \frac{\lambda_\ell \ell}{\lambda_s s} \times \frac{1}{p(w) - w p'(w)}.$$

This expression shows that the ratio β^P/β^E largely co-moves with the ratio of long-term over short-term investment, ℓ/s. The longer the time horizon of the firm's investment policies, the more contingent the manager's compensation is on the firm's stock price. Fig. 3.3 depicts a typical pattern of β^P/β^E and ℓ/s as functions of w. The manager's compensation depends the least on the firm's stock price after poor past performance and, in particular, under financial distress. In such instances, long-term investment is of little value so that managerial incentives primarily motivate short-term investment and rely on short-term accounting measures. Additionally, stock price volatility sharply increases under distress, whereas earnings volatility remains

[17] Note that P_t is the per-unit price of equity fully owned by the principal. The manager's value W_t, which could consist of restricted stock units or stock options, is held on an incentive account and is not traded in the market. As such, P_t reflects the market price of traded equity. Since P_t is the investors' firm valuation net of transfers dc to the manager, future dilution by stock vesting or managerial stock option exercise is already accounted for.

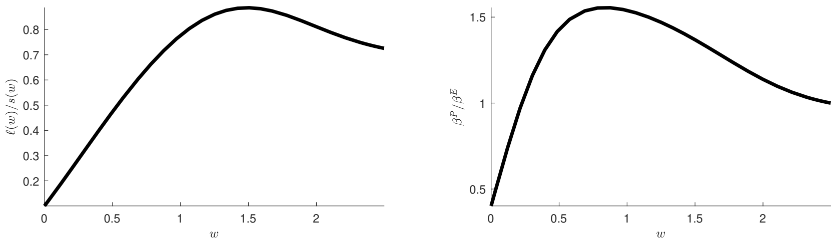

FIGURE 3.3: Numerical illustration of the dynamics of β^P/β^E and ℓ/s. The parameters are $\alpha = 0.25$, $\sigma_K = 0.25$, $\sigma_X = 0.2$, $\rho = 0$, $\mu = 0.025$, $r = 0.046$, $\gamma = 0.048$, $\delta = 0.125$, $\lambda_s = \lambda_\ell = 1$, and $R = 0.2$.

constant over time. This implies that relatively low β^P generates the required long-term incentives for w close to zero.

Remarkably, after sufficiently strong past performance, when $w = \overline{w}$, it holds that $\beta^P/\beta^E = 1$. As a result, the firm compensates the manager equally based on stock price and earnings, regardless of parameter values reflecting firm fundamentals. Thus, our model implies that time-varying agency conflicts—which are temporarily resolved at $\overline{w}$—account for the difference between stock-price-based and earnings-based compensation sensitivities.

3.6 Asymmetric pay in executive compensation

We now turn to analyze the dynamics of incentive provision and show that the optimal contract induces asymmetric pay. We assume throughout the section that the correlation ρ between short- and long-run shocks is zero. For clarity of exposition, we focus on a specification in which the investment cost $\mathcal{C}$ is linear:

$$\mathcal{C}(s, \ell) = \alpha\lambda_s s + \mu\lambda_\ell \ell. \tag{3.22}$$

As a consequence, investment follows a bang-bang solution (i.e., either full or no investment is optimal: $s \in \{0, s_{\max}\}$ and $\ell \in \{0, \ell_{\max}\}$). Equivalently, one could also specify that there is a linear adjustment cost to short-run (respectively long-run) investment up to some threshold $s_{\max}$ (respectively $\ell_{\max}$) and an infinite adjustment cost afterward. The Online Appendix shows that the results derived in this section also apply when the investment cost is convex.

Corner levels of investment are the only relevant cases in a model with binary effort choice (i.e., $s \in \{0, s_{\max}\}$ or $\ell \in \{0, \ell_{\max}\}$), as in He (2009b), or in a model with effort cost functions that are linear in effort levels, as in Biais et al. (2007) and DeMarzo et al. (2012). As a result, considering a linear cost function $\mathcal{C}$ allows us to directly compare our results with those in the models in which moral hazard is solely over the long or the short run and to clarify what outcomes are unique and novel to our model featuring both types of moral hazard. Finally, we assume that full short- and long-run investment is always optimal so that $s(w) = s_{\max}$ and $\ell(w) = \ell_{\max}$ for all w. Thus with the linear investment cost, the dynamics of optimal incentives are not confounded by changes to investment levels.

When the investment cost is linear, incentive compatibility requires

$$\beta^s \geq \lambda_s \text{ and } \beta^\ell \geq \lambda_\ell.$$

The objective of the principal when choosing the manager's exposure to firm performance is to maximize the value derived from the firm, given a promised payment w to the manager. To do so, the principal equivalently minimizes the agent's exposure to shocks, while maintaining incentive compatibility (see Eq. (3.11)). Minimizing risk exposure amounts to minimizing the instantaneous variance of the scaled promised payments:

$$\Sigma(w) = (\beta^s \sigma_X)^2 + (\beta^\ell - w)^2 \sigma_K^2 \text{ subject to } \beta^s \geq \lambda_s \text{ and } \beta^\ell \geq \lambda_\ell.$$

This leads to the following result:

Proposition 14 (Asymmetric pay in executive compensation) *When investment costs are linear and full investment is optimal (i.e., $s = s_{\max}$ and $\ell = \ell_{\max}$), we have that*

i) Incentives are given by $\beta^s = \lambda_s$ and $\beta^\ell = \lambda_\ell + \max\{0, w - \lambda_\ell\}$.
ii) $\beta^\ell(w) > \lambda_\ell$ arises only if $\sigma_X > 0$ and $\sigma_K > 0$.
iii) Assume $\sigma_X > 0$ and $\sigma_K > 0$. If $\gamma - r$, $\ell_{\max}$ or λ_ℓ is sufficiently low, $\overline{w} > \lambda_\ell$ and $\beta^\ell(w) > \lambda_\ell$ for $w \in (\lambda_\ell, \overline{w}]$.

The finding that the incentive compatibility constraint $\beta^s \geq \lambda_s$ in Proposition 14 binds is standard and intuitive. The principal needs to expose the agent to firm performance, but this is costly because this increases the risk of inefficient liquidation. Thus, the principal optimally exposes the agent to as little short-run risk as possible.

The finding that the incentive compatibility constraint $\beta^\ell \geq \lambda_\ell$ does not necessarily bind stems from the fact that the principal optimally wants to expose the manager's continuation payoff to long-run, permanent shocks. Indeed, and as noted above, a positive permanent shock $dZ^K > 0$ has two effects. First, the agent is rewarded for good performance and is promised higher future payments W, which increases the stake w by $\beta^\ell \sigma_K dZ^K$. Second, firm size K grows more than expected, thereby reducing the agent's stake in the firm by $-w\sigma_K dZ^K$. This second effect implies that the agent's stake w is exposed to dilution risk, which can be alleviated using long-run incentives β^ℓ.

When $w > \lambda_\ell$, the principal can fully eliminate dilution risk by setting $\beta^\ell = w$ while maintaining incentive compatibility. Under these circumstances, long-run incentives are effectively costless and the manager is exposed to long-run shocks beyond the level needed to incentivize long-term investment. By contrast, incentive compatibility prevents the principal from eliminating long-run risk when $w < \lambda_\ell$ and $\beta^\ell = \lambda_\ell$. Importantly, when $\lambda_\ell = 0$, there is no agency conflict over the long run and the agent is paid for productivity shocks without any incentive motive, just as in Hoffmann and Pfeil (2010) and DeMarzo et al. (2012).

An important implication of Proposition 14 is that, in our model with dual moral hazard, the agent receives asymmetric performance pay. In particular, the agent is provided minimal long-run incentives $\beta^\ell = \lambda_\ell > w$ for low w and higher-powered long-run incentives $\beta^\ell = w > \lambda_\ell$ after positive past performance, in which case sufficient slack w has been accumulated. In this region, incentives have option-like features and increase after positive performance. Our findings are consistent with evidence on the asymmetry of pay-for-performance in executive compensation (see, for example, Garvey and Milbourn, 2006; Francis et al., 2013). In contrast with the suggested explanations, the asymmetry in pay-for-performance is part of an optimal contract and is not due to managerial entrenchment.[18]

It should be stressed that the model predicts an asymmetry in pay sensitivity to long-run shocks but not to short-run shocks. With the implementation of the optimal contract using stock prices and earnings (see Section 5), the asymmetric sensitivity applies to stock prices but not to earnings.

Remarkably, asymmetric performance pay and strong long-run incentives $\beta^{\ell} \geq \lambda_{\ell}$ can only arise when $\sigma_X > 0$ and $\sigma_K > 0$ and there is a moral hazard over both time horizons, the short and long run. When $\sigma_X = 0$, the principal does not grant the agent a stake w larger than λ_{ℓ} in that payouts $dc > 0$ are made before the agent's stake can grow sufficiently large.

To close this section, note that our findings differ from those in He (2009). First, in his model the incentive condition—which corresponds to $\beta^{\ell} \geq \lambda_{\ell}$ in our paper's notation—is always tight. This occurs because of $w \leq \overline{w} < \lambda_{\ell}$ in that the payout boundary cannot exceed λ_{ℓ}. Intuitively, if it were to happen that $w \geq \lambda_{\ell}$, the firm could profit by paying the agent some small amount dc. This is because the subsequent increase in liquidation risk is negligible (i.e., of order $o(dc^2)$), while the cost of delaying payments—or equivalently the benefits of paying the agent earlier—are of order $o(dc)$. While in our model risks associated with permanent cash flow shocks dZ^K are also negligible whenever $w \simeq \lambda_{\ell}$, the firm remains exposed to substantial transitory cash flow shocks. As a consequence, it can be optimal to accumulate even more slack and to eliminate permanent cash flow risk if possible in that $\overline{w} > \lambda_{\ell}$ and $\beta^{\ell} = w > \lambda_{\ell}$ for $w \in (\lambda_{\ell}, \overline{w}]$.

Second, in He (2009) all risk from permanent cash-flow shocks is eliminated after sufficiently strong past performance only in the extreme case of an equally patient agent and principal. This implies that the firm eventually becomes riskless and the agent works forever. As a result, the first-best outcome can be achieved. In contrast, all risk from permanent cash flow shocks can be eliminated in our model even under the assumption that the agent is more impatient than the principal $\gamma > r$; yet the firm remains exposed to transitory cash flow shocks. As a consequence, only long-run agency conflicts may be temporarily harmless. Indeed, sufficiently adverse cash flow shocks may lower w, drive it below λ_{ℓ}, and even trigger liquidation, implying that first best will never be reached.

3.7 Robustness and extensions

3.7.1 Agent's limited wealth

The Online Appendix solves the model under the assumption that the agent has limited wealth and shows that our findings remain qualitatively unchanged in this alternative setting.

[18] In our model, the agent is essentially paid more for a positive shock than he is punished after a negative shock of the same size. Obviously, this statement is mathematically not exact since the agent's sensitivity to shocks dZ^K is locally symmetric but carries some meaning for shocks over a larger time interval. For a stark intuition, imagine, however, that at time t scaled continuation value equals $w_t = \lambda_{\ell} - \varepsilon$, and let $\Delta = 2\varepsilon > 0$. A shock $Z^K_{t+dt} - Z^K_t = \Delta > 0$ raises w_{t+dt} beyond λ_{ℓ} and therefore increases the agent's value by $W_{t+dt} - W_t > 2\varepsilon\lambda_{\ell}$. In contrast, a shock $Z^K_{t+dt} - Z^K_t = -\Delta < 0$ decreases the agent's value by $2\varepsilon\lambda_{\ell}$.

3.7.2 Private investment cost

In the model, we assume that the principal bears the investment cost $\mathcal{C}$ while the agent can divert funds for her private consumption. Alternatively, we could also assume that the effort (investment) cost $\mathcal{C}$ is private to the manager. In this alternative setting, incentivizing investment s, ℓ requires compensating this private cost to the manager by increasing the growth rate of the agent's scaled continuation value w. Hence, ignoring all other effects, increasing s, ℓ makes w drift up and therefore reduces the likelihood of termination. As a consequence, additional investment/effort cost $\mathcal{C}$ is actually beneficial for the principal when $p'(w) > 0$ or, equivalently, when w is low. As shown in DeMarzo, Livdan, and Tchistyi (2014) and Szydlowski (2015), this beneficial private cost effect may lead to overinvestment. For completeness, we solve our model with private investment cost in the Appendix and demonstrate that short- and long-termism can arise in this model as well.

In the baseline version of our model, the manager does not finance investment expenditures from her own pockets, and agency conflicts arise because of a misallocation or appropriation of funds allocated to investment. We believe that this setup is more realistic for most real-world environments. In addition, it allows us to clearly identify the drivers of short- and long-termism, compared to a model in which the cost of investment is private (see the Online Appendix for details).

3.8 Conclusion

We develop a continuous-time agency model in which the agent controls current earnings via short-term investment and firm growth via long-term investment. In this multitasking model, the principal optimally balances the costs and benefits of incentivizing the manager over the short or the long term. As shown in the paper, this can lead to optimal short- or long-termism (i.e., to short- or long-term investment levels above first best levels), depending on the severity of agency conflicts and firm characteristics. The model predicts that the nature of the risks facing firms is key in determining the corporate horizon. We show, for example, that the correlation between between shocks to earnings and to firm value leads to externalities between investment choices, which are necessary to generate short-termism. We additionally predict that firm performance should be positively related to the corporate horizon. In particular, firms should become more short-termist after bad performance.

Incentives are provided in the optimal contract by making the agent's compensation contingent on firm performance via exposure to the firm's stock price and earnings. Because the firm is subject to long-run, permanent shocks, it is optimal to introduce exposure to long-run volatility that is not needed to incentivize effort in the contract. In our model with multitasking, however, the principal needs to incentivize the manager to exert long-run effort. This generates the distinct prediction that extra pay-for-performance is introduced and the manager's wealth is fully exposed to permanent shocks only when her stake in the firm is large enough. Notably, when her stake is low, the extra pay-for-performance effect is shut down and the incentive compatibility constraint is binding. In other words, positive permanent shocks lead to additional pay-for-performance, and negative permanent shocks eventually eliminate this extra sensitivity to performance implied by the optimal contract. Our model therefore provides a rationale for the asymmetry of pay-for-performance observed in the executive compensation data.

Our model can be extended in several directions. First, we assume a risk-neutral manager who is optimally compensated by lumpy wages. Incorporating risk aversion would imply smooth, but time-varying, wage payments, which is arguably more consistent with empirical observations. Second, in our model, the outcome of long-term investment realizes instantaneously. It would be interesting to study a setup in which the impact of the manager's long-term investment decisions gradually realizes over time, giving rise to more involved incentive structures.

3.9 Appendix A

Without loss in generality, we consider throughout the whole appendix that the depreciation rate of capital δ equals zero. To ensure the problem is well-behaved, we impose the following regularity conditions:

a) Square integrability of the payout process $\{C\}$:

$$\mathbb{E}\left[\left(\int_0^\tau e^{-\gamma s} dC_s\right)^2\right] < \infty.$$

b) The processes $\{s\}$ and $\{\ell\}$ are of bounded variation.

c) The sensitivities $\{\beta^s\}$ and $\{\beta^\ell\}$ are almost surely bounded so that there exists $M > 0$ with $\mathcal{P}(\beta_t^K < M) = 1$ for each $t \geq 0$ and $K \in \{s, \ell\}$. We make this assumption for purely technical reasons and can choose $M < \infty$ arbitrarily large (enough) such that this constraint never binds at the optimum.

3.9.1 Proof of Proposition 1

The first-best investment levels (s^{FB}, ℓ^{FB}) maximize

$$\hat{p}(s, \ell) = \frac{1}{r + \delta - \mu\ell} \left[\alpha s - \mathcal{C}(s, \ell)\right].$$

For the case of quadratic cost, straightforward calculations lead to the desired expressions for s^{FB}, ℓ^{FB} and $p^{FB} \equiv \hat{p}(s^{FB}, \ell^{FB})$, where ℓ^{FB} satisfies the relation $\mu p^{FB} = \mathcal{C}_\ell(s^{FB}, \ell^{FB})$.

3.9.2 Proof of Proposition 8

Auxiliary results

We first show that each investment path $(\{s\}, \{\ell\})$ induces a probability measure under certain conditions. To begin with, fix a probability measure $\mathcal{P}^0$ such that

$$dX_t = \sigma_X K_t d\widetilde{W}_t^X \text{ and } dK_t = \sigma_K K_t d\widetilde{W}_t^K$$

with correlated standard Brownian motions $\{\widetilde{W}^X\}, \{\widetilde{W}^K\}$ under this measure, both progressive with respect to $\mathbb{F}$. The measure $\mathcal{P}^0$ corresponds to perpetual zero investment. Define $\widetilde{\mathbf{W}}_t \equiv \left(\widetilde{W}_t^X, \widetilde{W}_t^K\right)'$ and let the (unconditional) covariance matrix of $\widetilde{\mathbf{W}}_t$ under $\mathcal{P}^0$ be[19]

$$\mathbb{V}^0(\widetilde{\mathbf{W}}_t) = \mathbb{E}^0(\widetilde{\mathbf{W}}_t \widetilde{\mathbf{W}}_t') = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \times t \equiv \mathbf{C}t.$$

In this equation, $\mathbb{V}^0(\cdot)$ denotes the variance operator with respect to the measure $\mathcal{P}^0$. Let us employ a Cholesky decomposition to write $\mathbf{M}^{-1}(\mathbf{M}^{-1})' = \mathbf{C}$ or equivalently $\mathbf{M}'\mathbf{M} = \mathbf{C}^{-1}$

[19]For a matrix-valued random variable $\mathbf{Y} : \Omega \to \mathbb{R}^{n \times k}$, we denote the transposed random variable by $\mathbf{Y}' : \Omega \to \mathbb{R}^{k \times n}$.

for an invertible, deterministic matrix $\mathbf{M}$. Observe that

$$\mathbb{V}^0\big(\mathbf{M}\widetilde{\mathbf{W}}_t\big) = \mathbf{M}\mathbb{E}^0\big(\widetilde{\mathbf{W}}_t\widetilde{\mathbf{W}}_t'\big)\mathbf{M}' = \mathbf{M}\mathbf{C}\mathbf{M}' \cdot t = \mathbf{M}\big(\mathbf{M}'\mathbf{M}\big)^{-1}\mathbf{M}' \cdot t = \mathbf{I} \cdot t,$$

where $\mathbf{I} \in \mathbb{R}^{2\times 2}$ denotes the identity matrix. Because the two components of $\widetilde{\mathbf{W}}_t$ are jointly normal and uncorrelated, they are also independent in that the process $\{\widetilde{\mathbf{W}}^T\} \equiv \{\mathbf{M}\widetilde{\mathbf{W}}\}$ follows a bidimensional standard Brownian motion. We can now apply Girsanov's theorem to $\{\widetilde{\mathbf{W}}^T\}$, where all components, by definition, are mutually independent.

As a first step, we define

$$\Theta_t = \Theta_t(s,\ell) \equiv \Big(\frac{\alpha s_t}{\sigma_X}, \frac{\mu\ell_t}{\sigma_K}\Big)' \text{ and } \widetilde{\Theta}_t = \widetilde{\Theta}_t(s,\ell) \equiv \mathbf{M}\Theta_t(s,\ell).$$

Further, let

$$\Gamma_t' = \Gamma_t'(s,\ell) \equiv \exp\Big(\int_0^t \widetilde{\Theta}_u \cdot d\widetilde{\mathbf{W}}_u^T - \frac{1}{2}\int_0^t ||\widetilde{\Theta}_u||^2 du\Big),$$

where $||\cdot||$ denotes the Euclidean norm in $\mathbb{R}^2$ and

$$\int_0^t \widetilde{\Theta}_u \cdot d\widetilde{\mathbf{W}}_u^T = \int_0^t \sum_{i=1,2} \widetilde{\Theta}_{u,i} d\widetilde{\mathbf{W}}_{u,i}^T = \sum_{i=1,2}\int_0^t \widetilde{\Theta}_{u,i} d\widetilde{\mathbf{W}}_{u,i}^T.$$

Throughout the paper, we will assume that the processes $\{s\}, \{\ell\}$ are such that the Novikov condition is satisfied in that

$$\mathbb{E}^0\Big[\exp\Big(\frac{1}{2}\int_0^\tau ||\widetilde{\Theta}_t||^2(s,\ell)dt\Big)\Big] < \infty.$$

In fact, our regularity conditions imply the Novikov condition. Then, $\{\Gamma'\}$ follows a martingale under $\mathcal{P}^0$ rather than just a local martingale. Due to $\mathbb{E}^0\Gamma_t' = \mathbb{E}^0\Gamma_0' = 1$, the process $\{\Gamma'\}$ is a progressive density process and defines the probability measure $\mathcal{P}^{s,\ell}$ via the Radon-Nikodym derivative

$$\Big(\frac{d\mathcal{P}^{s,\ell}}{d\mathcal{P}^0}\Big)|_{\mathcal{F}_t} = \Gamma_t'.$$

By Girsanov's theorem,

$$\Big\{\mathbf{Z}_t^T = \widetilde{\mathbf{W}}_t^T - \int_0^t \widetilde{\Theta}_u du : t \geq 0\Big\}$$

follows a bidimensional, standard Brownian motion under the measure $\mathcal{P}^{s,\ell}$. The linearity of the (Riemann) integral implies

$$\mathbf{M}\Bigg(\begin{pmatrix} Z_t^X \\ Z_t^K \end{pmatrix}\Bigg) \equiv \mathbf{Z}_t^T = \mathbf{M}\Big(\widetilde{\mathbf{W}}_t - \int_0^t \Theta_u du\Big) = \mathbf{M}\Bigg(\begin{pmatrix} \widetilde{W}_t^X \\ \widetilde{W}_t^K \end{pmatrix} + \begin{pmatrix} \int_0^t \Theta_{u,1} du \\ \int_0^t \Theta_{u,2} du \end{pmatrix}\Bigg).$$

Therefore, for each $t \geq 0$,

$$dZ_t^X \equiv \frac{dX_t - K_t\alpha s_t dt}{K_t\sigma_X} \text{ and } dZ_t^K \equiv \frac{dK_t - K_t\mu\ell_t dt}{K_t\sigma_K}$$

are the increments of a standard Brownian motion under $\mathcal{P}^{s,\ell}$ with instantaneous correlation ρdt. In the following, we say the measure $\mathcal{P}^{s,\ell}$ is induced by the processes $\{s\}, \{\ell\}$. Note that all probability measures of the family $\{\mathcal{P}^{s,\ell}\}_{\{\{s\},\{\ell\}\}}$ are mutually equivalent in that they share the same null sets.

Proof of Proposition 8.1

Consider an incentive compatible contract $\Pi \equiv (\{C\}, \{s\}, \{\ell\}, \tau)$. Further, assume in the following without loss of generality that $\mathbb{F}$ is the filtration generated by $\{X\}, \{K\}$ in that

$\mathcal{F}_t = \sigma(X_s, K_s : 0 \leq s \leq t)$. Then, the agent's continuation utility at time t (under the principal's information) is defined by

$$W_t(\Pi) \equiv \mathbb{E}_t^{s,\ell}\left[\int_t^\tau e^{-\gamma(z-t)}dC_z + \int_t^\tau e^{-\gamma(z-t)}K_z\big(\mathcal{C}(s_z,\ell_z) - \mathcal{C}(\hat{s}_z,\hat{\ell}_z)\big)dz\right],$$

where $\mathbb{E}_t^{s,\ell}(\cdot)$ denotes the conditional expectation given $\mathcal{F}_t$, taken under the probability measure $\mathcal{P}^{s,\ell}$ induced by $\{s\}$ and $\{\ell\}$. Define for $t \leq \tau$:

$$\Gamma_t(\Pi) \equiv \mathbb{E}_t^{s,\ell}\left[W_0(\Pi)\right] = \int_0^t e^{-\gamma z}dC_z + \int_0^t e^{-\gamma z}K_z\big(\mathcal{C}(s_z,\ell_z) - \mathcal{C}(\hat{s}_z,\hat{\ell}_z)\big)dz + e^{-\gamma t}W_t(\Pi). \tag{3.23}$$

By construction, $\{\Gamma_t(\Pi) : 0 \leq t \leq \tau\}$ is a square-integrable martingale under $\mathcal{P}^{s,\ell}$, progressive with respect to $\mathbb{F}$. In the following, we will invoke incentive compatibility (i.e., $s_t = \hat{s}_t, \ell_t = \hat{\ell}_t$) whenever no confusion is likely to arise.

Next, observe that any sigma-algebra is invariant under an injective transformation of its generator. In particular, let $\mathbf{M} \in \mathbb{R}^{2\times 2}$ an invertible, deterministic matrix with $\det(\mathbf{M}) \neq 1$, and note that

$$\mathcal{F}_t = \sigma(X_s, K_s : s \leq t) = \sigma(Z_s^1, Z_s^2 : s \leq t) = \sigma(\mathbf{Z}_s : s \leq t) = \sigma(\mathbf{M}\cdot\mathbf{Z}_s : s \leq t)$$

with $\mathbf{Z}_t \equiv (Z_t^1, Z_t^2)'$. Here,

$$dZ_t^1 \equiv \frac{dX_t - K_t\alpha s_t dt}{K_t\sigma_X} \text{ and } dZ_t^2 \equiv \frac{dK_t - K_t\mu\ell_t dt}{K_t\sigma_K} \tag{3.24}$$

are the increments of a standard Brownian motion under the probability measure $\mathcal{P}^{s,\ell}$. Note that $dZ_t^1 = dZ_t^X$ and $dZ_t^2 = dZ_t^K$ whenever $a_t = \hat{a}_t$ for all $a \in \{s,\ell\}$.

As in the previous section, let the covariance matrix $\mathbb{V}(\mathbf{Z}_t) = \mathbf{C}t$ and employ a Cholesky decomposition $\mathbf{M}'\mathbf{M} = \mathbf{C}^{-1}$. We have already shown that $\{\mathbf{Z}_t^T \equiv \mathbf{M}\mathbf{Z}_t : 0 \leq t \leq \tau\}$ follows a bidimensional, standard Brownian motion under $\mathcal{P}^{s,\ell}$, where both components are mutually independent. By the martingale representation theorem (see, e.g., Shreve, 2004), there exists a bidimensional process $\{\mathbf{b}_t\}_{t\geq 0}$, progressively measurable with respect to $\mathbb{F}$ such that

$$d\Gamma_t(\Pi) = e^{-\gamma t}\mathbf{b}_t'\cdot d\mathbf{Z}_t^T = e^{-\gamma t}\mathbf{b}_t'\cdot\mathbf{M}\mathbf{M}^{-1}\cdot d\mathbf{Z}_t^T = e^{-\gamma t}K_t\big(\beta_t^s\sigma_X dZ_t^1 + \beta_t^\ell\sigma_K dZ_t^2\big),$$

where we exploit the linearity of the Itô integral—i.e., $d\big(\mathbf{M}\mathbf{Z}_t^T\big) = \mathbf{M}d\mathbf{Z}_t^T$—and set $(\beta_t^s\sigma_X, \beta_t^\ell\sigma_K) \equiv \mathbf{b}_t'\mathbf{M}/K_t$ for all t. Combining with Eq. (3.23), one can verify that

$$\begin{aligned} d\Gamma_t(\Pi) &= e^{-\gamma t}K_t\big(\beta_t^s\sigma_X dZ_t^1 + \beta_t^\ell\sigma_K dZ_t^2\big) \\ &= e^{-\gamma t} - \gamma e^{-\gamma t}W_t(\Pi)dt + e^{-\gamma t}dW_t(\Pi), \end{aligned}$$

and thus Eq. (3.7) holds after rearranging. Indeed, since the right-hand side of Eq. (3.7) satisfies a Lipschitz condition under the usual regularity conditions (i.e., square integrability of $\{C\}$ and $\{s\},\{\ell\}$ of bounded variation), $\{W\}$ is the unique strong solution to the stochastic differential Eq. (3.7).

Next, we provide necessary and sufficient conditions for the contract Π to be incentive compatible. For this purpose, let the recommended investment processes $\{s\}$ and $\{\ell\}$ and the expected payoff of the agent at time t be W_t when following the recommended strategy from time t onwards. Further, let $\{\hat{s}\}$ and $\{\hat{\ell}\}$ represent the actual investment processes, which may in principle differ from $\{s\}$ and $\{\ell\}$. We have

$$W_t \equiv \mathbb{E}_t^{s,\ell}\left[\int_t^\tau e^{-\gamma(z-t)}dC_z\right].$$

Recall that $\mathbb{E}_t^{s,\ell}$ denotes the expectation, conditional on the filtration $\mathcal{F}_t$, taken under the probability measure $\mathcal{P}^{s,\ell}$. As shown above, the process $\{W\}$ solves the stochastic differential equation:

$$dW_t = \gamma W_t dt + \beta_t^s \big(dX_t - K_t \alpha s_t dt\big) + \beta_t^\ell \big(dK_t - K_t \mu \ell_t dt\big) - dC_t.$$

We can rewrite this stochastic differential equation as

$$\begin{aligned} dW_t + dC_t &= \gamma W_t dt \\ &\quad + K_t \beta_t^s \big[\alpha(\hat{s}_t - s_t)dt + \sigma_X dZ_t^X\big] + K_t \beta_t^\ell \big[\mu(\hat{\ell}_t - \ell_t)dt + \sigma_K dZ_t^K\big] \end{aligned}$$

with

$$dZ_t^X \equiv \frac{dX_t - K_t \alpha \hat{s}_t dt}{K_t \sigma_X} \text{ and } dZ_t^K \equiv \frac{dK_t - K_t \mu \hat{\ell}_t dt}{K_t \sigma_K}.$$

Girsanov's theorem implies now that $dZ_t^X \equiv \frac{dX_t - K_t \alpha \hat{s}_t dt}{K_t \sigma_X}$ and $dZ_t^K \equiv \frac{dK_t - K_t \mu \hat{\ell}_t dt}{K_t \sigma_K}$ are the increments of a standard Brownian motion under the measure $\mathcal{P}^{\hat{s},\hat{\ell}}$ induced by $(\{\hat{s}\}, \{\hat{\ell}\})$.

Next, define the auxiliary gain process

$$g_t = g_t\big(\{\hat{s}\}, \{\hat{\ell}\}\big) \equiv \int_0^t e^{-\gamma z} dC_z - \int_0^t e^{-\gamma z} K_z (\mathcal{C}(\hat{s}_z, \hat{\ell}_z) - \mathcal{C}(s_z, \ell_z)) dz + e^{-\gamma t} W_t,$$

and recall that $W_\tau = 0$. Now, note that the agent's actual expected payoff under the strategy $(\{\hat{s}\}, \{\hat{\ell}\})$ reads

$$\begin{aligned} W_0' &\equiv \max_{\{\hat{s}\},\{\hat{\ell}\}} \mathbb{E}^{\hat{s},\hat{\ell}} \left[\int_0^\tau e^{-\gamma z} dC_z - \int_0^\tau e^{-\gamma z} K_z (\mathcal{C}(\hat{s}_z, \hat{\ell}_z) - \mathcal{C}(s_z, \ell_z)) dz \right] \\ &= \max_{\{\hat{s}\},\{\hat{\ell}\}} \mathbb{E}^{\hat{s},\hat{\ell}} \left[g_\tau\big(\{\hat{s}\}, \{\hat{\ell}\}\big) \right]. \end{aligned}$$

We obtain

$$\begin{aligned} e^{\gamma t} dg_t &= K_t \Big[\mathcal{C}(s_t, \ell_t) - \mathcal{C}(\hat{s}_t, \hat{\ell}_t) \Big] dt \\ &\quad + K_t \Big[\alpha \beta_t^s (\hat{s}_t - s_t) + \mu \beta_t^\ell (\hat{\ell}_t - \ell_t) \Big] dt + K_t \Big[\beta_t^s \sigma_X dZ_t^X + \beta_t^\ell \sigma_K dZ_t^K \Big] \\ &\equiv \mu_t^g dt + K_t \Big[\beta_t^s \sigma_X dZ_t^X + \beta_t^\ell \sigma_K dZ_t^K \Big]. \end{aligned}$$

It is now easy to see that when choosing $\hat{s}_t = s_t, \hat{\ell}_t = \ell_t$, the agent can always ensure that $\mu_t^g = 0$, in which case $\{g_z\}_{z \geq 0}$ follows a martingale under $\mathcal{P}^{s,\ell}$. Hence,

$$W_0' = \max_{\{\hat{s}\},\{\hat{\ell}\}} \mathbb{E}^{\hat{s},\hat{\ell}} \Big[g_\tau\big(\{\hat{s}\}, \{\hat{\ell}\}\big) \Big] \geq \mathbb{E}^{s,\ell} \Big[g_\tau\big(\{s\}, \{\ell\}\big) \Big] = W_0.$$

The inequality is strict if and only if there exist processes $\{\hat{s}\}, \{\hat{\ell}\}$ and a stopping time τ' with $\mathcal{P}^{\hat{s},\hat{\ell}}(\tau' < \tau) > 0$ such that $\mu_{\tau'}^g > 0$. This arises because then there also exists a set $\mathcal{A} \subseteq [0, \tau) \times \Omega$ with

$$\mu_t^G(\omega) > 0 \text{ for all } (t, \omega) \in \mathcal{A} \text{ and } \mathcal{L} \otimes \mathcal{P}^{\hat{s},\hat{\ell}}(\mathcal{A}) > 0,$$

where $\mathcal{L}$ is the Lebesgue measure on on the Lebesgue sigma-algebra in $\mathbb{R}$. Because $\mathcal{P}^{\hat{s},\hat{\ell}}(\tau < \infty)$ for all admissible $\{\hat{s}\}, \{\hat{\ell}\}$ it follows that $e^{-\gamma t} \mu_t^G(\omega) > 0$ for all $(t, \omega) \in \mathcal{A}$. Hence,

$$W_0' \geq \int_{\mathcal{A}} e^{-\gamma z} \mu_z^g(\omega) d\big(\mathcal{L}(z) \otimes \mathcal{P}^{s,\ell}(\omega)\big) + W_0 > W_0.$$

In case $W_0' > W_0$, either $\hat{s}_z(\omega) \neq s_z(\omega)$ or $\hat{\ell}_z(\omega) \neq \ell_z(\omega)$ on the set $\mathcal{A}$, which has positive measure so that Π is not incentive compatible.

Hence, for Π to be incentive compatible, it must, for all $t \geq 0$ (almost surely), hold that

$$\max_{\hat{s}_t,\hat{\ell}_t} \left\{ \alpha\beta_t^s(\hat{s}_t - s_t) + \mu\beta_t^\ell(\hat{\ell}_t - \ell_t) + \left[\mathcal{C}(s_t,\ell_t) - \mathcal{C}(\hat{s}_t,\hat{\ell}_t)\right] \right\} = 0,$$

or equivalently

$$(s_t,\ell_t) \in \arg\max_{\hat{s}_t,\hat{\ell}_t} \left\{ \alpha\beta_t^s(\hat{s}_t - s_t) + \mu\beta_t^\ell(\hat{\ell}_t - \ell_t) + \left[\mathcal{C}(s_t,\ell_t) - \mathcal{C}(\hat{s}_t,\hat{\ell}_t)\right] \right\},$$

for given β_t^s, β_t^ℓ. After going through the maximization, we obtain that this is satisfied if $\mathcal{C}_s(s_t,\ell_t) = \beta_t^s\alpha$ and $\mathcal{C}_\ell(s_t,\ell_t) = \beta_t^\ell\mu$, in case $(s_t,\ell_t) \in (0,s_{\max}) \times (0,\ell_{\max})$. If $a_t \in \{s_t,\ell_t\}$ is not interior in that $a_t = a_{\max}$ for $a \in \{s,\ell\}$, then $a_t = \hat{a}_t$ solves the above maximization problem if and only if $\beta_t^s\alpha \geq \mathcal{C}_s(s_t,\ell_t)$ if $a_t = s_t$, or $\beta_t^\ell\mu \geq \mathcal{C}_\ell(s_t,\ell_t)$ if $a_t = \ell_t$. It evidently suffices here to consider first-order optimality so that the result follows.

Proof of Proposition 8.2

In this section, we proceed as follows. First, we represent $P(W,K)$ as a twice continuously differentiable solution of a HJB equation and then show that there exists a function $p \in C^2$ such that $P(W,K) = Kp(w)$ and $p(w)$ solves the scaled HJB Eq. (3.11). Second, we verify that $P(W,K)$ or equivalently $p(w)$ with corresponding payout threshold $\overline{w}$ and $w_0 = w^*$ characterizes indeed the optimal contract Π^*. Since we focus on incentive compatible contracts, we will work in the following—unless otherwise mentioned—with the measure $\mathcal{P}^{s^*,\ell^*}$ induced by optimal investment $(\{s^*\},\{\ell^*\})$. For convenience, we will denote this measure by $\mathcal{P}$ if no confusion is likely to arise. We follow an analogous convention concerning the expectation operator where we just write $\mathbb{E}_t(\cdot)$ instead of $\mathbb{E}^{s^*,\ell^*}(\cdot|\mathcal{F}_t)$.

Scaling of the value function

Given the optimal control and stopping problem (3.6), suppose that the principal's value function $P(W,K)$ satisfies the HJB equation

$$\begin{aligned} rP = \max_{s,\ell,\beta^s,\beta^\ell} \Bigg\{ & \alpha sK - K\mathcal{C}(s,\ell) + P_W\gamma W + P_K\mu\ell K + \frac{1}{2}\Bigg(P_{WW}\Big[\left(\beta^s\sigma_X K\right)^2 \\ & + \left(\beta^\ell\sigma_K K\right)^2 + 2\rho\sigma_X\sigma_K K^2\beta^s\beta^\ell \Big] + P_{KK}\left(\sigma_K K\right)^2 + 2P_{WK}\Big[\left(\sigma_K K\right)^2\beta^\ell + \rho\sigma_X\sigma_K K^2\beta^s \Big] \Bigg) \Bigg\} \end{aligned}$$

in some region $\mathcal{S} \subset \mathbb{R}^2$, subject to the boundary conditions

$$P(0,K) = RK, P(W,0) = 0, P_W(\overline{W},K) = -1, P_{WW}(\overline{W},K) = 0.$$

Here, $\overline{W} \equiv \overline{W}(K) = \overline{w}K$ parametrizes the boundary of $\mathcal{S}$ on which $W,K > 0$. Taking the guess $P(W,K) = p(W/K)K$ for some function $p \in C^2$, we obtain

$$P_W = p'(w), P_K = p(w) - wp'(w), P_{WK} = -w/Kp''(w), P_{WW} = p''(w)/K, p_{KK} = w^2/Kp''(w),$$

which implies the HJB Eq. (3.11) and its boundary conditions.

In the following, we will assume that Eq. (3.11) admits an unique, twice continuously differentiable solution $p(\cdot)$ on $[0,\overline{w}]$. A formal existence proof is beyond the scope of the paper and is therefore omitted.[20]

[20] Indeed, the possible discontinuities of the functions $s(\cdot),\ell(\cdot)$ cause technical complications. If $s_{\max}, \ell_{\max}$ are sufficiently large, this problem is not present anymore. Then, the existence and uniqueness of the solution follow from the Picard-Lindelöf theorem since the required Lipschitz condition is evidently satisfied.

We first rewrite the principal's problem (3.6) in a convenient manner. Let

$$\Psi_t = \left(\rho\sigma_K t, \sigma_K t\right)' \text{ and } \widetilde{\Psi}_t = \mathbf{M}\Psi_t,$$

where $\mathbf{M}'\mathbf{M} = \mathbf{C}^{-1}$ and $\mathbf{C}t$ is the covariance matrix of (Z_t^X, Z_t^K). Next, define the equivalent, auxiliary probability measure $\widetilde{\mathcal{P}}$ according to the Radon-Nikodym derivative

$$\left(\frac{d\widetilde{\mathcal{P}}}{d\mathcal{P}}\right)|_{\mathcal{F}_t} \equiv \exp\left\{\int_0^t \widetilde{\Psi}_u du - \frac{1}{2}\int_0^t ||\widetilde{\Psi}_u||^2 du\right\}. \tag{3.25}$$

By arguments similar to those in Appendix 3.9.2, Girsanov's theorem implies that

$$\widetilde{Z}_t^X = Z_t^X - \rho\sigma_K t \text{ and } \widetilde{Z}_t^K = Z_t^K - \sigma_K t$$

are both standard Brownian motions with correlation ρt under $\widetilde{\mathcal{P}}$. An application of Itô's lemma consequently yields that the scaled continuation value $\{w\}$ evolves according to

$$dw_t + dc_t = (\gamma - \mu\ell_t)w_t dt + \beta_t^s \sigma_X d\widetilde{Z}_t^X + (\beta_t^\ell - w_t)\sigma_K d\widetilde{Z}_t^K$$

under $\widetilde{\mathcal{P}}$. Finally, for $\psi_t \equiv rt - \mu\int_0^t \ell_z dz$ we can rewrite the principal's problem (3.6) as

$$\max_{\{c\},\{s\},\{\ell\},w^*} \widetilde{\mathbb{E}}\left[\int_0^\tau e^{-\psi_t}\left(\alpha s_t - C(s_t,\ell_t)\right)dt - \int_0^\tau e^{-\psi_t}dc_t + e^{-\psi_\tau}R\middle| w_0 = w^*\right],$$

where the expectation $\widetilde{\mathbb{E}}[\cdot]$ is taken under the equivalent, auxiliary measure $\widetilde{\mathcal{P}}$. Here, $dc_t \equiv dC_t/K_t = \max\{w_t - \overline{w}, 0\}$. The stated integral expression is implied by following lemma.

Lemma 16 *Suppose $\{w\}$ is the unique, strong solution to the stochastic differential equation*

$$dw_t = \delta_t dt + \Delta_t w_t dt - dc_t + (\beta_t^\ell - w_t)\sigma_K dZ_t^K + \beta_t^s \sigma_X dZ_t^X$$

for $t \leq \tau$, standard Brownian motions $\{Z^X\}, \{Z^K\}$ with correlation ρ and progressive processes $\{\delta\}$, $\{\Delta\}$, $\{\beta^\ell\}$, $\{\beta^s\}$ of bounded variation.[21] *Assume that $dw_t = 0$ for $t > \tau$ where $\tau = \min\{t \geq 0 : w_t = 0\}$. Furthermore, $dc_t = \max\{w_t - \overline{w}, 0\}$ with threshold $\overline{w} > 0$. Let now $g : [0,\overline{w}] \to \mathbb{R}$ of bounded variation. Then the twice continuously differentiable function $f : [0,\overline{w}] \to \mathbb{R}$ (i.e., $f \in C^2$) solves the differential equation*

$$r_t f(w_t) = g(w_t) + f'(w_t)[\delta_t + \Delta_t w_t] + f''(w_t)\left[\sigma_K^2(\beta_t^\ell - w_t)^2 + (\beta_t^s\sigma_X)^2 + 2\rho\sigma_X\sigma_K\beta_t^s(\beta_t^\ell - w_t)\right] \tag{3.26}$$

with boundary conditions $f(0) = R$, $f'(\overline{w}) = -1$ if and only if

$$f(w) = \mathbb{E}\left[\int_0^\tau e^{-\int_0^t r_u du} g(w_t)dt - \int_0^\tau e^{-\int_0^t r_u du}dc_t + e^{-\int_0^\tau r_u du}R\middle| w_0 = w\right]$$

for a progressive discount rate process $\{r\}$ of bounded variation.

Suppose $f(\cdot)$ solves Eq. (3.26). Define

$$h_t \equiv \int_0^t e^{-\int_0^z r_u du} g(w_z)dz - \int_0^t e^{-\int_0^z r_u du}dc_z + e^{-\int_0^t r_u du}f(w_t).$$

[21] We call a process $\{Y\}$ "of bounded variation" if it can be written as the difference of two almost surely increasing processes. Similarly, a function $F \in \mathbb{R}^{[a,b]}$ is called "of bounded variation" if it can be written as the difference of two increasing functions on the interval $[a,b]$.

Applying Itô's lemma, we obtain

$$\begin{aligned}
e^{\int_0^t r_u dt} dh_t = &\Big\{ g(w_t) - r_t f(w_t) + \frac{f''(w_t)}{2} \left[\sigma_K^2 (\beta_t^\ell - w_t)^2 + (\beta_t^s \sigma_X)^2 + 2\rho\sigma_X\sigma_K \beta_t^s (\beta_t^\ell - w_t) \right] \\
&+ f'(w_t)(\delta_t + \Delta_t w_t) \Big\} dt \\
&- \left[(1 + f'(w_t)) dc_t \right] + f'(w_t) \left[dZ_t^X \beta_t^s \sigma_X + dZ_t^K (\beta_t^\ell - w_t)\sigma_K \right].
\end{aligned}$$

The first term in curly brackets equals zero because $f(\cdot)$ solves Eq. (3.26). Since $f'(\overline{w}) = -1$ and $dc_t = 0$ for all $w_t \leq \overline{w}$, the second term in square brackets equals also zero, and therefore $\{h\}$ follows a martingale up to time τ. As a result, we have

$$f(w_0) = f(w) = h_0 = \mathbb{E}\left[h_\tau\right] = \mathbb{E}\left[\int_0^\tau e^{-\int_0^t r_s ds} g(w_t) dt - \int_0^\tau e^{-\int_0^t r_s ds} dc_t + e^{-\int_0^\tau r_s ds} R \,\middle|\, w_0 = w \right].$$

The result follows.

Verification

Next, we verify the optimality of the contract Π^* among all contracts Π satisfying incentive compatibility. To do so, we show that the principal obtains under any contract $\Pi \in \mathbb{IC}$ at most (scaled) payoff $\tilde{G}(\Pi)/K \leq p(w^*)$ with equality if and only if $\Pi = \Pi^*$. Here, $p(\cdot)$ solves the HJB Eq. (3.11) with corresponding payout threshold $\overline{w}$ and $w_0 = w^*$.

Consider any incentive-compatible contract $\Pi \equiv (\{C\}, \{s\}, \{\ell\}, \tau)$. For any $t \leq \tau$, define its auxiliary gain process G as

$$G_t(\Pi) = \int_0^t e^{-ru} \big(dX_u - \mathcal{C}(s_u, \ell_u) du \big) - \int_0^t e^{-ru} dC_u + e^{-rt} P(W_t, K_t),$$

where the agent's continuation payoff evolves according to Eq. (3.7). Recall that $w_t = \frac{W_t}{K_t}$ and $P(W_t, K_t) = K_t p(w_t)$. Itô's lemma implies that for $t \leq \tau$

$$\begin{aligned}
&e^{rt} \frac{dG_t(\Pi_1)}{K_t} \\
&= \Bigg[- (r - \mu\ell_t) p(w_t) + \alpha s_t - \mathcal{C}(s_t, \ell_t) + p'(w_t) w_t (\gamma - \mu\ell_t) \\
&+ \frac{p''(w_t)}{2} \left[(\beta_t^s \sigma_X)^2 + \sigma_K^2 (\beta_t^\ell - w_t)^2 + 2\rho\sigma_X\sigma_K \beta_t^s (\beta_t^\ell - w_t) \right] \Bigg] dt - (1 + p'(w_t)) dc_t \\
&+ \sigma_K \big(p(w_t) + p'(w_t)(\beta_t^\ell - w_t) \big) dZ_t^K + \sigma_X \big(1 + \beta_t^s p'(w_t) \big) dZ_t^X.
\end{aligned}$$

Under the optimal investment and incentives, the first term in square brackets stays at zero always. Other investment and incentive policies will make this term negative (owing to the concavity of p). The second term is non-positive since $p'(w_t) \geq -1$ but equal to zero under the optimal contract. Therefore, for the auxiliary gain process, we have

$$dG_t(\Pi) = \mu_G(t) dt + e^{-rt} K_t \left[\sigma_K \big(p(w_t) + p'(w_t)(\beta_t^\ell - w_t) \big) dZ_t^K + \sigma_X \big(1 + \beta_t^s p'(w_t) \big) dZ_t^X \right],$$

where $\mu_G(t) \leq 0$. Due to our assumption of bounded sensitivities $\{\beta^s\}, \{\beta^\ell\}$, it follows that

$$\mathbb{E}\left(\int_0^t e^{-ru} \big(p(w_u) + p'(w_u)(\beta_u^\ell - w_u) \big) dZ_u^K \right) = \mathbb{E}\left(\int_0^t e^{-ru} \big(1 + \beta_u^s p'(w_u) \big) dZ_u^X \right) = 0,$$

which implies that $\{G_t\}_{t\geq 0}$ follows a supermartingale. Furthermore, under Π, investors' expected payoff is

$$\tilde{G}(\Pi) \equiv \mathbb{E}\left[\int_0^\tau e^{-ru}\big(dX_u - \mathcal{C}(s_u,\ell_u)du\big) - \int_0^\tau e^{-ru}dC_u + e^{-r\tau}RK_\tau\right].$$

As a result, we have

$$\begin{aligned}
&\tilde{G}(\Pi) = \mathbb{E}\left[G_\tau(\Pi)\right] \\
&= \mathbb{E}\left[G_{\tau\wedge t}(\Pi) + \mathbf{1}_{\{t\leq\tau\}}\left(\int_t^\tau e^{-rs}\big(dX_s - dC_s - \mathcal{C}(s_s,\ell_s)ds\big) + e^{-r\tau}RK_\tau - e^{-rt}P(W_t,K_t)\right)\right] \\
&= \mathbb{E}\left[G_{\tau\wedge t}(\Pi)\right] \\
&+ e^{-rt}\mathbb{E}\left[\mathbf{1}_{\{t\leq\tau\}}\mathbb{E}_t\left(\int_t^\tau e^{-r(s-t)}\big(dX_s - dC_s - \mathcal{C}(s_s,\ell_s)ds\big) + e^{-r(\tau-t)}RK_\tau - P(W_t,K_t)\right)\right] \\
&\leq G_0 + e^{-rt}\mathbb{E}\left[P^{FB}(K_t) - W_t - P(W_t,K_t)\right] \\
&\leq G_0 + e^{-rt}\big(p^{FB} - R\big)\mathbb{E}\left[K_t\right],
\end{aligned}$$

where $p^{FB} \equiv \frac{P^{FB}(K_t)}{K_t}$ is the (scaled) first-best value. The inequalities follow from the supermartingale property of G_t, the fact that the value of the firm with agency is below first best, and the fact that $p^{FB} - w - p(w) \leq p^{FB} - R$. Since $\mu\ell_{\max} < r$, it follows that $\lim_{t\to\infty} e^{-rt}\mathbb{E}\left[K_t\right] = 0$. Therefore, letting $t \to \infty$ yields $\tilde{G}(\Pi) \leq G_0 \equiv P(W_0,K_0) = p(w_0)K_0$ for all incentive compatible contracts. For the optimal contract Π^*, the investors' payoff $\tilde{G}(\Pi^*)$ achieves $P(W_0,K_0) = p(w_0)K_0$ since the above weak inequality holds in equality when $t \to \infty$. This completes the argument.

Proof of Proposition 8.3

Auxiliary results

In this section, we prove the following auxiliary lemma, which is key for establishing the concavity of the value function.

Lemma 17 *Let $p(\cdot)$ the unique, twice continuously differentiable solution to the HJB Eq. (3.11) on the interval $[0,\overline{w}]$ subject to the boundary conditions $p(0) = R$, $p'(\overline{w}) = -1$, and $p''(\overline{w}) = 0$. Further, assume the processes $\{s\},\{\ell\}$ are of bounded variation. Then it follows for any $w_1 \in (0,\overline{w}]$ with $p''(w_1) = 0$, that $p'(w_1) < 0$ and the policy functions $s(\cdot),\ell(\cdot)$ are continuous in a neighborhood of w_1.*

We start with an important observation. Because the processes $\{s\},\{\ell\}$ are, by hypothesis, of bounded variation, they can be written as the difference of two almost surely increasing processes such that $a_t = a_t^1 - a_t^2$ for all $t \geq 0$, $a \in \{s,\ell\}$ and $a^j_{(\cdot)}$ increases almost surely. By Froda's theorem,[22] each of the processes $\{\hat{a}\}$ has no essential discontinuity and at most countably many jump-discontinuities with probability one. Since $\{w\}$ follows a Brownian semimartingale, this implies that any point of discontinuity of $a(\cdot)$ can neither be an essential discontinuity nor can the set of discontinuity points of $a(\cdot)$ be dense in $[0,\overline{w}]$ for all $a \in \{s,\ell\}$.

We first prove that $p'(w_1) < 0$. Let us suppose to the contrary $p'(w_1) \geq 0$ hence $w_1 < \overline{w}$. Note that for any $\delta > 0$ exists $z \in (w_1-\delta, w_1+\delta)$ such that $s(\cdot),\ell(\cdot)$ are continuous in a neighborhood of z because discontinuity points do not form a dense set. Since $p'(\cdot),p''(\cdot)$ are continuous, for any $\varepsilon > 0$, we can choose $\delta > 0$ and $z \in (w_1-\delta, w_1+\delta)$ such that

[22]Froda's theorem states that each real-valued, monotone function has at most countably many points of discontinuity. It is clear that such a function cannot have an essential discontinuity (i.e., a point of oscillation).

$\min\{p'(z), p''(z)\} > -\varepsilon$. The HJB Eq. (3.11) and the fact that $\ell(z) = \ell^{FB}$ is not necessarily optimal imply

$$(r - \mu\ell^{FB})p(z) \geq \max_{s\in[0,s_{\max}]} \Big\{\alpha s + p'(z)(\gamma - \mu\ell^{FB})z - \mathcal{C}(s, \ell^{FB}) + p''(z)\Sigma(z)\Big\}$$
$$\geq \max_{s\in[0,s_{\max}]} \Big\{\alpha s - \varepsilon\big(\gamma - \mu\ell^{FB})z - \mathcal{C}(s, \ell^{FB}) + \Sigma(z)\big]\Big\}.$$

Sending $\varepsilon, \delta \to 0$ such that $s = s(z) = s_{\max} \geq s^{FB}$ and in particular for $z = w_1$

$$\alpha s - \mathcal{C}(s, \ell^{FB} \geq \alpha s^{FB} - \mathcal{C}(s^{FB}, \ell^{FB}) \geq (r - \mu\ell^{FB})p^{FB}.$$

Hence, there exists $z \in [0, \overline{w}]$ such that $p(z) \geq p^{FB}$, a contradiction.

Next, let us prove that $\ell(\cdot)$ must be continuous in a neighborhood of w_1, and assume to the contrary that there is no neighborhood of w_1 on which $\ell(\cdot)$ is continuous. Since the set of discontinuities of $\ell(\cdot)$ must be discrete (not dense), it is immediate that

$$\ell_- \equiv \lim_{w\uparrow w_1} \ell(w) \neq \lim_{w\downarrow w_1} \ell(w) \equiv \ell_+,$$

i.e., $\ell(\cdot)$ has a jump discontinuity at w_1 itself. Without loss of generality, we will assume that $\ell_- < \ell_+$ and $w_1 < \overline{w}$.[23]

Note that for all $\varepsilon > 0$, there exists $\delta > 0$ such that for all $z \in (w_1, w_1 + \delta)$, it holds that $|\ell(z) - \ell_+| < \varepsilon$. The optimality of $\ell(z)$ requires that $\frac{\partial p(z)}{\partial \ell}|_{\ell=\ell(z)} \geq 0$ with equality if $\ell(z)$ is interior. Due to the continuity of $p''(\cdot)$, the limit $\varepsilon \to 0$ yields $\Gamma_\ell(w_1) \geq 0$ for

$$\Gamma_\ell(w) = p(w) - p'(w)w - \mathcal{C}_\ell(s, \ell_+) \text{ with } \mathcal{C}_\ell(s, \ell_+) = \frac{\partial \mathcal{C}(s, \ell)}{\partial \ell}|_{\ell=\ell_+}.$$

In addition, for all $\varepsilon > 0$, it must be that there exists $\delta > 0$ such that for all $x \in (w_1 - \delta, w_1)$, it holds that $|\ell(x) - \ell_-| < \varepsilon$. Hence, for $\varepsilon > 0$ sufficiently small, $\ell(x) < \ell_{\max}$ and therefore $\frac{\partial p(x)}{\partial \ell}|_{\ell=\ell(x)} = 0$, which implies together with the continuity of $p''(\cdot)$ that $\widehat{\Gamma}_\ell(w_1) = 0$ for

$$\widehat{\Gamma}_\ell(w) = p(w) - p'(w)w - \mathcal{C}_\ell(s, \ell_-).$$

Next, observe that

$$0 \leq \Gamma_\ell(w_1) - \widehat{\Gamma}_\ell(w_1) = -\lambda_\ell(\ell_+ - \ell_-).$$

Then, it follows that $\ell_- \geq \ell_+$, a contradiction.

Finally, assume that there is no neighborhood of w_1 on which $s(\cdot)$ is continuous. Since the set of discontinuity points of $s(\cdot)$ is discrete, this is equivalent to $s_- \equiv \lim_{w\uparrow w_1} s(w) \neq \lim_{w\downarrow w_1} s(w) \equiv s_+$. Without loss of generality, suppose $s_+ > s_-$. Then, for all $\varepsilon > 0$, there exists $\delta > 0$ such that for all $z \in (w_1, w_1 + \delta)$, it holds that $|s(z) - s_+| < \varepsilon$. Optimality requires $\frac{\partial p(z)}{\partial s}|_{s=s(z)} \geq 0$. Taking the limit $\varepsilon \to 0$, we obtain $\Gamma_s(w_1) \geq 0$ for $\Gamma_s(w) = \alpha s(w) - \mathcal{C}_s(s_+, \ell)$. Similarly, $\widehat{\Gamma}_s(w_1) = 0$ for $\widehat{\Gamma}_s(w) = s(w) + p'(w)\mathcal{C}_s(s_-, \ell)$. Hence,

$$0 \leq \Gamma_s(w_1) - \widehat{\Gamma}_s(w_1) = -\lambda_s(s_+ - s_-).$$

Then, it follows that $s_- \geq s_+$, a contradiction.

Concavity of the value function

Since $p''(\cdot)$ is continuous on $[0, \overline{w}]$ and $\{s\}, \{\ell\}$ are of bounded variation, it follows that the mappings $s(\cdot), \ell(\cdot)$ are continuous on $[0, \overline{w}]$ up to a discrete set with (Lebesgue) measure zero. On the set, where $s(\cdot), \ell(\cdot)$ are continuous, the envelope theorem implies now that

[23] Since $p(\cdot)$ is extended linearly to the right of $\overline{w}$, discontinuity to the right of $\overline{w}$ is not an issue.

$p'''(\cdot)$ exists and is given by

$$p'''(w) = \frac{(r-\gamma)p'(w) - p''(w)\left(w(\gamma-\mu\ell) - \sigma_K^2(\beta^\ell - w) - \rho\sigma_X\sigma_K\beta^s\right)}{\frac{1}{2}\left((\beta^s\sigma_X)^2 + \sigma_K^2(\beta^\ell - w)^2 + 2\rho\sigma_X\sigma_K\beta^s(\beta^\ell - w)\right)}.$$

We have to show that $p''(w) < 0$ for all $0 \leq w < \overline{w}$.

By Lemma 17 we know that $s(\cdot), \ell(\cdot)$ are continuous in a neighborhood of $\overline{w}$. Then, we observe that $p'''(\overline{w}) \propto \gamma - r > 0$ due to $\beta^s \geq \lambda_s s > 0$, and thus $p'''(\cdot) > 0$ in a neighborhood of $\overline{w}$. Hence, $p''(w) < 0$ on an interval $(\overline{w} - \varepsilon, \overline{w})$ with appropriate $\varepsilon > 0$.

Next, suppose there exists $w_0 \in [0, \overline{w}]$ with $p''(w_0) > 0$, and define $w_1 \equiv \sup\{w \in [0, \overline{w}) : p''(w) \geq 0\}$. By the previous step and continuity, it follows that $p''(w_1) = 0$ and $w_1 < \overline{w}$. We obtain now from Lemma 17 that $s(\cdot), \ell(\cdot)$ are continuous in a neighborhood of w_1 and $p'(w_1) < 0$. However, this implies $p'''(w_1) > 0$ and therefore $p'''(\cdot) > 0$ in a neighborhood of w_1. Thus, there exists $w' > w_1$ with $p''(w') > 0$, a contradiction to the definition of w_1. This completes the proof.

Last, let us state the following corollary, which proves useful in some instances:

Corollary 5 *If $\gamma - r$ and σ_K^2 are sufficiently small, then $p'''(w) > 0$ for any $w \in [0, \overline{w})$.*

It is immediate from the above given the expression of $p'''(w)$.

3.9.3 Proofs of Propositions 9 and 10

The expressions for $s = s(w), \ell = \ell(w)$ follow directly from the maximization of $p(w)$ over $s \in [0, s_{\max}]$ and $\ell \in [0, \ell_{\max}]$ for a given w as indicated by the HJB Eq. (3.11). Interior levels $s(w), \ell(w)$ must solve the respective first-order conditions of maximization (i.e., $\frac{\partial p(w)}{\partial s}|_{s=s(w)} = 0$ and $\frac{\partial p(w)}{\partial \ell}|_{\ell=\ell(w)} = 0$). After rearranging the first-order conditions, one arrives at the desired expressions.

Due to $p''(w) < 0$ for all $w < \overline{w}$ and $p''(\overline{w}) = 0$, it is immediate to see that $s(w) \leq s^{FB}$, with the inequality holding as equality if and only if $w = \overline{w}$. When $\gamma - r$ and σ_K are sufficiently small, then $p'''(w) > 0$ (see Corollary 5) for all w, and due to

$$sign\left(\frac{\partial s(w)}{\partial w}\right) = sign(p'''(w)),$$

short-run investment increases in w under these circumstances.

Evaluating the HJB equation at the boundary under the optimal controls yields

$$(r - \mu\ell)p(\overline{w}) + (\gamma - \mu\ell)\overline{w} = \alpha s - \mathcal{C}(s, \ell).$$

Hence, owing to $\gamma > r$ and agency-induced termination, $\mathcal{P}(\tau < \infty) = 1$:

$$p(\overline{w}) + \overline{w} < \left(\frac{\alpha s - \mathcal{C}(s,\ell)}{r - \mu\ell}\right) \leq p^{FB}.$$

Since $\mathcal{C}_\ell(s^{FB}, \ell^{FB}) = \mu p^{FB}$ and $\mathcal{C}_\ell(s(\overline{w}), \ell(\overline{w})) = \mu(p(\overline{w}) + \overline{w})$, it is clear that $\ell(\overline{w}) < \ell^{FB}$ and therefore by continuity that $\ell(w) < \ell^{FB}$ in a left neighborhood of $\overline{w}$.

3.9.4 Proof of Proposition 11

We prove Proposition 11 i) in two parts. Part I shows that either $\sigma_X = 0$ or $\sigma_K = 0$ implies $\ell(w) < \ell^{FB}$. Part I shows that there exist parameter values so that $\ell(w) > \ell^{FB}$ once $\sigma_X, \sigma_K > 0$. We start the proof with an auxiliary lemma.

Proof of Proposition 11—auxiliary results

Lemma 18 *Under the optimal contract for an arbitrary parameter $\theta \notin \{r, \mu\}$, it holds*

$$\frac{\partial p(w)}{\partial \theta} = \mathbb{E}\left\{ \int_0^\tau e^{-rt+\mu \int_0^t \ell_s ds} \left[\frac{\partial \alpha}{\partial \theta} s_t - \frac{\partial \mathcal{C}(s_t, \ell_t)}{\partial \theta} + p'(w_t) w_t \frac{\partial (\gamma - \mu \ell_t)}{\partial \theta} - p(w_t) \frac{\partial (r - \mu \ell_t)}{\partial \theta} \right.\right.$$
$$\left.\left. + \frac{p''(w_t)}{2} \frac{\partial}{\partial \theta} \left[(\beta_t^s \sigma_X)^2 + \sigma_K^2 (\beta_t^\ell - w_t)^2 + 2\rho \sigma_X \sigma_K \beta_t^s (\beta_t^\ell - w_t) \right] \right] dt \middle| w_0 = w \right\}.$$

Let $w \in [0, \overline{w}]$, $\theta \notin \{r, \mu\}$ and $s = s(w), \ell = \ell(w), \beta^s = \beta^s(w), \beta^\ell = \beta^\ell(w)$ be determined by the HJB Eq. (3.11).[24] Then, by the envelope theorem,

$$\frac{\partial p(w)}{\partial s} \frac{\partial s(w)}{\partial \theta} = \frac{\partial p(w)}{\partial \ell} \frac{\partial \ell(w)}{\partial \theta} = 0,$$

and therefore total differentiation of the HJB Eq. wrt. θ yields

$$\begin{aligned}(r - \mu \ell) \frac{\partial p(w)}{\partial \theta} + \frac{\partial (r - \mu \ell)}{\partial \theta} p(w) =& \frac{\partial \alpha}{\partial \theta} s - \frac{\partial \mathcal{C}(s, \ell)}{\partial \theta} + p'(w) w \frac{\partial (\gamma - \mu \ell)}{\partial \theta} + w(\gamma - \mu \ell) \frac{\partial}{\partial w} \frac{\partial p(w)}{\partial \theta} \\ &+ \frac{p''(w)}{2} \frac{\partial}{\partial \theta} \left[(\beta^s \sigma_X)^2 + \sigma_K^2 (\beta^\ell - w)^2 + 2\rho \sigma_X \sigma_K \beta^s (\beta^\ell - w) \right] \\ &+ \frac{\partial^2}{\partial w^2} \frac{\partial p(w)}{\partial \theta} \left[(\beta^s \sigma_X)^2 + \sigma_K^2 (\beta^\ell - w)^2 + 2\rho \sigma_X \sigma_K \beta^s (\beta^\ell - w) \right].\end{aligned}$$

Note that we used

$$\frac{\partial^k}{\partial w^k} \frac{\partial p(w)}{\partial \theta} = \frac{\partial}{\partial \theta} \frac{\partial^k p(w)}{\partial w^k} \text{ for } k \in \{1, 2\};$$

i.e., we changed the order of (partial) differentiation, which is possible since p is sufficiently smooth. The above ordinary differential equation admits a unique solution subject to the boundary conditions

$$\frac{\partial p(w)}{\partial \theta}|_{w=0} = 0 \text{ and } \frac{\partial p'(w)}{\partial \theta}|_{w=\overline{w}} = \frac{\partial}{\partial w} \frac{\partial p(w)}{\partial \theta}|_{w=\overline{w}} = 0,$$

and we can invoke Lemma 16 to arrive at the desired expression.

Proof of Proposition 11 i)—part I

Let us assume $\sigma_X = 0$ and state the following lemma:

Lemma 19 *Assume $\sigma_X = 0$. Hence, short-run investment $s(w)$ is contractible and constant over time. Then, it must be that $\beta^\ell > w$.*

The proof is split in several parts. Part a) shows that $\beta^\ell(\overline{w}) \neq \overline{w}$. Part b) shows that $\beta^\ell(w) \neq w$, and part c) concludes by showing $\beta^\ell(w) > w$ for all $w \in [0, \overline{w}]$.

a) Let us first show that $\beta^\ell(\overline{w}) = \lambda_\ell \ell(\overline{w}) \neq \overline{w}$. Define $\ell := \ell(\overline{w})$ and suppose to the contrary $\lambda_\ell \ell = \overline{w}$. Then

$$p(\overline{w}) = \frac{1}{r - \mu \ell} \left(\alpha s - \frac{1}{2} (\lambda_s^2 \alpha s + \lambda_\ell \ell^2 \mu) - \overline{w}(\gamma - \mu \ell) \right).$$

Let $\varepsilon > 0$ and consider the Taylor expansion of $p(\overline{w} - \varepsilon)$ around $p(\overline{w})$, given by $p(\overline{w} - \varepsilon) = p(\overline{w}) + \varepsilon + o(\varepsilon^3)$. Further, define $\ell_\varepsilon := \ell(\overline{w} - \varepsilon)$ and note that in optimum $\beta^\ell(\overline{w} -$

[24] For convenience, we suppress the dependence of $p(\cdot), \overline{w}$ on θ in the notation.

$\varepsilon) = \lambda_\ell \ell_\varepsilon + o(\varepsilon)$ by continuity. Hence

$$\begin{aligned}(r - \mu\ell_\varepsilon)p(\overline{w} - \varepsilon) &= \alpha s - \frac{\lambda_s \alpha s^2}{2} - \frac{1}{2}\lambda_\ell \ell_\varepsilon^2 \mu + p'(\overline{w} - \varepsilon)\Big((\gamma - \mu\ell_\varepsilon)(\overline{w} - \varepsilon)\Big) \\ &+ \frac{\sigma_K^2(\lambda_\ell \ell_\varepsilon + o(\varepsilon) - \overline{w} + \varepsilon)^2}{2} p''(\overline{w} - \varepsilon) \\ &= \alpha s - \frac{\lambda_s \alpha s^2}{2} - \frac{1}{2}\lambda_\ell \ell_\varepsilon^2 \mu + \big(-1 + o(\varepsilon^2)\big)\Big((\gamma - \mu\ell_\varepsilon)(\overline{w} - \varepsilon)\Big) \\ &+ \frac{\sigma_K^2(\lambda_\ell \ell_\varepsilon - \overline{w} + o(\varepsilon))^2}{2} p''(\overline{w} - \varepsilon),\end{aligned}$$

where we used that $p'(\overline{w} - \varepsilon) = p'(\overline{w}) - \varepsilon p''(\overline{w}) + o(\varepsilon^2)$.
Combining the above and using the Taylor expansion for $p(\overline{w} - \varepsilon)$ around $p(\overline{w})$ yields

$$\begin{aligned}p(\overline{w} - \varepsilon)\mu(\ell_\varepsilon - \ell) &= \varepsilon(r - \mu\ell) + (\gamma - \mu\ell_\varepsilon)(\overline{w} - \varepsilon) - \overline{w}(\gamma - \mu\ell) \\ &+ \frac{1}{2}\mu\lambda_\ell\big(\ell_\varepsilon^2 - \ell^2\big) - \frac{\sigma_K^2(\lambda_\ell \ell_\varepsilon - \overline{w} + o(\varepsilon))^2}{2} p''(\overline{w} - \varepsilon) + o(\varepsilon^2) + o(\varepsilon^3).\end{aligned}$$

Next, note that $\ell = \ell_\varepsilon + \varepsilon\ell'(\overline{w} - \varepsilon) + o(\varepsilon^2)$, in case $\ell(\cdot)$ is differentiable, which is guaranteed for $\varepsilon > 0$ sufficiently small. This yields

$$\begin{aligned}\mu p(\overline{w} - \varepsilon)(-\varepsilon\ell'(\overline{w} - \varepsilon)) &= \varepsilon(r - \gamma) - \overline{w}\mu\varepsilon\ell'(\overline{w} - \varepsilon) + o(\varepsilon^2) \\ &\iff o(\varepsilon) - \mu\big(p(\overline{w} - \varepsilon) + \overline{w}\big)\ell'(\overline{w} - \varepsilon) = r - \gamma.\end{aligned}$$

If $\ell(\overline{w}) = \ell_{\max}$, then it must be either that $\ell'(\overline{w} - \varepsilon) = o(\varepsilon)$ for ε sufficiently small, which leads to $\gamma - r = o(\varepsilon)$ and thereby a contradiction, or $\lim_{w\uparrow\overline{w}} \ell'(w) > 0$.
If $\ell(\overline{w}) < \ell_{\max}$ or $\lim_{w\uparrow\overline{w}} \ell'(w) > 0$, then $\ell(\overline{w} - \varepsilon)$ solves the following first-order condition of maximization, $\frac{\partial p(\overline{w}-\varepsilon)}{\partial \ell} = 0$. Moreover, $\ell(\overline{w}) < \ell_{\max}$ also solves the FOC at $w = \overline{w}$:

$$\mu p(\overline{w}) + \mu\overline{w} - \lambda_\ell \mu \ell(\overline{w}) = 0 \iff p(\overline{w}) + \overline{w} - \lambda_\ell \ell(\overline{w}) = 0.$$

Invoking the implicit function theorem, we can differentiate the above identity with respect to $w = \overline{w}$ so as to obtain $\ell'(\overline{w}) = 0$ as well as $\ell''(\overline{w}) = 0$. Then, by Taylor's theorem, which is applicable owing to $p \in C^2$, we get $\ell'(\overline{w} - \varepsilon) = o(\varepsilon^2)$, which yields the desired contradiction. This concludes the proof.

b) Let us assume that there exists now $w < \overline{w}$ with $\beta^\ell = \lambda_\ell \ell(w) = w$ optimal, in which case the HJB equation under the optimal control reads:

$$(r - \mu\ell(w))p(w) = \alpha s(w) - \frac{\lambda_s \alpha s(w)^2}{2} - \frac{\lambda_\ell \mu \ell(w)^2}{2} + p'(w)w(\gamma - \mu\ell(w)).$$

Due to $p'(w) \geq -1$ (i.e., since scaled payouts at rate $w(\gamma - \mu\ell)$ and this way keeping $w_t = w$ constant for all future times t is always an option but not necessarily optimal), it follows that

$$p(w) \geq \frac{1}{r - \mu\ell(w)}\Big(\alpha s(w) - \frac{\lambda_s \alpha s(w)^2}{2} - \frac{\lambda_\ell \ell(w)^2 \mu}{2} - w(\gamma - \mu\ell(w))\Big).$$

Likewise, due to the fact that $\ell(w)\lambda_\ell = w$ is optimal, it also must hold that

$$p(w) \geq \max_{s,\ell} \frac{1}{r - \mu\ell}\Big(\alpha s - \frac{\lambda_s \alpha s^2}{2} - \frac{\lambda_\ell \ell^2 \mu}{2} - w(\gamma - \mu\ell)\Big).$$

Then

$$\begin{aligned} p(w) < p(\overline{w}) - (w - \overline{w}) &= \max_{s,\ell} \frac{1}{r - \mu\ell}\Big(\alpha s - \frac{\lambda_s \alpha s^2}{2} - \frac{\lambda_\ell \ell^2 \mu}{2} - \overline{w}(\gamma - r) - w(r - \mu\ell)\Big) \\ &< \max_{s,\ell} \frac{1}{r - \mu\ell}\Big(\alpha s - \frac{\lambda_s \alpha s^2}{2} - \frac{\lambda_\ell \ell^2 \mu}{2} - w(\gamma - r) - w(r - \mu\ell)\Big) \\ &= \max_{s,\ell} \frac{1}{r - \mu\ell}\Big(\alpha s - \frac{\lambda_s \alpha s^2}{2} - \frac{\lambda_\ell \ell^2 \mu}{2} - w(\gamma - \mu\ell)\Big), \end{aligned}$$

where the first inequality is due to strict concavity and the second one due to $w < \overline{w}$. This yields the desired contradiction.

c) Eventually, let us assume that to the contrary $\beta^\ell = \ell(w)\lambda_\ell < w$ for at least one point w and define for this sake the function $\chi(w) = \beta^\ell(w) - w$. In a neighbourhood of $w = 0$ (i.e., on $(0, \varepsilon)$ for appropriate $\varepsilon > 0$), it is evident that $\chi(w) > 0$ because $0 < R = p(0) \Rightarrow \ell(0) > 0$.

If there is w' such that $\chi(w') < 0$, then there exists also w such that $\chi(w) = 0$ by continuity. If $\chi(w) = 0$ for some $w > 0$, then it must either be that $\chi(\overline{w}) = 0$, which contradicts part a), or $\chi(w) = 0$ for $0 < w < \overline{w}$, which contradicts part b). Hence, $\beta^\ell = \ell(w)\lambda_\ell > w$ for all $w \in [0, \overline{w}]$, which eventually proves the lemma.

Proof of Proposition 11 i)—part II

Here, we prove that $\sigma_K = 0$ and $\sigma_X > 0$ imply that $\ell(w) < \ell^{FB}$ provided investment is not at the corner.

For interior levels, $\ell = \ell(w)$ solves the first-order condition of maximization $\frac{\partial p(w)}{\partial \ell} = 0$ so that

$$\mu(p(w) - p'(w)w) - \lambda_\ell \mu \ell = 0.$$

Because of $p(w) - wp'(w) < p^{FB}$ and ℓ^{FB} solves $\mu p^{FB} - \lambda_\ell \mu \ell = 0$, it is immediate to see that $\ell(w) < \ell^{FB}$ for all $w \in [0, \overline{w}]$. For corner levels, a similar argument applies, which readily yields $\ell(w) \leq \ell^{FB}$ with the inequality being strict if $\ell_{\max} > \ell^{FB}$.

Proof of Proposition 11 ii)

Let θ denote an arbitrary set of model parameters and denote the family of solutions to the principal's problem by $\{p_\theta, \overline{w}_\theta\}_\theta$. By Berge's maximum theorem, $\overline{w}_\theta$ is continuous wrt. (the value of) θ, in the standard Euclidean metric space on $\mathbb{R}$ and p_θ is continuous in θ on $\mathbf{A}^{\mathbf{B}}$ with respect to the topology, induced by the norm $||\cdot,\cdot||_\infty$ where

$$||f||_\infty = \sup_{x \in \mathbf{A}} |f(x)|.$$

Here, $\mathbf{A}$, $\mathbf{B}$ are some compact subsets of $\mathbb{R}$ that satisfy all necessary regularity conditions and possibly depend on θ. We choose $\mathbf{A}$ sufficiently large so that $\overline{w}_\theta \in \mathbf{A}$ and $0 \in \mathbf{A}$ for all considered θ. We may choose $\mathbf{B}$ so that $p_\theta(w) \in \mathbf{B}$ for all $w \in [0, \overline{w}_\theta]$ for all considered θ. For brevity, we omit a formal introduction of the sets $\mathbf{A}, \mathbf{B}$ and the associated notation in the following.

Without loss of generality, we assume throughout that the constraint $\ell \leq \ell_{\max}$ is never tight. The proof goes through as long as the first-best level is interior (i.e., $\ell_{\max} > \ell^{FB}$). Formally dealing within the proof with corner levels would merely complicate the notation.

Let us start by considering the limit case $\mu \to 0$ holding the remaining parameters fixed. That is, we study the family $\{p_\mu, \overline{w}_\mu\}_{\mu \geq 0}$ and take the limit $\mu \to 0$. The model in the limit case $\mu \to 0$ is well behaved and features a value function p_0 with reflecting boundary $\overline{w}_0 > 0$. Due to continuity in μ, it follows that $p_\mu \to p_0$ and $\overline{w}_\mu \to \overline{w}_0$ as $\mu \to 0$. As a consequence, for any $w < \overline{w}_\mu$

$$\lim_{\mu\to 0} \ell(w) = \frac{-wp_0''(w)\lambda_\ell\sigma_K^2}{-p_0''(w)(\lambda_\ell\sigma_K)^2} = \frac{w}{\lambda_\ell} \wedge \ell_{\max} > 0,$$

where we omit for simplicity indexing for the optimal controls (e.g., for $\ell = \ell_\mu$). It can be verified for $w < \overline{w}_\mu$ that

$$\mathbb{V}(dw) = (\beta^s\sigma_X)^2 dt + (\beta^\ell - w)^2\sigma_K^2 dt = o(\sigma_X^2)dt,$$

when $\mu \to 0$ because $\beta^\ell \to \lambda_\ell\ell(w) = w$. If it were $\ell(w) = \ell_{\max}$, then it is easy to verify that $\beta^\ell = w$ becomes optimal.

As a consequence, the joint limit $\sigma_X, \mu \to 0$ would lead to a solution where $\{w\}$ has no volatility and accordingly $\overline{w}_0 = 0$. That is, $\lim_{\sigma_X\to 0}\overline{w}_0 = 0$. Hence, $\sigma_X = 0 \Rightarrow \overline{w}_0 = 0$. Since the limit case $\mu \to 0$ corresponds (effectively) to the model of DeMarzo and Sannikov (2006), we know that $\sigma_X > 0 \Leftrightarrow \overline{w}_0 > 0$. To take the limit $\lim_{\mu\to 0}\ell^{FB}$, we have to use L'Hospital's rule, which yields

$$\lim_{\mu\to 0}\frac{1}{\mu}\left[r - \sqrt{r^2 - \frac{\mu\alpha}{(\lambda_s\lambda_\ell)}}\right] = \lim_{\mu\to 0}\frac{1}{2\sqrt{r^2 - \frac{\mu\alpha}{(\lambda_s\lambda_\ell}}}\frac{\alpha}{\lambda_s\lambda_\ell} = \frac{\alpha}{2\lambda_s\lambda_\ell r}.$$

We prove now that there is long-termism, $\ell(w) > \ell^{FB}$, when $\gamma > r$ and $\mu > 0$ are sufficiently small. To avoid cluttering with subscripts, we omit indexing model quantities by μ or γ when it does not cause confusion. We evaluate the HJB equation at the boundary:

$$(r - \mu\ell(\overline{w}))p(\overline{w}) = \frac{\alpha}{2\lambda_s} - (\gamma - \mu\ell(\overline{w}))\overline{w} - \frac{\lambda_\ell\mu\ell(\overline{w})^2}{2},$$

and totally differentiate wrt. γ. Using $dp(\overline{w})/d\gamma = p'(\overline{w})\partial\overline{w}/\partial\gamma + \partial p(\overline{w})/\partial\gamma$ and the boundary condition $p'(\overline{w}) = -1$, we obtain

$$\frac{\partial\overline{w}}{\partial\gamma} = -\frac{1}{\gamma - r}\left(\overline{w} + (r - \mu\ell(\overline{w}))\frac{\partial p(\overline{w})}{\partial\gamma}\right).$$

Lemma 18 implies

$$\begin{aligned}\frac{\partial p(\overline{w})}{\partial\gamma} &= \mathbb{E}\left(\int_0^\tau e^{-rt+\mu\int_0^t \ell_u du}p'(w_t)w_t dt\middle| w_0 = \overline{w}\right)\\ &> -\mathbb{E}\left(\int_0^\tau e^{-rt+\mu\int_0^t \ell_u du}w_t dt\middle| w_0 = \overline{w}\right),\end{aligned} \tag{3.27}$$

where the inequality uses $p'(w) \geq -1$ with $p'(w) > -1$ for all $w < \overline{w}$.

We obtain that

$$\begin{aligned}\mathcal{A} &:= \overline{w} + (r - \mu\ell(\overline{w}))\frac{\partial p(\overline{w})}{\partial\gamma}\\ &> \overline{w}r\mathbb{E}\left(\int_0^\tau e^{-rt}dt\middle| w_0 = \overline{w}\right) - (r - \mu\ell(\overline{w}))\mathbb{E}\left(\int_0^\tau e^{-rt+\mu\int_0^t \ell_u du}w_t dt\middle| w_0 = \overline{w}\right) \qquad (3.28)\\ &= r\mathbb{E}\left(\int_0^\tau e^{-rt}(\overline{w} - w_t)dt\middle| w_0 = \overline{w}\right) + o(\mu).\end{aligned}$$

for all $\gamma > r$ and $\mu > 0$, where the first inequality uses that $\tau < \infty$ (almost surely) and (3.27). The remainder terms of order μ in the last line of (3.28) are denoted by $o(\mu)$ and tend to zero as $\mu \to 0$.

Since the process $\{w\}$ has—because of $\sigma_X > 0$ and $\beta^s(\overline{w}) = \lambda_s s^{FB}$—strictly positive volatility at the boundary $\overline{w}$, the payout boundary $\overline{w}$ cannot constitute an absorbing (or attracting) state. This holds true for any $\gamma > r$ and any $\mu > 0$, so that

$$\mathcal{E} := \mathbb{E}\left(\int_0^\tau e^{-rt}(\overline{w} - w_t)dt \big| w_0 = \overline{w}\right) > 0.$$

Due to that and the fact that the stochastic process $\{w\}$ possesses strictly positive volatility almost everywhere on $(0, \overline{w})$, it cannot be that the above expectation $\mathcal{E}$ tends to zero as $\gamma \to r$ or $\mu \to 0$, so that $\lim_{\gamma \downarrow r} \mathcal{E} > 0$ for any $\mu > 0$. Likewise, $\lim_{\mu \to 0} \mathcal{E} > 0$ for any $\gamma > r$, and

$$\lim_{\gamma \downarrow r} \lim_{\mu \to 0} \mathcal{E} > 0.$$

The above limit possibly takes value ∞.

It follows that $\mathcal{A} > 0$ when $\mu > 0$ is sufficiently small and $\lim_{\mu \to 0} \mathcal{A} > 0$. That is,

$$\lim_{\mu \to 0} \mathcal{A} \geq \lim_{\mu \to 0} r\mathbb{E}\left(\int_0^\tau e^{-rt}(\overline{w} - w_t)dt \bigg| w_0 = \overline{w}\right) > 0.$$

From there it follows that

$$\lim_{\mu \to 0} \frac{\partial \overline{w}}{\partial \gamma} = \frac{\lim_{\mu \to 0} \mathcal{A}}{-(\gamma - r)} \leq \frac{\lim_{\mu \to 0} r\mathbb{E}\left(\int_0^\tau e^{-rt}(\overline{w} - w_t)dt \bigg| w_0 = \overline{w}\right)}{-(\gamma - r)} < 0,$$

and accordingly

$$\lim_{\gamma \downarrow r} \lim_{\mu \to 0} \frac{\partial \overline{w}}{\partial \gamma} = \lim_{\gamma \downarrow r} \frac{\lim_{\mu \to 0} \mathcal{A}}{-(\gamma - r)} \leq \lim_{\gamma \downarrow r} \frac{\lim_{\mu \to 0} r\mathbb{E}\left(\int_0^\tau e^{-rt}(\overline{w} - w_t)dt \bigg| w_0 = \overline{w}\right)}{-(\gamma - r)} = -\infty,$$

as $\lim_{\gamma \downarrow r} \lim_{\mu \to 0} \mathcal{E} > 0$. Thus, it follows that

$$\lim_{\gamma \downarrow r} \lim_{\mu \to 0} \overline{w} = \lim_{\gamma \downarrow r, \mu \to 0} \overline{w} = \infty.$$

Take $\varepsilon > 0$ sufficiently small and define $\tilde{w} := \overline{w} - \varepsilon$, so that

$$\lim_{\mu \to 0} \ell(\tilde{w}) = \frac{\tilde{w}}{\lambda_\ell} \quad \text{and} \quad \lim_{\gamma \downarrow r} \lim_{\mu \to 0} \tilde{w} = \lim_{\gamma \downarrow r, \mu \to 0} \tilde{w} = \infty.$$

Therefore,

$$\lim_{\gamma \downarrow r, \mu \to 0} \ell(\tilde{w}) > \lim_{\gamma \downarrow r, \mu \to 0} \ell^{FB} = \lim_{\mu \to 0} \ell^{FB} = \frac{\alpha}{2\lambda_s \lambda_\ell r}.$$

By continuity, when $\mu > 0$ and $\gamma > r$ are sufficiently small, there exists $\tilde{w} < \overline{w}$ with $\ell(\tilde{w}) > \ell^{FB}$.

Let $w^H \equiv \sup\{w : \ell(w) > \ell^{FB}\}$ and $w^L \equiv \sup\{w : \ell(w) > \ell^{FB}\}$. Since $\ell(\overline{w}) < \ell^{FB}$ for any $\mu > 0, \gamma > r$, it must be that $w^H < \overline{w}$ with $\lim_{\mu \to 0, \gamma \to r} w^H = \overline{w}$. In addition, $\lim_{\mu \to 0, \gamma \to r} w^L = \frac{\alpha}{2\lambda_s r}$. It follows then that in the limit $\mu \to 0, \gamma \to r$ it must be that the set $\{w : \ell(w) = w/\lambda_\ell + o(\mu) \wedge \ell_{\max} > \ell^{FB}\}$ is convex. By continuity, when $\mu > 0$ and $\gamma > r$ are sufficiently small, the set $\{w : \ell(w) > \ell^{FB}\}$ is convex, thereby concluding the proof.

Proof of Proposition 11 iii)

We consider parameters are such that $\sup\{\ell(w) : w \in [0, \overline{w}]\} = \ell^{FB}$, and let $\mu > 0$ so that $\ell(w) > 0$ for any $w \in (0, \overline{w}]$. Let us evaluate the HJB equation at the boundary:

$$(r - \mu\ell(\overline{w}))p(\overline{w}) = \frac{\alpha}{2\lambda_s} - (\gamma - \mu\ell(\overline{w}))\overline{w} - \frac{\lambda_\ell \mu \ell(\overline{w})^2}{2}.$$

Differentiating this identity wrt. σ_i for $i \in \{X, K\}$ leads to

$$\frac{\partial \overline{w}}{\partial \sigma_i} = \frac{r}{r-\gamma}\frac{\partial p(\overline{w})}{\partial \sigma_i}.$$

Lemma 18 then implies:

$$\frac{\partial p(\overline{w})}{\partial \sigma_K} = \mathbb{E}\left(\int_0^\tau e^{-rt+\mu\int_0^t \ell_u du} p''(w_t)(\beta_t^\ell - w_t)^2 \sigma_K dt \middle| w_0 = \overline{w}\right) < 0,$$
$$\frac{\partial p(\overline{w})}{\partial \sigma_X} = \mathbb{E}\left(\int_0^\tau e^{-rt+\mu\int_0^t \ell_u du}\left(p''(w_t)(\beta_t^s)^2\sigma_X\right) dt \middle| w_0 = \overline{w}\right) < 0$$

so that $\overline{w}$ increases in σ_i for $i \in \{X, K\}$. The claim follows due to continuity in parameter values $\{\sigma_X, \sigma_K\}$.

3.9.5 Proof of Proposition 12 and 13

We prove the two propositions separately. In both cases claim i) is trivial since $\sigma_K = 0$ precludes risk externalities between short- and long-run incentives.

Proof of Proposition 12 ii)

The proof of Proposition 12 ii) is split in two parts. The first part of the proof shows that there is short-termism, $s(w) > s^{FB}$, for σ_X sufficiently small; the second one points out under which circumstances $\{w : s(w) > s^{FB}\}$ is convex.

Let us assume that correlation ρ is negative. Let us fix all parameters and consider the family of solution $\{p_{\sigma_X}, \overline{w}_{\sigma_X}\}$, which is–by Berge's maximum theorem—continuous in σ_X wrt. an appropriate topology, already discussed before. In the limit case $\sigma_X \to 0$, we have $s(w) \to s^{FB}$ for all $w \in [0, \overline{w}_{\sigma_X}]$. In addition, for any $\sigma_X \geq 0$, including the limit case $\sigma_X \to 0$, we have $p''_{\sigma_X}(0) < 0$, as $\overline{w}_{\sigma_X} > 0$ due to $\sigma_K > 0$. For notational convenience, we omit indexing model quantities by σ_X when no confusion is likely to arise.

We can write

$$s(w) = \frac{\alpha + p''(w)\rho\sigma_X\sigma_K\lambda_s(\lambda_\ell \ell(w) - w)}{\lambda_s\alpha - p''(w)(\lambda_s\sigma_X)^2} = \frac{\alpha + p''(w)\rho\sigma_X\sigma_K\lambda_s\lambda_\ell\ell(w)}{\lambda_s\alpha - p''(w)(\lambda_s\sigma_X)^2} + o(w).$$

From there it follows immediately that

$$\frac{\partial s(w)}{\partial p''(w)} = o(\sigma_X).$$

Thus

$$\begin{aligned}
\frac{ds(w)}{d\sigma_X} &= \frac{\partial s(w)}{\partial \sigma_X} + \frac{\partial s(w)}{\partial p''(w)}\frac{\partial p''(w)}{\partial \sigma_X} \\
&\propto \Big[\lambda_s\alpha - p''(w)(\lambda_s\sigma_X)^2\Big] p''(w)\rho\sigma_K\lambda_s(\lambda_\ell\ell(w) - w) \\
&\quad + 2\Big[\alpha + p''(w)\rho\sigma_X\sigma_K\lambda_s(\lambda_\ell\ell(w) - w)\Big] p''(w)\lambda_s^2\sigma_X + o(\sigma_X) \\
&= p''(w)\rho\sigma_K\lambda_s^2(\lambda_\ell\ell(w) - w)\alpha + o(\sigma_X) = p''(w)\rho\sigma_K\lambda_s^2\lambda_\ell\ell(w)\alpha + o(\sigma_X) + o(w),
\end{aligned}$$

where $\propto$ means "has the same sign as."

Because of $R > 0$, we have $\ell(0) > 0$. This implies $\ell(w) > 0$ close to zero and $\ell(w) \notin o(w)$. Hence, it holds that $\lambda_\ell\ell(w) > w$ in a neighborhood of zero, implying short-run investment $s(w)$ increases in σ_X, provided $\sigma_X > 0$ and w are sufficiently close to zero. This follows from $\lim_{\sigma_X \to 0} p'' \not\equiv 0$ and $p''(0) < 0$ because $\sigma_K > 0$ guarantees a nontrivial boundary $\overline{w} > 0$, even in the limit $\sigma_X \to 0$. Because of $s(w) = s^{FB}$, if $\sigma_X = 0$, there exists $\sigma_X > 0$ and $w \in [0, \overline{w}]$ so that $s(w) > s^{FB}$, which concludes the first part of the proof.[25]

The second part of the proof establishes the convexity of the set $\{w : s(w) > s^{FB}\}$ under certain parameters conditions. Let us calculate

$$\frac{\partial s(w)}{\partial w} \equiv s'(w) \propto p'''(w)\rho\sigma_X\sigma_K\lambda_s(\lambda_\ell \ell(w) - w) + p''(w)\rho\sigma_X\sigma_K\lambda_s \frac{\partial(\lambda_\ell \ell(w) - w)}{\partial w} + o(\sigma_X^2).$$

If $\gamma - r$ (and possibly σ_K^2) is sufficiently small, then $p'''(w) \geq 0$ (see Corollary 5) so that the first term is negative for $w < \lambda_\ell \ell(w)$ (i.e., for w close to zero). If λ_ℓ is sufficiently small, then

$$\frac{\partial(\lambda_\ell \ell(w) - w)}{\partial w} = \lambda_\ell \ell'(w) - 1 < 0$$

so that the second term is also negative. The remainder is negligible for σ_X sufficiently small. Under these conditions, $s'(w) < 0$ for $w \leq \lambda_\ell \ell(w)$.

Let us conclude the proof by demonstrating $\{w : s(w) > s^{FB}\}$ must be a convex set, containing zero when in addition to σ_X also λ_ℓ and $\gamma - r$ are sufficiently small so as to ensure $\partial s(w)/\partial w < 0$ for $w < \lambda_\ell \ell(w)$. Without loss of generality, assume that $\{w : s(w) > s^{FB}\}$ is nonempty. If the set is not convex, it must be that there exists $w' \in [0, \overline{w}]$ with $s(w') = s^{FB}$ and $s'(w') > 0$ such that $w' < \overline{w}$. Next, let us take a look at

$$s(w) = \frac{\alpha + p''(w)\rho\sigma_X\sigma_K\lambda_s(\lambda_\ell \ell(w) - w)}{\lambda_s\alpha - p''(w)(\lambda_s\sigma_X)^2},$$

and notice that for $s(w') \geq s^{FB}$ being optimal it is necessary that $\lambda_\ell \ell(w') > w'$, as $\rho < 0$ and $p''(w') < 0$. This implies $s'(w') < 0$ when λ_ℓ and $\gamma - r$ are sufficiently small, a contradiction.

Next, assume the set does not contain zero (i.e., $s(0) \leq s^{FB}$). It follows that $s'(\hat{w}) > 0$ for $\hat{w} = \inf\{w \geq 0 : s(w) > s^{FB}\}$. By continuity $s(\hat{w}) = s^{FB}$. However, due to $s(\overline{w}) = s^{FB}$, it must be that $\hat{w} < \overline{w}$. For $s(\hat{w}) = s^{FB}$ being optimal, it must be that $\lambda_\ell \ell(\hat{w}) > \hat{w}$. This implies $s'(\hat{w}) < 0$ when λ_ℓ and $\gamma - r$ are sufficiently small, a contradiction. This concludes the proof.

Proof of Proposition 13 ii)

Fix $\sigma_X > 0$ and consider $\gamma - r$ sufficiently small, such that there exists $w < \overline{w}$ with $w > \lambda_\ell \ell(w)$. This is possible as $\overline{w} \to \infty$ for $\gamma \to r$ and because there exists a left neighborhood of $\overline{w}$, where $\ell(w) < \ell^{FB} < \infty$ for any $\gamma > r$.

Note that this holds for any $\sigma_X > 0$. Therefore, we can choose σ_X sufficiently small and $\gamma - r$ sufficiently small so that there exists $w < \overline{w}$ with $p''(w) < 0$ and $s(w) > s^{FB}$ because of

$$s(w) = \frac{\alpha + p''(w)\rho\sigma_X\sigma_K\lambda_s(\lambda_\ell \ell(w) - w)}{\lambda_s\alpha + o(\sigma_X)^2},$$

that is, because the incentive cost of short-run investment is of order σ_X^2, while the incentive externality is of order σ_X. Taking the limit $\sigma_X \to 0$ is innocuous only because $\sigma_K > 0$ guarantees a nontrivial solution in this limit. To be more rigorous, one could mimick and adapt the argument of the proof of Proposition 12 ii).

Since in the limit $\mu \to 0$ for arbitrary $\sigma_X > 0$, long-term investment satisfies $\ell(w) \to \frac{w}{\lambda_\ell}$, it follows that $s(w) \to \hat{s}(w) < s^{FB}$ for $w < \overline{w}$, as $\mu \to 0$. From there, it follows readily that there exist $\mu > 0$, $\gamma - r$, and σ_X sufficiently small such that $\{w : s(w) > s^{FB}\}$ is nonempty and convex with its infimum exceeding zero and its supremum equal to $\overline{w}$.

---

[25] If we did not have $R > 0$, the proof is still valid as long as there exists a point $w < \overline{w}$ satisfying $\ell(w) > w/\lambda_\ell$. The existence of such a point can be ensured by appropriate λ_ℓ.

3.9.6 Proof of Proposition 14

Claim i) is straightforward, directly follows from the HJB equation, and is already explained in the main text.

Claim ii) is implied by the proof of Proposition 11 i), where we show that $\lambda_\ell \ell(w) > w$ for all w when $\sigma_X = 0$. The proof can be easily adjusted for linear cost (compare, e.g., He, 2009b).

Claim iii) relies on the premise that $\overline{w}$ increases in $1/(\gamma - r)$ with $\lim_{\gamma \downarrow r} \overline{w} = \infty$ and can be proven to mimick the argument of the proof of Proposition 11 ii). Moreover, the limit $\lambda_\ell \to 0$ leads to a well-behaved solution with a strictly positive payout threshold. Hence, it follows by continuity of the solution $\{p_{\lambda_\ell}, \overline{w}_{\lambda_\ell}\}_{\lambda_\ell \geq 0}$ that there exists w with $\beta^\ell(w) = w > \lambda_\ell$ for $0 < w < \overline{w}_{\lambda_\ell}$ when λ_ℓ is sufficiently small.

Chapter 4

Delegated Monitoring and Contracting[1]

4.1 Introduction

Financial intermediaries—such as private equity funds, hedge funds, and banks—play a crucial role in firm-level governance by both monitoring management's activities and directly influencing managerial compensation contracts. While these intermediaries may possess unique capacities to contain agency conflicts within firms, they are subject to agency frictions of their own. Thus, intermediation—specifically, delegated monitoring—is a double agency problem. In this paper, we analyze the joint presence of agency conflicts at the firm and intermediary levels and their effects on the design of incentive contracts.

To do so, we formulate a dynamic contracting model with intermediation in which investors (the principal) finance a firm run by a manager (the agent, she) and monitored by an intermediary (he). The agent controls the firm's output via costly but unobservable effort, which gives rise to firm-level moral hazard. The intermediary also affects the firm's output via costly effort capturing the intermediary's monitoring activity or his direct influence on firm performance. In addition, the intermediary offers a compensation contract to the agent, while the principal only contracts with the intermediary. Thus the intermediary faces a two-task problem, monitoring and contracting. Because both the intermediary's effort and the contract offered to the agent are unobservable to investors, moral hazard at the intermediary level arises. The moral hazard problems interact. On the one hand, agency conflicts at the firm level make it harder to discern the impact of intermediary's effort on firm's output and thereby can exacerbate agency conflicts at the intermediary level. That is, moral hazard propagates from the firm to the intermediary level. On the other hand, because the principal determines the agent's contract indirectly via the intermediary, the intermediary's moral hazard affects the agent's incentives. That is, moral hazard propagates from the intermediary to the firm level.

The model can broadly represent various forms of delegated monitoring and contracting, but the setting where it most directly applies is private equity investment. General partners (GPs, represented by the intermediary in the model) raise funds from limited partners (LPs, the principal) to invest in a target firm run by a manager (the agent). The LPs are usually passive but can offer risk-sharing opportunities for the GPs. Actions of both the manager and the GPs affect the firm's performance but are subject to moral hazard. The GPs are in charge of monitoring and governance, including compensation contracts of the manager (governance engineering of Kaplan and Strömberg (2009)), but can also directly influence the performance of the firm (operational engineering). An extensive literature studies the incentive contracts of GPs and of target firm managers, yet these two are commonly treated as separate and isolated incentives problems. There is little understanding if and how the position in the investment relationship affects the structure of incentive contracts. How to incentivize monitoring by GPs? How do GP incentives affect the contract of the manager? When do GPs and managers contribute more to firm performance, after good or poor performance? Should LPs be more directly involved in firm governance, bypassing GPs?

[1] This Chapter is based on Gryglewicz and Mayer (2019).

To answer such questions, it is essential to understand how the incentives of the agent and the intermediary interact with each other. The key feature of the optimal contracts that address the double moral hazard problem is compensation sensitive to the observable firm performance. However, the incentive role of performance pay is different for the agent and the intermediary. These differences reflect the positions of the agent and the intermediary in the hierarchy of the investment relationship. The intermediary's performance pay motivates the intermediary to monitor and also to incentivize managerial effort. As such, investors can indirectly influence the agent's incentives by providing incentives to the intermediary, who passes part of his incentives through the agent's contract. That is, incentives *trickle down* from the intermediary to the agent. The agent's performance pay incentivizes the agent's own effort, but it also indirectly affects the intermediary. As the intermediary increases monitoring effort, the resulting gains in the firm's output partially accrue to the agent due to the agent's performance pay. When the agent's stake in the firm is large, this effect leads to an *agency overhang* problem in the sense that the intermediary is reluctant to invest in monitoring effort as most of the benefits are reaped by the agent. Conversely, when the agent's stake in the firm is low after poor performance, the effect generates additional incentives for the intermediary to exert effort to avoid agency-induced distress. In both cases, incentives *trickle up* from the agent to the intermediary. The optimal set of contracts for the investment relationship accounts for both the trickle-up and trickle-down effects of incentives as well as for risk-sharing with investors.

We analyze the optimal design of incentives addressing the double moral hazard problem and the intermediary's multitasking problem. To address the agent's moral hazard problem, the agent's stake in the firm increases after good firm performance and decreases after poor performance. After sufficiently bad performance, the agent's contract is terminated, and the firm enters agency-induced distress. The sensitivity of the agent's stake to firm performance determines the agent's incentives to exert effort. It is costly to expose the agent to performance when in distress, and so accordingly, the trickle-down effect is weak, and the intermediary passes little incentives in the agent's contract. Conversely, it is relatively cheap to incentivize the agent away from distress, the trickle-down effect is strong, and the intermediary passes strong incentives in the agent's contract. As the sensitivity of compensation to performance increases after good performance, the agent's incentives are convex or option-like.

The shape of the intermediary's incentive contract is driven by the trickle-up effect of the agent's incentives on the intermediary. As the sign of the trickle-up effect depends on the agent's stake, the trickle-up effect generates disincentives for the intermediary after good performance (due to the agency overhang problem) and positive incentives after poor performance (to mitigate agency-induced distress). To counter disincentives after good performance, the intermediary's contract amplifies his exposure to cash flow shocks (akin to a leveraged position vis-à-vis the principal). To curb the excessive trickle-up incentives after poor performance, the intermediary's contract reduces his exposure to cash flow shocks (akin to risk sharing with the principal, but, notably, present also without risk-sharing motive). Taken together, the intermediary's exposure to cash flow shocks increases in firm performance. This means that the intermediary's monetary exposure to firm performance is convex and exhibits option-like features. Remarkably, the total incentives of the intermediary, which consist of the monetary exposure to cash flow shocks and the indirect exposure via the trickle-up effect, are no longer convex in firm performance as they are the highest in distress due to the strong positive trickle-up incentives. The mechanism is that the intermediary benefits from saving the firm from agency-induced distress, and this generates incentives without direct exposure to cash flows.

We also study the impact of the agent's and the intermediary's efforts on firm performance. Because the intermediary passes his incentives on the agent, one could have expected that the two efforts move in accord. That is, when the intermediary's contract strongly exposes the intermediary to firm performance, the intermediary would both exert high effort and pass strong incentives to exert effort in the agent's contract. The concurrence of efforts (and idleness) could have had a destabilizing effect on the firm. We show that this is not the case and the reason for this is that the trickle-up and trickle-down effects are time-varying and performance-sensitive. The agent's and the intermediary's efforts are proportional to the incentives as discussed above. Consequently, the agent puts most effort after good performance, and the intermediary puts most effort in distress. The model thus implies that the intermediary is primarily active in the firm when the agent's role is diminished after poor performance and steps back when the agent role is increased. The interaction of the two incentive problems endogenously generates stability in firm performance.

We extend the model to consider the effects of increased investor participation in firm-level governance. Such increased participation means that investors more directly influence managerial contract terms. To capture this in our setting, we consider a variant of our model in which the contract between the intermediary and the agent is publicly observable and contractible. That is, investors can essentially contract directly with the agent and merely have to incentivize the intermediary for his monitoring activity. We show that direct investor participation changes the level of incentives for the agent and the intermediary, especially after poor performance. The agent's pay sensitivity to performance increases and the intermediary's decreases relative to the case with delegated contracting. This happens because the intermediary—now facing only one task of his own effort—can be effectively insulated from the agent's incentive problem. Whereas agency-induced distress made the intermediary to exert more of his own effort and to delegate less in the agent's contract, this effect can be eliminated under direct contracting. Increased investor participation offers little benefit when the agency conflicts at the intermediary and the firm levels are symmetric. When the agency conflicts are significantly larger at either level, it can propagate to the other level under delegated contracting. It is then worthwhile to separate the agency problems by direct contracting with the agent. We also study a setting in which it is possible for investors to temporally increase or decrease participation in firm governance in response to past firm performance. Increased investor participation is optimal after poor firm performance when incentives under delegated contracting are distorted by weak trickle-down and strong trickle-up effects.

Finally, we show how the model can be adapted to settings in which the intermediary does not accept negative payments, i.e., does not inject funds in the firm except, possibly, for the initial investment. With such an assumption, the model fits other applications apart from our leading one of private equity investment. In particular, the intermediary can represent boards of directors who monitor and set executives' contracts on behalf of shareholders. In this setting, a shift from delegated to direct contracting can represent the introduction of say-on-pay regulations that increased shareholder participation in determining executive compensation. The model predicts that say-on-pay regulations raise the level of executives' performance pay and increase the sensitivity of pay to poor realizations of performance, consistently with empirical evidence. Shareholders' say-on-pay is particularly beneficial after poor performance when delegated contracting via the board leads to largest distortions in incentive provision.

Our theory focuses on the monitoring function of financial intermediaries and complements previous agency-based models of intermediation that consider other functions. Bhattacharya and Pfleiderer (1985) study delegated portfolio management within a one-period model with hidden information, while Ou-Yang (2003) studies portfolio management in a continuous-time model with moral hazard. Cuoco and Kaniel (2011), Kaniel and Kondor (2012), and Guerrieri and Kondor (2012) study the impact of delegated portfolio management on asset prices. Likewise, He and Krishnamurthy (2011, 2013) analyze financial intermediaries that facilitate access to risky assets in general equilibrium models with asset pricing implications.

The fact that the monitoring function of financial intermediaries is limited by their own moral hazard has been studied in the banking literature, starting with Diamond (1984). Other related contributions include Hellwig (2000) and Bond (2004).[2] More closely related to our paper is Holmstrom and Tirole (1997), who consider monitoring by financial intermediaries in an agency model. Their focus is on intermediaries' financial constraints and their effect on the provision of loans and on equilibrium interest rates. In contrast to these

[2] Bebchuk, Cohen, and Hirst (2017) discuss agency problems of various institutional investors other than banks, such as passive or active mutual funds and hedge funds.

theories, our model is dynamic, and its objective is to examine the provision of incentives for both the intermediary and the agent (firm manager).

Our paper is also related to the more general literature on multi-layered moral hazard problems. Strausz (1997) and Rahman (2012) study the optimally of delegation of monitoring to a supervisor, but without delegation of contracting. Macho-Stadler and Pérez-Castrillo (1998) compare a decentralized organizational structure with delegated contracting and a centralized structure (direct contracting with all agents) with possible side contracting between the agents. Baliga and Sjöström (1998) analyze the advantages of delegated contracting when the supervisor's effort is observable to the agent. Buffa, Liu, and White (2020) study when to delegate contracting in a model with two agents, in which the principal cannot fully commit to privately observed contracts. In contrast to our paper, all these models are static and primarily focus on the optimal choice of the organizational structure rather than on incentives and their interactions along the hierarchy. More broadly, our paper is related to the literature on hierarchies in which the agency friction is adverse selection instead of moral hazard (see Mookherjee, 2013, for a review). A large part of this literature focuses on contrasting various organizational forms and identifying conditions when delegation can dominate centralized organizational structures (e.g., Faure-Grimaud, Laffont, and Martimort, 2003; Mookherjee and Tsumagari, 2004). More closely related to our paper is the observation in Melumad, Mookherjee, and Reichelstein (1995) that in a static model of delegation under adverse selection, the middle agent tends to assign a higher production task to his favor at the expense of a lower production task assigned the lower agent. This is parallel to incentives trickling down in our model of delegation under moral hazard.

These multi-layered agency models and ours are distinct from two-sided agency problems, in which both the agent and the principal are subject to moral hazard, as in the static model of Bhattacharyya and Lafontaine (1995) or in the dynamic model of Hori and Osano (2013). Our theory also relates to papers considering optimal monitoring in two-player agency models in both dynamic (see, e.g., Piskorski and Westerfield, 2016; Halac and Prat, 2016; Dilmé and Garrett, 2019; Malenko, 2019a; Varas, Marinovic, and Skrzypacz, 2020b) and static settings (see, e.g., Lazear, 2006; Eeckhout, Persico, and Todd, 2010). The main difference of our model is that the monitoring party is an agent too who exerts effort and contracts with the ultimate agent.

Our paper is part of the growing literature on dynamic contracting models as in, among others, Holmstrom and Milgrom (1987), DeMarzo and Sannikov (2006), Biais et al. (2007), Sannikov (2008), Biais et al. (2010), He (2011), DeMarzo et al. (2012), Zhu (2012), He, Wei, Yu, and Gao (2017), Marinovic and Varas (2019a), Szydlowski (2015), and Gryglewicz et al. (2020). We contribute to this literature by adding an intermediary and considering delegated contracting.

The paper is organized as follows. Section 4.2 describes the model. Section 4.3 presents its solution and analysis. Section 4.4 studies increased investor participation in corporate governance by analyzing a variant of the model with an observable and contractible managerial contract. Section 4.5 considers a setting in which the intermediary is financially constrained. We also discuss the interpretation of the intermediary as the board of directors. Section 4.6 concludes. Proofs and technical details are deferred to the appendix.

4.2 Model

4.2.1 Setup

Time t is continuous on $[0, \infty)$. There are three players: the principal ("they" or "player P"), the intermediary ("he" or "player I"), and the agent ("she" or "player A"). The principal allocates funds to an intermediary who finances a firm operated by the agent. The agent affects firm performance with her effort a^A. The intermediary also contributes to firm performance with his effort a^I, which may capture the intermediary's monitoring activity or direct influence on firm performance.

For theoretical clarity, we assume that the agent's and the intermediary's efforts have independent impact on the firm's cash flows. That is, the cash flow process until firm liquidation (at endogenous time $\tau \in [0, \infty])$ is given by

$$dX_t = (a_t^A + a_t^I)dt + \sigma dZ_t, \tag{4.1}$$

where Z is a standard Brownian Motion and $\sigma > 0$ the constant volatility. Cash flows dX_t are publicly observable and contractible, whereas no player observes cash flow shocks dZ_t and cash flow shocks dZ_t are not contractible. Effort a_t^A is only observed by the agent and effort a_t^I is only observed by the intermediary; both efforts are not contractible. In addition, both efforts are bounded, $a_t^A, a_t^I \in [0, A]$ with a constant $A > 0$. Effort is costly in that the agent and the intermediary incur private flow costs of $g^A(a_t^A) := \frac{1}{2}\delta(a_t^A)^2$ and $g^I(a_t^I) := \frac{1}{2}\lambda(a_t^I)^2$, respectively. This specification gives rise to moral hazard at *both* the firm and the intermediary level. We focus on parameter configurations that imply optimal interior effort levels $a_t^A, a_t^I \in (0, A)$ at all times $t \geq 0$ until firm liquidation.

All players discount the future at rate $r > 0$. The principal representing diversified outside investors is risk-neutral and has deep pockets. The agent and the intermediary are risk-averse with CARA preferences and constant absolute risk-aversion $\theta^A \geq 0$ and $\theta^I \geq 0$ respectively. Risk aversion may reflect the agent's and intermediary's undiversified exposure to the firm. For $j = A, I$, the flow utility of player j is then given by

$$u^j(c_t^j, a_t^j) = -\frac{1}{\theta^j} \exp\left[-\theta^j \left(c_t^j - g^j(a_t^j)\right)\right],$$

with consumption c_t^j and cost of effort $g^j(a_t^j)$. Both the intermediary and the agent can smooth consumption and privately save and borrow at the rate r, with S_t^A (S_t^I) denoting the balance of the agent's (intermediary's) savings account. Consumption c_t^j and the savings balance S_t^j are not restricted to be positive and savings satisfy the standard transversality condition $\lim_{t\to\infty} \mathbb{E}e^{-rt}S_t^j = 0$, ruling out Ponzi schemes. In addition, the initial balance of the savings accounts is normalized to zero (i.e., $S_{0-}^A = S_{0-}^I = 0$).

All players have limited liability or, put differently, limited commitment, in that any player's continuation value from following the contractual relationship must exceed the player's respective outside option at any time $t \geq 0$. We stipulate that the outside option of the agent and the principal has zero monetary value. In addition, we normalize the recovery value upon firm liquidation to zero. The intermediary is residual claimant on total firm value and can always liquidate the firm, so the intermediary's outside option (in monetary terms) is equal to the firm's liquidation value, $R \geq 0$.

4.2.2 Contracting problem

A concrete application of our model is private equity: general partners (the intermediary) raise funds from limited partners (the principal) and invest at their discretion into a target firm operated by a manager (the agent). Private equity funds generally acquire target firms via limited-liability special propose vehicles which effectively grants general partners limited liability. Motivated by this application, the contracting problem is as follows. The intermediary as the firm's owner collects the firm's cash flows dX_t and offers a contract Π^P to the principal, specifying payouts to the principal, and a contract Π^A to the agent, specifying the agent's compensation. The intermediary is the residual claimant on total firm value and can extract all surplus from both the principal and the agent, reflecting that limited partners as outside investors are competitive and the agent as a manager has little or no bargaining power. Upon firm liquidation at time τ, the intermediary seizes the liquidation value worth R dollars. With a slight abuse of notation, we write $dX_\tau = R$ and $dX_t = 0$ for $t > \tau$, while dX_t follows (4.1) for times $t < \tau$. Note that while the intermediary has sufficiently deep pockets himself to finance the firm, the intermediary is risk-averse and benefits from sharing risk with the risk-neutral principal via the contract Π^P.[3] The terms of the contract Π^A offered to the agent are not observable to the principal and are not contractible between the principal and the intermediary. Section 4.4 studies the model when the contract Π^A is publicly observable and contractible, e.g., because the principal directly engages in the firm's governance.

The contract offered to the agent $\Pi^A = (w^A, \hat{a}^A, \hat{a}^I)$ specifies i) prescribed effort $\hat{a}^A$ for the agent, ii) prescribed effort for the intermediary $\hat{a}^I$, and iii) cumulative payouts (wages) w^A to the agent. Likewise, the contract offered to the principal $\Pi^P = (w^P, \bar{a}^A, \bar{a}^I)$ specifies i) prescribed effort $\bar{a}^A$ for the agent, ii) prescribed effort for the intermediary $\bar{a}^I$, and iii) cumulative payouts w^P to the principal. Throughout the paper, we consider *incentive compatible* contracts Π^A and Π^P that respect any player's limited liability and induce the intermediary and the agent to exert the prescribed effort levels (so that $\hat{a}^A = \bar{a}^A = a^A$ and $\hat{a}^I = \bar{a}^I = a^I$).

Because a firm's manager typically is not paid negative wages, we stipulate that payouts to the agent must be positive, in that $dw_t^A \geq 0$. In contrast, payouts to the principal dw_t^P can be negative. The interpretation of negative payouts $dw_t^P < 0$ is that the intermediary raises new funds from outside investors. In the context of private equity $dw_t^P < 0$ can reflect that the general partners call capital from the limited partners. Like in DeMarzo and Sannikov (2006), as the agent cannot be paid negative wages and is protected by limited liability, incentive provision may require to terminate the agent's contract Π^A at some time τ, leading firm liquidation and $dX_t = dw_t^A = dw_t^P = 0$ for $t > \tau$.[4] Firm liquidation is inefficient.

Given contract Π^A, the agent chooses her effort a^A and her consumption c^A to maximize

$$\begin{aligned} U_0^A(\Pi^A) &:= \max_{a^A, c^A} \mathbb{E}\left[\int_0^\infty e^{-rt} u^A(c_t^A, a_t^A) dt\right] \\ \text{s.t.} \quad dS_t^A &= rS_t^A dt - c_t^A dt + dw_t^A \\ S_{0^-}^A &= 0 \ \text{ and } \ \lim_{t\to\infty} \mathbb{E} e^{-rt} S_t^A = 0. \end{aligned} \tag{4.2}$$

Note that the agent's savings S_t^A grow with interest earnings $rS_t^A dt$ and wage payouts dw_t^A, but fall with her consumption $c_t^A dt$.

The intermediary chooses his effort a^I, his consumption c^I, the agent's contract Π^A, and the principal's contract Π^P to maximize

$$\begin{aligned} U_0^I(\Pi^I) &:= \max_{a^I, c^I, \Pi^A, \Pi^P} \mathbb{E}\left[\int_0^\infty e^{-rt} u^I(c_t^I, a_t^I) dt\right] \\ \text{s.t.} \quad dS_t^I &= dX_t - dw_t^P - dw_t^A + rS_t^I dt - c_t^I dt \\ S_{0^-}^I &= 0 \ \text{ and } \ \lim_{t\to\infty} \mathbb{E} e^{-rt} S_t^I = 0. \end{aligned} \tag{4.3}$$

Observe that the intermediary collects net dollar payoffs $dw_t^I := dX_t - dw_t^P - dw_t^A$ from financing the firm, which is the firm's cash flows after payouts to the principal and the agent. Finally, the principal's payoff derived from the contract Π^P is equal to the expected discounted stream of future payouts dw_t^P:

$$v_0^P := \mathbb{E}\left[\int_0^\infty e^{-rt} dw_t^P\right]. \tag{4.4}$$

[3]Section 4.5 discusses the model solution when the intermediary is financially constrained and does not have deep pockets.

[4]Admittedly, the intermediary could continue running the firm without the agent and exerting effort $a^I > 0$ after the agent's contract is terminated. For simplicity, we consider that the liquidation value satisfies $R \geq R_L$, where R_L is characterized in (4.76) in Appendix. As we show in Appendix 4.10.8, $R \geq R_L$ implies that the intermediary prefers liquidation over running the firm without the agent. Appendix 4.13.2 discusses this assumption in more detail and argues that is has no qualitative effects on the model solution.

Generally, the principal, the intermediary, and the agent may have different private information and therefore apply potentially different probability measures to evaluate payoffs. For simplicity, we do not distinguish between these probability measures in the main text and provide a formal discussion of this issue to Appendix 4.7.

Note that from a theory point of view, it is irrelevant whether the principal offers a contract Π^I to the intermediary and collects the firm's cash flows or the intermediary offers a contract to the principal Π^P and collects the firm's cash flows. The reason is that both the principal and the intermediary would like to minimize agency costs and maximize firm value. In particular, an equivalent formulation of the contracting problem is that the principal collects the firm's cash flows dX_t and offers the intermediary a contract $\Pi^I = (w^I, w^A, \bar{a}^A, \bar{a}^I)$, stipulating cumulative (net) payouts w^I to the intermediary and payouts w^A to the agent, while the intermediary offers a contract Π^A to the agent and intermediates funds (payouts) w^A from the principal to the agent.[5] Hence, our model describes *delegated contracting*, in that the principal delegates contracting with the agent to the intermediary.[6] To better match the private equity application, we lay out the problem such that the intermediary offers a contract to the principal instead of the other way around.

4.3 Model solution and analysis

4.3.1 First-best benchmark

We start by considering the *first-best* benchmark in which efforts a^A and a^I are publicly observable and contractible. The first-best solution is as follows. The risk-neutral principal absorbs all cash flow shocks via his payout scheme dw_t^P and the intermediary collects the residual cash flows $dX_t - dw_t^P$, hence $dw_t^P = \sigma dZ_t$ and $dX_t - dw_t^P = (a_t^A + a_t^I)dt$. In addition, the intermediary compensates the agent for the flow costs of effort, i.e., $dw_t^A = \frac{1}{2}\delta(a_t^A)^2 dt$. Optimal efforts (a_t^A, a_t^I) maximize the firm's expected cash flows net the costs of effort

$$a_t^A + a_t^I - \frac{\delta(a_t^A)^2}{2} - \frac{\lambda(a_t^I)^2}{2},$$

leading to

$$a_t^A = a_{FB}^A \equiv \frac{1}{\delta} \quad \text{and} \quad a_t^I = a_{FB}^I \equiv \frac{1}{\lambda}. \tag{4.5}$$

Note that in the first-best benchmark, optimal payouts and efforts are constant over time and the firm is never liquidated. Moreover, the total firm value reads

$$F^{FB} = \max_{a_t^A, a_t^I} \frac{1}{r}\left(a_t^A + a_t^I - \frac{\delta(a_t^A)^2}{2} - \frac{\lambda(a_t^I)^2}{2}\right) = \frac{1}{2r}\left(\frac{1}{\delta} + \frac{1}{\lambda}\right). \tag{4.6}$$

The reason for the stationarity of the first-best solution is that the financially-constrained agent is not exposed to firm risk. In the remainder of the section, we provide the solution for the full model with agency conflicts where the optimal contract requires exposing the intermediary and the agent to firm risk.

4.3.2 Optimal effort and consumption

Martingale representation. In the following, "player j" refers to "the intermediary or the agent," in that $j = A, I$. For any time $t < \tau$ and contracts Π^A and Π^P, we define player

[5] With $\Pi^P = (w^P, \bar{a}^A, \bar{a}^I)$ and $\Pi^A = (w^A, \hat{a}^A, \hat{a}^I)$ solving the intermediary's problem (4.3), one can define $\Pi^I = (w^I, w^A, \bar{a}^A, \bar{a}^I)$ with $dw_t^I = dX_t - dw_t^P - dw_t^A$. Then, Π^I is the optimal contract the principal offers to the intermediary under the alternative formulation of the problem, while the agent is offered the contract Π^A.

[6] Admittedly, in many applications of delegated contracting, the intermediary may not have deep pockets. Section 4.5 discusses the model solution when the intermediary is financially constrained, in a sense that $dw_t^I \geq 0$ at all times $t \geq 0$.

j's continuation utility as

$$U_t^j := \mathbb{E}_t \left[\int_t^\infty e^{-r(s-t)} u^j(c_s^j, a_s^j) ds \right] \tag{4.7}$$

for $j = A, I$. Using the martingale representation theorem, we can derive that

$$dU_t^j = rU_t^j dt - u^j(c_t^j, a_t^j)dt + (-\theta^j rU_t^j)\beta_t^j (dX_t - \hat{a}_t^A dt - \hat{a}_t^I dt), \tag{4.8}$$

where $dX_t - \hat{a}_t^A dt - \hat{a}_t^I dt = \sigma dZ_t$ when $\hat{a}_t^A = a_t^A$ and $\hat{a}_t^I = a_t^I$. Here, β_t^j captures the endogenous exposure to cash flows dX_t and is determined by the contracts Π^A and Π^P. Note that the intermediary affects the agent's continuation utility not only through the choice of Π^A and β_t^A but also via his unobservable and non-contractible effort a_t^I. For instance, by reducing his effort level a_t^I below the prescribed level $\hat{a}_t^I$ expected by the agent, the intermediary reduces the firm's realized cash flows dX_t by amount $(\hat{a}_t^I - a_t^I)dt$, thereby reducing the agent's continuation payoff U_t^A by amount $(-\theta^A rU_t^A)\beta_t^A(\hat{a}_t^I - a_t^I)dt$.

Likewise, for the principal's continuation payoff

$$v_t^P := \mathbb{E}_t \left[\int_t^\infty e^{-r(s-t)} dw_s^P \right], \tag{4.9}$$

we obtain

$$dv_t^P = rv_t^P dt - dw_t^P + \beta_t^P (dX_t - \bar{a}_t^A dt - \bar{a}_t^I dt), \tag{4.10}$$

where β_t^P is the principal's endogenous exposure to cash flow shocks and $dX_t - \bar{a}_t^A dt - \bar{a}_t^I dt = \sigma dZ_t$ when $\bar{a}_t^A = a_t^A$ and $\bar{a}_t^I = a_t^I$.

Optimal consumption. As player j can save and borrow at risk-free rate r, player j optimally smoothes consumption, implying that the marginal utility $\frac{\partial}{\partial c_t^j} u^j(c_t^j, a_t^j)$ is a martingale. As shown in the Appendix, player j's first order condition with respect to consumption implies that $\frac{\partial}{\partial c_t^j} u^j(c_t^j, a_t^j) = -\theta^j rU_t^j > 0$. By direct differentiation, $\frac{\partial}{\partial c_t^j} u^j(c_t^j, a_t^j) = -\theta^j u^j(c_t^j, a_t^j)$, so that $u_t^j(c_t^j, a_t^j) = rU_t^j$. Plugging this relation back into (4.8) implies $\mathbb{E}\, dU_t^j = 0$, hence U_t^j is a martingale. Also note that because $(-\theta^j rU_t^j)$ is the marginal utility, we can interpret β_t^j in equation (4.8) as the dollar sensitivity of U_t^j to cash flow shocks.

Optimal effort. Player j chooses effort a_t^j to maximize the sum of flow utility and the (expected) change in continuation utility, so that

$$a_t^j = \arg\max_{a^j} \left(u^j(c_t^j, a^j) + (-\theta^j rU_t^j)\beta_t^j a^j \right) dt. \tag{4.11}$$

Raising effort by one unit over $[t, t+dt)$ increases cash flows dX_t by $1dt$, thereby increasing continuation utility by $(-\theta^j rU_t^j)\beta_t^j dt$. At the same time, player j incurs higher costs of effort, reducing utility by $\frac{\partial}{\partial a_t^j} u^j(c_t^j, a_t^j)dt = -\frac{\partial}{\partial c_t^j} u^j(c_t^j, a_t^j)(g^j)'(a_t^j)dt$ which equals $(-\theta^A rU_t^j)(g^j)'(a_t^j)dt$. As a result, the incentive condition (4.12) in the following Lemma pins down the intermediary's and the agent's effort, and ensures $\hat{a}_t^j = a_t^j$.[7]

Lemma 20 *Optimal consumption is characterized by $u(c_t^j, a_t^j) = rU_t^j$ and optimal effort is characterized by*

$$a_t^A = \hat{a}_t^A = \frac{\beta_t^A}{\delta} \quad \textit{and} \quad a_t^I = \hat{a}_t^I = \frac{\beta_t^I}{\lambda}. \tag{4.12}$$

[7] With $(-\theta^j rU_t^j)$ being the marginal utility, the player j's first order condition with respect to effort becomes $\frac{\partial}{\partial a_t^j} u^j(c_t^j, a_t^j) + \beta_t^j(-\theta^j rU_t^j) = (-\theta^j rU_t^j)(\beta_t^j - (g^j)'(a_t^j)) = 0$.

That is, the exposure to cash flow shocks β_t^j makes player j's continuation utility U_t^j sensitive to the firm's cash flows and therefore provides incentives to exert effort. The sensitivity β_t^j quantifies player j's *incentives*. Note that the principal's exposure to cash flow shocks β_t^P plays a different role. Stipulating $\beta_t^P > 0$ transfers risk to the principal and therefore facilitates risk-sharing between the intermediary and the principal.

4.3.3 Certainty equivalent and deferred payouts

It is convenient to transform the continuation utility U_t^j into monetary units by considering the certainty equivalent (i.e., the continuation payoff in monetary units)

$$V_t^j = -\frac{\ln(-\theta^j r U_t^j)}{\theta^j r}. \tag{4.13}$$

Using the optimal consumption policy, $u^j(c_t^j, a_t^j) = rU_t^j$, we obtain

$$c_t^j = rV_t^j + g^j(a_t^j). \tag{4.14}$$

Thus, consumption net of the cost of effort is equal to the "interest" r on the certainty equivalent. The interpretation is that at time t, player j obtains the same continuation payoff as from retiring with savings worth V_t^j dollars.

With (4.8) and the $u_t^j(c_t^j, a_t^j) = rU_t^j$, we invoke Itô's Lemma to get

$$dV_t^j = \frac{1}{2}\theta^j r(\beta_t^j \sigma)^2 dt + \beta_t^j(dX_t - \hat{a}_t^A dt - \hat{a}_t^I dt). \tag{4.15}$$

Note that player j demands a risk compensation $\frac{1}{2}\theta^j r(\beta_t^j\sigma)^2 > 0$ for being exposed to cash-flow shocks so that the certainty equivalent increases on average, $\mathbb{E}\, dV_t^j > 0$.

The certainty equivalent payoff V_t^j consists two sources: i) the savings S_t^j player j has accumulated up to time t and ii) future payouts player j expects to receive after time t with value v_t^j. That is, the value of player j's deferred compensation or deferred payouts is given by $v_t^j \equiv V_t^j - S_t^j$. Using (4.2), (4.3), and (4.15), one can calculate

$$dv_t^j = rv_t^j dt + \frac{1}{2}\theta^j r(\beta_t^j\sigma)^2 dt + g^j(a_t^j)dt - dw_t^j + \beta_t^j(dX_t - \hat{a}_t^A dt - \hat{a}_t^I dt). \tag{4.16}$$

Condition (4.16) is the so-called *promise-keeping* constraint. The intuition is that the sum of expected payouts at time t and the change in future expected payouts must compensate player j for the time preference, risk, and effort cost, in that

$$\mathbb{E}[dw_t^j + dv_t^j] = \left(rv_t^j + \frac{1}{2}\theta^j r(\beta_t^j\sigma)^2 + g^j(a_t^j)\right) dt.$$

We can also derive the following integral expression for the value of deferred payouts v_t^j:

$$v_t^j = \mathbb{E}_t\left[\int_t^\infty e^{-r(s-t)}\left(dw_s^j - \frac{1}{2}\theta^j r(\beta_s^j\sigma)^2 ds - g^j(a_s^j)ds\right)\right]. \tag{4.17}$$

Indeed, v_t^j is the expected discounted value of payouts dw_s^j player j receives at future times $s \geq t$, adjusted for risk and effort cost.

Observe that v_t^j is the additional value player j gains from continuing within the contractual relationship on top of the private savings S_t^j. As a result, the agent's limited liability requires that $v_t^A \geq 0$ at any time t, because otherwise the agent would be better off leaving the firm. Likewise, the principal's limited liability requires that $v_t^P \geq 0$ for all times $t \geq 0$. Also note that the intermediary can always liquidate the firm, seize the recovery value of R dollars, and renege on the promised payouts to the principal and the agent. Thus, the intermediary's outside option is equal to R, so that limited liability requires $v_t^I \geq R$ to hold at all times $t \geq 0$. This leads to the following Lemma:

Lemma 21 *At all times $t \geq 0$, the agent's limited liability requires $v_t^A \geq 0$, the principal's limited liability requires $v_t^P \geq 0$, and the intermediary's limited liability requires $v_t^I \geq R$.*

4.3.4 Optimal contracting

To solve for the optimal contracts Π^A and Π^P, we apply dynamic programming techniques to maximize the intermediary's continuation payoff in utility terms U_0^I or equivalently in monetary terms $V_0^I = v_0^I + S_0^I$. Due to CARA preferences, there are no wealth effects in the intermediary's utility function and the intermediary selects the same course of action (except for the level of consumption) irrespective of the private savings level S_t^I along the life of the contracts. Thus, for the purpose of deriving the optimal contracts, it is without loss of generality to maximize the value of the intermediary's deferred payouts v_t^I.

The intermediary's maximization problem and the value of his deferred payouts v_t^I depend now on three state variables: i) the principal's continuation value v_t^P, ii) the agent's continuation value V_t^A, and iii) the agent's savings S_t^A. Again, due to CARA preferences, there are no wealth effects in the agent's utility function and the absolute levels of V_t^A and S_t^A can be replaced by the value of deferred compensation $v_t^A = V_t^A - S_t^A$ to solve the intermediary's problem. As a result, the value of intermediary's deferred compensation v_t^I is a function of v_t^A and v_t^P, in that $v_t^I = \hat{F}(v_t^A, v_t^P)$. In what follows, we omit time subscripts to simplify notation.

To characterize the optimal contracts Π^P and Π^A, we proceed in three steps. First, we discuss the optimal timing of payouts to the principal and the agent. Second, we characterize the optimal provision of incentives to the agent β^A and the optimal risk-sharing with the principal, β^P. Third, we apply the dynamic programming principle to derive the HJB equation for the intermediary's problem.

Optimal payout timing

Note that by (4.10), the intermediary can always increase or decrease the principal's continuation payoff v_t^P by stipulating $dw_t^P < 0$ or $dw_t^P > 0$. Thus, in optimum, $\hat{F}(v_t^A, v_t^P - dw_t^P) - dw_t^P = \hat{F}(v_t^A, v_t^P)$. Since this must hold for any v_t^P and dw_t^P, it follows that $\frac{\partial}{\partial v_t^P}\hat{F}(v_t^A, v_t^P) = -1$. Hence, $\hat{F}(v^A, v^P) = F(v^A) - v^P$ for some function $F(v^a)$.

Payments to the agent dw_t^A must be positive. Consequently, it is always possible to decrease but not to increase the agent's deferred compensation v^A with payouts to the agent $dw^A \geq 0$. Thus, $\hat{F}(v^A - dw^A, v^P) - dw^A \leq \hat{F}(v^A, v^P)$. Using $\hat{F}(v^A, v^P) = F(v^A) - v^P$ and taking $dw^A \to 0$, it follows that $F'(v^A) \geq -1$ with equality if $dw^A > 0$.

When v^A falls down to zero, the volatility of dv^A must vanish to respect the agent's limited liability constraint, $v^A \geq 0$. Hence, $\beta^A = 0$ and the agent does not receive incentives and payouts anymore at all future times, which means that the agent's contract is terminated. Termination of the agent's contract also implies firm liquidation so that $v^P = 0$ and

$$F(0) = R. \tag{4.18}$$

Conversely, incentive provision (i.e., $\beta^A > 0$) requires to delay payouts to the agent, so that $v^A > 0$. Providing incentives β^A to the agent raises the volatility of v^A and, therefore, the risk of firm liquidation, which is costly. To reduce these agency costs, the intermediary delays payouts to the agent until the firm's distance to liquidation is sufficiently large and v^A exceeds the payout boundary $\overline{v}$, so $dw^A = \max\{0, v^A - \overline{v}\}$. At the payout boundary, the smooth pasting condition

$$F'(\overline{v}) = -1 \tag{4.19}$$

holds. As the intermediary and the agent discount at the same rate $r > 0$, delaying payouts to the agent is not costly for the intermediary but reduces the risk of firm liquidation, which is beneficial. Thus, the intermediary optimally delays payouts to the agent as much as his limited liability allows, and the limited liability constraint

$$v^I = \hat{F}(v^A, v^P) = F(v^A) - v^P \geq R, \tag{4.20}$$

holds with equality for $v^A = \overline{v}$.

As the intermediary and the principal discount at the same rate r, the intermediary is indifferent between paying the principal now or later. To minimize the firm's liquidation risk and hence to maximize $\overline{v}$, the intermediary designs the principal's compensation to relax his limited liability constraint (4.20), $F(v^A) - v^P \geq R$. As a result, it is optimal to set v^P at its minimum (i.e., $v_t^P = 0$ for all $t \geq 0$), leading to $\hat{F}(v^A, v^P) = F(v^A)$. Therefore, at the payout boundary $\overline{v}$, the limited liability constraint (4.20) holds with equality with $v^P = 0$, i.e.,

$$F(\overline{v}) = R. \tag{4.21}$$

Note that $F(v^A)$ is the firm value net of the agent's deferred compensation and is simply called firm value.[8] Observe that $v_t^P = 0$ for all $t \geq 0$ pins down by (4.10) the payouts to the principal

$$dw^P = \beta^P(dX - \bar{a}^A dt - \bar{a}^I dt). \tag{4.22}$$

Hence, the principal absorbs fraction β^P of the firm's cash flow shocks σdZ. We conclude this section with the following proposition.

Lemma 22 *The following holds*

1. *Payouts to the principal satisfy* $dw_t^P = \beta^P(dX_t - \bar{a}_t^A dt - \bar{a}_t^I dt)$*, leading to* $v_t^P = 0$ *and* $\hat{F}(v^A, v^P) = F(v^A)$*.*
2. *Payouts to the agent satisfy* $dw^A = \max\{v^A - \overline{v}, 0\}$ *where* $\overline{v}$ *is characterized by* $F'(\overline{v}) - 1 = F(\overline{v}) - R = 0$*.*
3. *The firm is liquidated when* $v^A = 0$*, leading to* $F(0) = R$*.*

Incentives and risk-sharing

In this section, we discuss the mechanism jointly determining the levels of risk-sharing with the principal and of incentives provided to the agent and the intermediary. Specifically, we show that the interaction between the incentives of the agent and the intermediary can be understood as a combination of trickle-down and trickle-up effects, in which the intermediary's incentives trickle down to the agent and the agent's incentives trickle up to influence the intermediary's incentives.

Note that according to (4.16), $\beta^I\sigma$ is the volatility of the intermediary's instantaneous payoff $dw^I + dv^I$, consisting of instantaneous dollar payoffs $dw^I = dX - dw^P - dw^A$ and change in future payoffs dv^I. When $v^A < \overline{v}$, there are no payouts to the agent (i.e., $dw^A = 0$) and the intermediary's instantaneous dollar payoff is $dX - dw^P$, implying that the intermediary absorbs fraction $1 - \beta^P$ of the firm's cash flow shocks dX. By Itô's Lemma, $dv^I = dF(v^A)$ has volatility $F'(v^A)\beta^A\sigma$, where $F'(v^A)$ captures the sensitivity of the intermediary's payoff to changes in the agent's compensation. That is, β^I can be decomposed as[9]

$$\beta^I = \underbrace{1 - \beta^P}_{\substack{\text{Direct exposure} \\ \text{to cash flow risk}}} + \underbrace{F'(v^A)\beta^A}_{\substack{\text{Exposure to} \\ \text{endogenous risk}}}. \tag{4.23}$$

The intermediary's direct exposure to cash flow risk is equal to $1 - \beta^P$, while his exposure to endogenous risk or his indirect exposure to cash flow risk through the agent's compensation contract is equal to $F'(v^A)\beta^A$.

[8] Strictly speaking, total firm value also involves the agent's stake v^A and is given by $F(v^A) + v^A$.

[9] Note that (4.23) also holds at the payout boundary when $v^A = \overline{v}$ and $F'(\overline{v}) = -1$. When $v^A = \overline{v}$, (4.16) implies that $dv^I + dw^A$ has volatility $(1 - \beta^P)\sigma$, as $dX - dw^P$ has volatility $1 - \beta^P$. Itô's Lemma implies that $dv^I + dw^A$ has volatility $F'(\overline{v})\beta^A\sigma = -\beta^A\sigma$, which leads to (4.23).

The indirect exposure of the intermediary to cash flow risk via the agent's incentives means that incentives *trickle up* from the agent to the intermediary. The intuition behind this finding is as follows. By exerting more effort a^I, the intermediary increases the firm's cash flows dX. Because the intermediary's effort a^I is not observable to the agent and not contractible, the increase in cash flows dX raises the value of the agent's continuation payoff V^A and of her deferred payouts v^A, so part of the gains generated by the intermediary's effort accrue to the agent.[10] The increase in v^A has two opposing effects as it reduces the firm's liquidation risk but increases the costs of compensating the agent. When $F'(v^A) > 0$ ($F'(v^A) < 0$), the first (second) effect dominates, leading to additional trickle up *incentives* (*dis-incentives*). That is, $F'(v^A)$ quantifies magnitude of the trickle-up effect.

The intermediary's direct exposure to cash flow shocks $1 - \beta^P$ influences the level of incentives the intermediary provides to the agent, creating what we term the *trickle-down* effect. To quantify it, observe that at any point in time t, the intermediary chooses the agent's incentives β^A to maximize the sum of expected cash flows after payouts to the principal and the change in value, adjusted for the cost of effort and the risk he bears. That is, the intermediary solves

$$\max_{\beta^A \geq 0} \left(\underbrace{(1-\beta^P)(a^A + a^I)dt + \beta^P(\bar{a}^A + \bar{a}^I)dt}_{=\mathbb{E}\,[dX - dw^P]} + \mathbb{E}[dF(v^A)] - \frac{\lambda(a^I)^2}{2}dt - \frac{\theta^I r}{2}(\beta^I\sigma)^2 dt \right), \tag{4.24}$$

taking into account the effort incentive constraints (4.12) and the composition of incentives in (4.23). Note that the expectation $\mathbb{E}\,[dX - dw^P]$ is taken under the intermediary's information that accounts for a potential difference in actual effort levels a^j and prescribed effort levels $\bar{a}^j$ anticipated by the principal, while $a_j = \hat{a}_j$ due the incentive condition (4.12). Specifically, under the intermediary's information, $dX_t - a^A dt - a^I dt$ has expectation zero.[11]

The intermediary only internalizes fraction $1 - \beta^P$ of the *actual* expected output $(a^A + a^I)dt$ because the principal attributes changes in cash flows due to unobservable deviations as realizations of cash flow shocks of which she receives fraction β^P. Solving the maximization problem (4.24) leads to the following lemma.

Lemma 23 *The intermediary's incentives satisfy* (4.23). *The agent's incentives satisfy* $\beta^A = (1 - \beta^P)\pi^I$ *with*

$$\pi^I = \pi^I(v^A) := \frac{\overbrace{1}^{\textit{Cash flows}} \; \overbrace{-\delta\theta^I r\sigma^2 F'(v^A)}^{=C_2}}{\underbrace{\frac{\delta(F'(v^A))^2}{\lambda}}_{=D(>0)} + \underbrace{\theta^I r\sigma^2(F'(v^A))^2}_{=C_1(>0)} \underbrace{-(1+\delta\theta^A r\sigma^2)F'(v^A)}_{=B} \underbrace{-\delta\sigma^2 F''(v^A)}_{=A(>0)}}. \tag{4.25}$$

The equation $\beta^A = (1-\beta^P)\pi^I$ *combined with* (4.23) *also implies that*

$$\beta^I = (1 + F'(v^A)\pi^I)(1 - \beta^P) \tag{4.26}$$

and after substituting $1 - \beta^P = \frac{\beta^A}{\pi^I}$*:*

$$\beta^A = \beta^I \left(\frac{\pi^I}{1 + F'(v^A)\pi^I} \right). \tag{4.27}$$

[10] For a more formal argument rewrite (4.16) for $j = A$ to $dv_t^A + dw_t^A = (rv_t^A + g^A(a_t^A) + \frac{1}{2}\theta^A r(\beta_t^A\sigma)^2)dt + \beta_t^A(a_t^I - \hat{a}_t^I)dt + \beta_t^A(dX_t - \hat{a}_t^A dt - a_t^I dt)$ and note that $\beta_t^A(dX_t - \hat{a}_t^A dt - a_t^I dt)$ has expectation zero under the intermediary's information, as the intermediary observes both the prescribed effort $\hat{a}_t^A$ and his own effort a_t^I. Thus, under the intermediary's information: $\frac{\partial}{\partial a_t^I}\mathbb{E}[dv_t^A + dw_t^A] = \beta_t^A dt$.

[11] Formally, this expectation is taken under the probability measure $\mathbb{P}^I$ that is induced by efforts $a^I = \hat{a}^I$ and $a^A = \hat{a}^A$. Under this probability measure, $(dX - a^A dt - a^I dt)/\sigma$ is the increment of a standard Brownian motion and $dX - (a^A + a^I)dt$ has expectation zero. For details, see Appendix 4.7. For a derivations of (4.24), we use (4.22) to get $dX - dw^P = (1-\beta^P)dX + \beta^P(\hat{a}^A + \hat{a}^I)dt$, which is in expectation (under the probability measure $\mathbb{P}^I$) $(1-\beta^P)(a^A + a^I)dt + \beta^P(\bar{a}^A + \bar{a}^I)dt$.

The first part of the lemma states that the agent's incentives $\beta^A = (1-\beta^P)\pi^I$ increase with the intermediary's incentives $1-\beta^P$ generated by his direct exposure to cash flow shocks, in that incentives trickle down from the intermediary to the agent. The strength of this trickle-down effect is captured by π^I given in (4.25). This coefficient reflects that by raising the agent's incentives, the intermediary boosts the agent's effort a^A and cash flows but also increases the risk of liquidation (term A), the cost of compensating the agent (term B), and the intermediary's risk exposure (terms C_1 and C_2). Lastly, the choice of the agent's incentives affects the intermediary's incentives to exert effort via the trickle-up effect (term D). The trickle-up effect in term D decreases π^I and undermines trickle-down incentives because it inadvertently moves the intermediary's incentives and effort, which is in principle costly. In other words, term D captures the shadow cost of constraint (4.23) linking the intermediary's and agent's incentives. The shadow cost is lower when δ/λ is low, that is, when the agent's effort is relatively cheaper and the intermediary focuses more on efficient incentive provision to the agent than on distortions to his own effort.

The combination of trickle-up and trickle-down incentives induces a feedback loop between the agent's and the intermediary's incentives, leading to the fixed point provided in the second part of the lemma, equation (4.26). As a result, the intermediary's total incentives β^I reflect both the trickle-up and the trickle-down effect via $F'(v^A)\pi^I$. When $F'(v^A) > 0$, trickle-up and trickle-down incentives reinforce each other and amplify the intermediary's incentives from direct exposure to cash flow shocks $1-\beta^P$. When $F'(v^A) < 0$, incentives trickle down from the intermediary to the agent but generate trickle-up dis-incentives, dampening the intermediary's incentives from direct cash flow exposure $1-\beta^P$.

4.3.5 HJB equation

We now derive the HJB equation for the value of the intermediary's deferred payouts, $F(v^A)$. To begin with, recall that the integral expression for v_t^I in (4.17) implies that

$$F(v^A) = \mathbb{E}_t\left[\int_t^\infty e^{-r(s-t)}\left(dX_s - dw_s^P - dw_s^A - \frac{\lambda(a_s^I)^2}{2}ds - \frac{\theta^I r(\beta_s^I\sigma)^2}{2}ds\right)\bigg| v_t^A = v^A\right].$$

Note that when $v^A \in [0,\overline{v}]$, there are no payouts to the agent (i.e., $dw^A = 0$), hence the intermediary's expected instantaneous payoff equals $f(v^A)dt := \mathbb{E}\,dX - \frac{\lambda(a^I)^2}{2}dt - \frac{\theta^I r(\beta^I\sigma)^2}{2}dt$, due to $\mathbb{E}\,dw^P = 0$. By the dynamic programming principle, the expected flow payoff $f(v^A)dt$ and the expected change in payoff $\mathbb{E}\,dF(v^A)$ must in optimum compensate the intermediary for his time preference $rF(v^A)dt$, so that $rF(v^A)dt = f(v^A)dt + \mathbb{E}\,dF(v^A)$. Using Itô's Lemma to calculate $\mathbb{E}\,dF(v^A)$, we can derive that $F(v^A)$ solves on $[0,\overline{v}]$ the following HJB equation

$$\begin{aligned} rF(v^A) = \max_{\beta^A,\beta^I}\Bigg\{ & a^A + a^I - \frac{\lambda(a^I)^2}{2} - \frac{\theta^I r(\beta^I\sigma)^2}{2} \\ & + F'(v^A)\left(rv^A + \frac{\theta^A r(\beta^A\sigma)^2}{2} + \frac{\delta(a^A)^2}{2}\right) + \frac{F''(v^A)(\beta^A\sigma)^2}{2}\Bigg\}, \end{aligned} \tag{4.28}$$

subject to the boundary conditions $F(0) = R$ (liquidation), $F'(\overline{v}) = -1$ (smooth pasting), and $F(\overline{v}) = R$ (limited liability). The choice of β^A and β^I is subject to the effort incentive constraint (4.12) the characterization of the agent's incentives (4.27). As $\beta^I = (1+F'(v^A))(1-\beta^P)$, maximizing (4.28) over β^I by choosing the intermediary's optimal incentives is equivalent to maximizing (4.28) over β^P by choosing the optimal risk-sharing and the intermediary's optimal exposure $1-\beta^P$ to cash flow shocks. Also note that when the incentive conditions (4.12) and (4.27) hold, prescribed and actual effort levels coincide and, unlike in the objective (4.24), we can directly write $\mathbb{E}\,dX = (a^A + a^I)dt$.

Optimal risk-sharing β^P and, equivalently, optimal incentives for the intermediary β^I are then obtained by solving the first order condition $\frac{\partial F(v^A)}{\partial \beta^I} = 0$ taking into account (4.12) and (4.27). This yields

$$1-\beta^P = \frac{\overbrace{\frac{1+F'(v^A)\pi^I}{\lambda}}^{=A} + \overbrace{\frac{\pi^I}{\delta}}^{=C}}{\underbrace{(1+F'(v^A)\pi^I)^2\left(\frac{1}{\lambda}+\theta^I r\sigma^2\right)}_{=B}\underbrace{-F'(v^A)(\pi^I)^2\left(\frac{1}{\delta}+\theta^A r\sigma^2\right)}_{=D}\underbrace{-F''(v^A)(\pi^I\sigma)^2}_{=E}}, \tag{4.29}$$

with π^I defined in (4.25) and $\beta^I = (1-\beta^P)(1+F'(v^A)\pi^I)$. The optimal provision of incentives to the intermediary reflects the trickle-down and trickle-up effects of incentives in addition to the direct impact of incentives on intermediary's effort. Reducing risk-sharing β^P or, equivalently, raising $1-\beta^P$ increases the intermediary's incentives β^I and effort a^I, which increases cash flows (term A) but also increases the intermediary's required compensation for risk and effort costs (term B). Moreover, incentives $1-\beta^P$ trickle down to the agent, which increases cash flows through the agent's effort (term C) as well as the cost of compensating the agent (term D) and the firm's liquidation risk (term E).

Finally, combining the findings of this section with the previous Lemmata 20-23, we can complete the characterization of the optimal contracts Π^A and Π^P with the following proposition.

Proposition 15 *In optimum, the following holds:*

1. *The value of the intermediary's deferred payouts $F(v^A)$ solves* (4.28) *subject to the boundary conditions $F(0)-R = F'(\overline{v})-1 = F(\overline{v}) = R$.*
2. *Payouts satisfy $dw^A = \max\{v^A-\overline{v},0\}$ and $dw^P = \beta^P\sigma dZ$, leading to $v^P=0$.*
3. *The function F is strictly concave, in that $F''(v^A)<0$ for all $v^A \in [0,\overline{v}]$.*
4. *The sensitivities β^A, β^P and β^I are characterized by* (4.23), (4.27), *and* (4.29)*. Effort is characterized by* (4.12).

Figure 4.1 presents a numerical example of the solution under our baseline parameters. Following He (2011), we use the discount rate $r = 0.05$ and the agent's risk aversion coefficient $\theta^A = 5$ and we normalize volatility and agency cost parameters to one (i.e., $\delta = \lambda = \sigma = 1$). For simplicity, we consider a risk-neutral intermediary with $\theta^I = 0$ and the recovery value is set to $R = 12.5$.[12] The upper left panel of Figure 4.1 shows that $F(v^A)$ is hump-shaped and concave. The concavity of $F(v^A)$ reflects that increasing the volatility of v^A by providing stronger incentives β^A to the agent is costly because it increases the risk of firm liquidation. Also note that an increase in v^A has two opposing effects, as it reduces the firm's liquidation risk but also increases for the intermediary the cost of compensating the agent. When v^A is small (large), the first (second) effect dominates and $F'(v^A) > 0$ ($F'(v^A) < 0$). Observe that $F'(v^A)$ switches sign exactly once when $v^A = v^*$ with $F'(v^*) = 0$ and $F(v^A)$ has its peak, which is denoted in Figure 4.1 by the dashed red line.

4.3.6 Dynamics of incentives

We study the dynamics of the agent's and the intermediary's incentives, β^A and β^I. By the incentive constraint (4.12), effort a^j is directly proportional to incentives β^j, hence the dynamics of effort and incentives follow a similar pattern. The upper right panel of Figure 4.1 illustrates the standard result that the agent's incentives increase in v^A, that is, after positive cash flow realizations. The reason is that when v^A is low, the firm's distance to liquidation is small. Therefore, the cost of providing incentives to the agent as captured by $-F''(v^A) > 0$ is high and the strength of the trickle-down incentives π^I is low, which hampers incentive provision to the agent.

[12] Appendix 4.13.3 studies the impact of intermediary risk-aversion and shows that the risk-aversion coefficient θ^I has no qualitative effect on the optimal contracts and the shape of the value function.

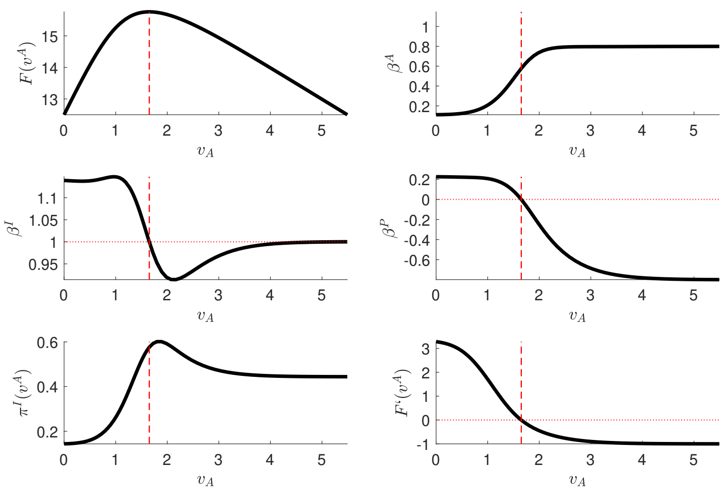

FIGURE 4.1: Numerical example of the model solution. The dashed red line denotes the value of v^A where $F'(v^A) = 0$ and $F(v^A)$ has the peak.

The middle left panel of Figure 4.1 plots the intermediary's incentives β^I against v^A. For low values of v^A when $F'(v^A) > 0$, strong trickle-up incentives lead to $\beta^I > 1$ and therefore to the overprovision of the intermediary's effort $a^I = \frac{\beta^I}{\lambda}$ above the first-best level $\frac{1}{\lambda}$. To curb excessive incentives β^I in the presence of trickle-up incentives, the intermediary reduces his direct exposure to cash flow shocks $1 - \beta^P$ by sharing risk with the principal and setting $\beta^P > 0$, remarkably, even though there is no risk-sharing motive (see the middle right panel).[13]

Note that for larger values of v^A with $F'(v^A) < 0$, the trickle-up effect weakens the intermediary's incentives and leads to dis-incentives, reflecting an *agency overhang* problem. The intuition is that when v^A is large, the agent possesses a large stake in the firm. Hence, the gains generated by the intermediary's effort mostly accrue to the agent, undermining the intermediary's incentives. To maintain the intermediary's incentives in the presence of trickle-up dis-incentives, the intermediary's direct exposure to cash flow shocks $1 - \beta^P$ must increase, leading to $\beta^P < 0$ (see the middle right panel). That is, the principal has negative exposure to cash flow shocks and the intermediary's direct exposure to cash flow shocks $1 - \beta^P$ exceeds one. Thus, the principal contributes funds via $dw^P < 0$ after positive cash-flow shocks $dZ > 0$ and receives payouts $dw^P > 0$ after negative cash-flow shocks $dZ < 0$ meaning that the intermediary effectively takes leverage on his stake within the firm.

In summary, the intensity of risk-sharing reflects the trickle-up effect: the intermediary shares risk with the principal (i.e., $\beta^P > 0$) if and only if the trickle-up effect strengthens his incentives, whereas he takes leverage (i.e., $\beta^P < 0$) when the trickle-up effect weakens his incentives. The following corollary establishes analytical results regarding the effects discussed above when the intermediary's risk-aversion is small and his effort cost is large.[14]

[13]Notably, we do not need risk aversion in the model to obtain most of the results. As we show, in the optimal contract, even a risk-neutral financially-unconstrained intermediary will not take all the residual risk. The reason is that the intermediary faces two tasks, monitoring and contracting, and risk-sharing with the principal can, in general, improve the intermediary's incentives for these tasks. The agent's risk-aversion is also not necessary. We present the simplest risk-neutral version of the model in an appendix.

[14]Appendix 4.13.3 presents additional numerical comparative statics results with respect to the agent's and the intermediary's costs of effort.

Corollary 6 *Let $v^* > 0$ the unique value with $F'(v^*) = 0$ and suppose that $1/\lambda > 0$ and $\theta^I > 0$ are sufficiently small. Then, the following holds:*

1. *In the limit, as $\frac{1}{\lambda} \to 0$ and $\theta^I \to 0$, it holds that $\beta^A = \pi^I$, $\beta^I = 1 + F'(v^A)\pi^I$, and $\beta^P = 0$.*
2. *The intermediary's incentives and effort satisfy $\beta^I > 1$ and $a^I > a^I_{FB} = \frac{1}{\lambda}$ for $v^A < v^*$ and $\beta^I < 1$ and $a^I < a^I_{FB} = \frac{1}{\lambda}$ for $v^A > v^*$.*

Implications for private equity

In the private equity example, the intermediary represents the general partners, the principal the limited partners, and the agent the manager of a portfolio firm. The state variable, the value of the agent's deferred compensation v_A, more broadly proxies for firm performance and financial health as it increases after positive cash flow shocks and decreases after negative shocks. When v_A is low, the firm is threatened by liquidation. Accordingly, v_A can be interpreted as a measure of financial slack of the portfolio firm. With this interpretation at hand, our model has the following implications.

First, the manager's incentives and the general partner's monetary incentives increase in firm performance. Specifically, both β^A and $1 - \beta^P$ increase in v_A. If the sensitivity of compensation to performance increases after good performance, incentives are convex or option-like. This prediction is consistent with the high-powered incentive schemes frequently adopted in practice. Private equity compensation contracts commonly feature carry hurdle provisions, which generate convexity of incentives for general partners (Metrick and Yasuda (2010)).

Second, the general partners' actual incentives β^I (including monetary and indirect incentives) are the highest in distress due to the trickle-up effect. In other words, the general partners benefit most from saving the firm from liquidation and this generates endogenous incentives without direct exposure to cash flows. Thus, because the general partners' value is concave in firm performance, their actual incentives are no longer convex despite the convex exposure to cash flows.

Third, the general partners' effort is the highest in distress, whereas the manager's effort is the highest after good performance. Recall that optimal efforts are proportional to β^I and β^A, respectively, so this prediction is implied by the previous two points. This pattern of efforts means that the optimal contract generates incentives for the intermediary to step in when the manager's role is diminished in distress and to step back when the manager's role increases after good performance.

4.4 Observable contract terms and direct contracting

In this section, we consider that the contract terms Π^A between the intermediary and the agent are publicly observable and contractible between the principal and the intermediary. Thus, the principal can directly influence the agent's contract Π^A via the contract Π^P she signs with the intermediary. In other words, observable and contractible contract terms Π^A reflect increased participation of investors (i.e., the principal) in the firm's governance. In practice, increased investor participation can represent various forms of investor activism. In the context of private equity, this can arise when limited partners directly co-invest in portfolio firms outside the fund structure. The trend towards increased co-investment, documented in Fang, Ivashina, and Lerner (2015), is viewed as evidence of disintermediation in private equity investment.

The specification of our model with observable and contractible contract terms Π^A can also be called *direct contracting* because an equivalent formulation is that the principal *directly* offers both a contract Π^A to the agent and a contract Π^I to the intermediary. Likewise, our baseline specification can be called *delegated contracting* because the principal contracts only with the intermediary, while contracting with the agent is *delegated* to the intermediary.

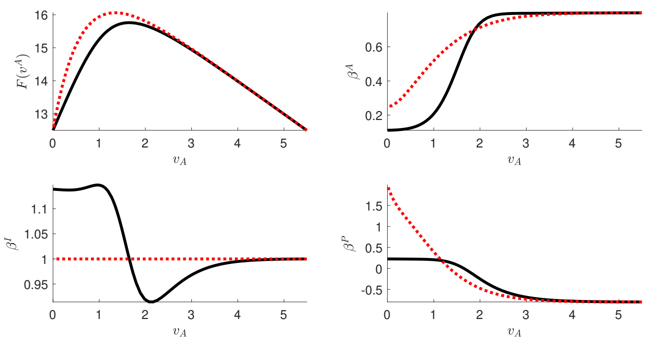

FIGURE 4.2: Model solution with direct contracting. The dashed red lines represent the model with direct contracting, while the solid back lines represent the baseline model with delegated contracting. The payout boundaries under delegated and direct contracting are quantitatively similar, with $\overline{v} \approx 5.5$.

4.4.1 Solution

When the agent's contract Π^A is observable and contractible, the intermediary's value function—which we denote for simplicity, but with a slight abuse of notation, also by $F(v^A)$—solves the HJB equation (4.28) too. However, as the agent's incentives β^A and her prescribed effort $\hat{a}^A$ are observable and contractible, the agent's and the intermediary's incentives β^A and β^I are no more linked via the constraint (4.27) and, therefore, can be chosen independently to maximize $F(v^A)$ subject to the effort incentive constraints (4.12). Thus, the optimal incentives are obtained by solving the first-order conditions $\frac{\partial F(v^A)}{\partial \beta^A} = 0$ and $\frac{\partial F(v^A)}{\partial \beta^I} = 0$ (or equivalently $\frac{\partial F(v^A)}{\partial \beta^P} = 0$) taking into account (4.12). This leads to

$$\beta^A = \frac{1}{-(\delta\theta^A r\sigma^2 + 1)F'(v^A) - \delta\sigma^2 F''(v^A)} \tag{4.30}$$

and

$$\beta^I = \frac{1}{1 + \lambda\theta^I r\sigma^2}. \tag{4.31}$$

Combining (4.23) and (4.31), we are able to characterize optimal risk-sharing with the principal and as such the intermediary's direct exposure to cash flow risk

$$1 - \beta^P = \frac{1}{1 + \lambda\theta^I r\sigma^2} - F'(v^A)\beta^A \tag{4.32}$$

under direct contracting. Note that while the intermediary's total incentives β^I are constant under direct contracting, the agent's incentives β^A and the intermediary's risk-exposure $1 - \beta^P$ are state-dependent.

Figure 4.2 plots the agent's incentives β^A, the intermediary's incentives β^I, and the risk-sharing intensity β^P against v^A both in the baseline version (solid black line) and under direct contracting (dotted red line). Because β^A and β^I are no more linked via (4.27), incentives no more trickle down from the intermediary to the agent. However, incentives continue to trickle up from the agent to the intermediary. Because it is optimal to implement a constant level of incentives β^I and effort a^I, risk-sharing between the principal and the intermediary (i.e., β^P) exactly offsets the trickle-up effect. Note that the intermediary's incentives become stronger (weaker) relative to the baseline when v^A is low (high), while the agent's incentives become stronger relative to the baseline for low values of v^A.

In the context of private equity, the model implies that the effect of increased participation of limited partners in firm governance will depend on past performance and the state of the firm. When the target firm is in financial distress, increased participation of limited partners reduces general partners' incentives, transfers risk to limited partners, and increases the incentives of the portfolio firm's manager. In contrast, when the target firm is financially sound after strong performance, increased participation of limited partners has little effects on risk-sharing and the manager's incentives, yet it may increase general partners' incentives. We conclude with the following proposition formalizing the results discussed in this section.

Proposition 16 *When the contract Π^A is publicly observable and contractible, the following holds:*

1. *The value of the intermediary's deferred payouts $F(v^A)$ solves* (4.28) *subject to the boundary conditions $F(0) - R = F'(\overline{v}) - 1 = F(\overline{v}) = R$. Payouts satisfy $dw^A = \max\{v^A - \overline{v}, 0\}$ and $dw^P = \beta^P \sigma dZ$, leading to $v^P = 0$. The function F is strictly concave.*
2. *The sensitivities β^A, β^P and β^I are characterized by* (4.23), (4.30) *and* (4.31). *Effort is characterized by* (4.12). *The agent's incentives and effort are increasing for $v^A \geq v^*$ and the intermediary's incentives and effort are constant in v^A.*
3. *Suppose that θ^I and $\frac{1}{\lambda}$ are sufficiently small. When $v^A < v^*$ ($v^A > v^*$) the intermediary's incentives and effort are lower (higher) relative to the base case scenario with unobservable Π^A.*

4.4.2 Direct contracting versus delegated contracting

In this section, we analyze when increased investor participation adds value or, equivalently, under what circumstances direct contracting is preferred over delegated contracting.

Note that direct contracting creates value above delegated contracting by facilitating more efficient incentive provision to both the agent and the intermediary. While it is clear that the firm value under direct contracting is always higher than the firm value under delegated contracting, in practice, there are costs associated with increased investor participation (e.g., costs of attention, costs of information acquisition and processing). In the following, we study the additional value generated from direct contracting above delegated contracting, denoted by Δ, and observe that when Δ is low, delegated contracting can be preferred over direct contracting.

To isolate the effects, we first study the limit $\sigma \to 0$. As $\sigma \to 0$, the firm's liquidation risk vanishes, so it is optimal to set $\overline{v} \to 0$. Then v_t^A is at $\overline{v}$ with probability one and the model solution converges to become time-stationary. The agent receives payouts continuously and $F'(v^A) = -1$, implying that the trickle-up effect leads to dis-incentives for the intermediary.

Under delegated contracting, we take the limit $\sigma \to 0$ in the expressions for π^I and β^I evaluated at $v^A = \overline{v} = 0$ and obtain the closed-form expressions:

$$\pi^I = \frac{\lambda}{\delta + \lambda}, \quad \beta^I = \frac{\delta^3 + \lambda^2\delta}{\delta^3 + \lambda^3}, \quad \text{and} \quad \beta^A = \frac{\lambda^3 + \delta^2\lambda}{\delta^3 + \lambda^3}. \tag{4.33}$$

In addition, the incentive constraints (4.12) apply, in that $a^A = \frac{\beta^A}{\delta}$ and $a^I = \frac{\beta^I}{\lambda}$. Recall that first-best efforts are given by $a^A_{FB} = \frac{1}{\delta}$ and $a^I_{FB} = \frac{1}{\lambda}$, so that effort a^j is at its first-best level if and only if $\beta^j = 1$. Under delegated contracting, the optimal incentives β^j are generally not equal to one and so the optimal efforts a^j are generally not equal to the first-best efforts for $j = A, I$. In contrast, under direct contracting, the limit $\sigma \to 0$ yields $\beta^j \to 1$ and $a^j \to a^j_{FB}$ for $j = A, I$, hence the first-best outcome is attained.

Observe that $\lambda = \delta$ implies $\beta^A = \beta^I = 1$ in (4.33) and the first-best outcome is also attained under delegated contracting. Clearly, the distortion in effort levels under delegated contracting is related to the differences in costs λ and δ. When the intermediary's effort cost λ exceeds the agent's effort cost δ, then $\beta^I < 1 < \beta^A$, leading to under-provision of the intermediary's effort (i.e., $a^I = \frac{\beta^I}{\lambda} < \frac{1}{\lambda} = a^I_{FB}$) and over-provision of the agent's effort (i.e., $a^A = \frac{\beta^A}{\delta} > \frac{1}{\delta} = a^A_{FB}$). Evidently, because the intermediary's incentives are linked to the

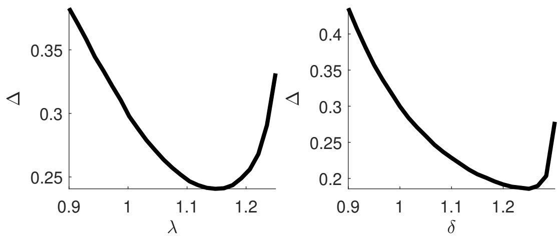

FIGURE 4.3: Value of direct versus delegated contracting. On the y-axis, Δ denotes the additional value direct contracting generates above delegated contracting. On the x-axis, λ and δ are the coefficients of the effort cost functions of the intermediary and the agent, respectively.

agent's incentives, the intermediary loads most of the incentives on the more cost-efficient party. The reason for this is that when λ is large relative to δ, then the shadow cost of the trickle-up constraint is low (see term D in (4.25), which equals $\frac{\delta}{\lambda}$ as $\sigma \to 0$), which leads to high π^I and high trickle-down incentives. Symmetrically, when $\lambda < \delta$, trickle-down incentives are weak and $\beta^I > 1 > \beta^A$, inducing over-provision of the intermediary's effort (i.e., $a^I > a^I_{FB}$) and under-provision of the agent's effort (i.e., $a^A < a^A_{FB}$).

In sum, incentive provision under delegated contracting becomes less efficient when the difference between δ and λ increases. The intuition is related to the fact that under delegated contracting, the intermediary has two tasks, own effort and incentive provision to the agent. As the two incentive problems are connected in this way, when λ (or δ) is low and δ (respectively, λ) is large, severe moral hazard at the firm (respectively, intermediary) level propagates to the intermediary (respectively, firm) level and so leads to distortions in incentive provision. Under these circumstances, direct contracting becomes valuable as it eliminates the adverse propagation of agency problems.

Having gained this intuition, we now move to the general model with σ away from zero. Figure 4.3 plots the additional value of direct contracting Δ against varying levels of λ and δ. The two panels show that Δ is U-shaped in δ and λ. As shown above, delegated contracting becomes harder, and the value of direct contracting rises when the difference between λ and δ is large, leading to the U-shaped pattern of Δ in both λ and δ.

4.4.3 Dynamic delegation

Investor activism, including engagements in corporate governance, is often intermittent rather than constant and continuous (Dimson, Karakaş, and Li, 2015). If investors can sporadically increase their participation in firm governance, when would they optimally do it? To answer this question within our model, this section allows the intermediary to temporally commit to prescribed effort $\hat{a}^A_t$ and to make $\hat{a}^A_t$ observable and contractible with the principal. This is costly at flow cost $\eta \geq 0$. Thus, η captures the intermediary's cost of transparency or commitment towards the principal. Note that it is immaterial whether the principal or the intermediary bears the cost η, as both can split this cost via the contract Π^P in any arbitrary way. As such, an alternative interpretation is that by paying the cost η, the principal can temporarily affect the agent's contract terms directly. Therefore, we say that there is direct contracting at time t when the flow cost η is expended, and there is delegated contracting otherwise. The solution details of this model extension are relegated to Appendix 4.13.1.

In what follows, we denote the value function in this model extension by $F(v^A)$. As in the baseline version of the model, incentives β^A and β^I are chosen to maximize the sum of expected cash flows $a^A + a^I$ and the expected change in value $\mathbb{E}dF(v^A)/dt$, adjusted for risk

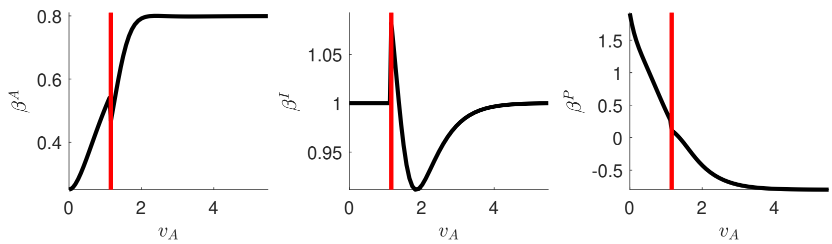

FIGURE 4.4: Dynamic delegation. We use our baseline parameters and set $\eta = 0.02$. For small values of v^A to the left of the solid red line, the principal contracts directly with the agent. For large values of v^A to the right of the solid red line, the principal delegates contracting.

and effort costs. This leads to the maximization problem

$$\max_{\beta^A,\beta^I} \left\{ a^A + a^I - \frac{\lambda(a^I)^2}{2} - \frac{\theta^I r(\beta^I \sigma)^2}{2} + \frac{\mathbb{E}dF(v^A)}{dt} \right\} \tag{4.34}$$

that is subject to the effort incentive constraints (4.12) and, additionally, to constraint (4.27) if and only if the agent's prescribed effort $\hat{a}^A$ is not observable to the principal and not contractible. Hence, incurring the flow cost η makes the agent's prescribed effort and incentives observable to the principal and so removes the constraint (4.27) from the instantaneous maximization problem (4.34).

Direct contracting is optimal if and only if $B^*(v^A) - \eta \geq A^*(v^A)$ where

$$A^*(v^A) \equiv \max_{\beta^A,\beta^I} \left\{ a^A + a^I - \frac{\lambda(a^I)^2}{2} - \frac{\theta^I r(\beta^I \sigma)^2}{2} + \frac{\mathbb{E}dF(v^A)}{dt} \right\} \text{ s.t. (4.12) and (4.27)}$$

$$B^*(v^A) \equiv \max_{\beta^A,\beta^I} \left\{ a^A + a^I - \frac{\lambda(a^I)^2}{2} - \frac{\theta^I r(\beta^I \sigma)^2}{2} + \frac{\mathbb{E}dF(v^A)}{dt} \right\} \text{ s.t. (4.12).}$$

Because direct contracting facilitates efficient incentive provision to both the agent and the intermediary, it is preferred to delegated contracting exactly when the trickle-up and trickle-down effects prevent efficient incentive provision under delegated contracting. Figure 4.1 illustrates that efficient incentive provision is especially hard for low values of v^A, as strong trickle-up incentives lead to over-provision of incentives to the intermediary and weak trickle-down incentives curb the agent's incentives. Under these circumstances, direct contracting becomes optimal as it can mitigate the over-provision (under-provision) of incentives to the intermediary (agent). This pattern is illustrated in Figure 4.4, showing that direct contracting is optimal for low values of v^A (to the left of the solid red line) and delegated contracting is optimal for high values of v^A (to the right of the solid red line). As direct contracting reduces the intermediary's incentives and increases the agent's incentives for low values of v^A, the agent's (intermediary's) incentives β^A (β^I) exhibit a downward (upward) jump, once v^A crosses the solid red line and the principal stops to directly influence firm governance (i.e., the contract Π^A).

In sum, we find that direct contracting in dynamic settings becomes more valuable after poor past performance (i.e., when moral hazard exacerbates), while the delegation of contracting becomes more valuable after good past performance (i.e., when moral hazard becomes less severe).

4.5 Alternative model applications and say-on-pay

Our baseline model features a financially unconstrained, albeit possibly risk averse, intermediary. This set of assumptions is consistent with the application to private equity in that general partners raise additional funds from limited partners (instead of using their own funds or borrowing) to diversify risk but, if required, can contribute more financing. In other applications, intermediaries may be more financially constrained, without deep pockets, and unwilling or unable to inject funds after poor performance. In one such important application, the intermediary represents a board of directors. In their traditional roles, shareholders delegate to boards both monitoring of and contracting with firm managers, as in our baseline model of delegated contracting.

To show that the model can accommodate financially constrained intermediaries and be consistent with the board of directors application, we solve the model under the additional constraint that the cash payments $dw_t^I := dX_t - dw_t^P - dw_t^A$ the intermediary collects over $[t, t+dt)$ must be positive (i.e., $dw_t^I \geq 0$ for all $t > 0$), while the intermediary is able to inject funds only at the initial round of financing at time $t = 0$ (i.e., $dw_0^I < 0$ is possible). Under these circumstances, the intermediary does not inject funds into the firm out of his own pocket but rather intermediates funds from the principal to the firm and the agent. To be able to solve the model with the constraint dw_t^I, we require a slight change to our limited liability assumption. In particular, we stipulate that the joint continuation value of the intermediary and the principal, that is, $F(v^A)$, must exceed the recovery value of the firm R.

The solution details are presented in Appendix 4.13.4. Notably, the model solution remains mostly unchanged compared with the baseline. In particular, the agent's payouts satisfy $dw^A = \max\{v^A - \overline{v}, 0\}$, and the value function $F(v^A)$ solves the HJB equation equation (4.28) subject to the boundary conditions $F(0) = F(\overline{v}) = R$ and $F'(\overline{v}) = -1$. As a result, optimal incentives to the agent and the intermediary and optimal efforts are the same as in the baseline. That is, the model implications regarding the dynamics of incentives and risk-sharing discussed in the previous sections remain unchanged compared with the baseline.

When the intermediary is interpreted as the board of directors, we find that directors' incentives and monitoring activity (as captured by effort a^I) are high when the firm undergoes financial distress and are low after good past performance (see Figure 4.1). Interestingly, the board of directors application also allows a clear interpretation of the model setup with observable and contractible contract terms Π^A. Adoption of various say-on-pay regulations altered these traditional roles and brought about an increase in shareholders' direct participation in arranging executive compensation. These regulations shifted the shareholders-board-manager relationship towards the alternative setting of our model studied Section 4.4 in which shareholders determine the manager's compensation directly.

The model has several implications with respect to say-on-pay regulations. First, say-on-pay regulations increase managerial incentives but decrease board members' incentives and monitoring activity under financial distress, whereas they increase board members' incentives and monitoring activity for firms with sufficient financial slack (Section 4.4.1 and Figure 4.2). Overall, say-on-pay regulations reduce the variation in directors' incentives and monitoring. Second, suggests that say-on-pay regulations are particularly useful for firm under financial distress (Section 4.4.3 and Figure 4.4). Third, say-on-pay regulations are valuable when agency conflicts at the manager level are sufficiently different (larger or smaller) than those at the board level (Section 4.4.2 and Figure 4.3).

4.6 Conclusion

We study a dynamic agency model in which investors finance a firm run by an agent and monitored by an intermediary. Both the agent and the intermediary are subject to moral hazard. Because contracting is delegated to the intermediary, the intermediary receives incentives to both monitor and contract with the agent. The incentive problems of the agent and the intermediary interact with each other and distinctly depend on the position in the investment structure. The intermediary's incentives are partially passed through to the agent. The rate at which the intermediary is willing to pass through the incentives depends on past performance and is the weakest after poor performance. When the agent's incentives are strong after good performance, they create an agency overhang problem, which decreases the intermediary's incentives to monitor. Conversely, after poor past performance, the intermediary is willing to exert a high effort to save the firm from agency-induced distress. Taken together, the optimal compensation schemes of the agent and the intermediary are convex in firm performance. However, the total incentives of the intermediary are not convex, and the intermediary is mostly strongly incentivized in distress. We show that increased investor participation in firm governance can be particularly valuable to prevent agency conflicts from spreading from the agent to the intermediary (or vice versa) or after poor performance when the interaction of the incentive problems leads to the largest distortions.

Appendix

4.7 Preliminaries

Uncertainty is modeled via the complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ that satisfies the usual conditions and is equipped with the filtration $\mathbb{F} := (\mathcal{F}_t)_{t\geq 0}$. Here, $\mathcal{F}_t = \sigma(X_s : s \leq t)$ is the public information that is available at time t to all players, as all players observe cash flow realizations dX_t. In the following, we work with four different probability measures, denoted $\mathbb{P}$ and $\mathbb{P}^k$ for $k = A, I, P$. The expectation $\mathbb{E}_t$ is taken under the probability measure $\mathbb{P}$ conditional on time-t information. The expectation $\mathbb{E}_t^k$ is taken under the probability measure $\mathbb{P}^k$ conditional on time-t information. Let us discuss the four probability measures in more detail.

1. The measure $\mathbb{P}^A$ is induced by the efforts a^A and $\hat{a}^I$, so that $dZ_t^A = \frac{dX_t - a_t^A dt - \hat{a}_t^I dt}{\sigma}$ is the increment of a standard Brownian Motion under the measure $\mathbb{P}^A$. Note that the agent observes her own effort a_t^A and the prescribed effort $\hat{a}_t^I$ via the contract Π^A, so the expectation $\mathbb{E}_t^A$ is taken under (i.e., conditional on) the agent's information.

2. The measure $\mathbb{P}^I$ is induced by the efforts $\hat{a}^A$ and a^I, so that $dZ_t^I = \frac{dX_t - \hat{a}_t^A dt - a_t^I dt}{\sigma}$ is the increment of a standard Brownian Motion under the measure $\mathbb{P}^I$. Note that the intermediary observes his own effort a_t^I and the prescribed effort $\hat{a}_t^A$ via the contract Π^A, so the expectation $\mathbb{E}_t^I$ is taken under (i.e., conditional on) on the intermediary's information.

3. The measure $\mathbb{P}^P$ is induced by the efforts $\bar{a}^A$ and $\bar{a}^I$, so that $dZ_t^P = \frac{dX_t - \bar{a}_t^A dt - \bar{a}_t^I dt}{\sigma}$ is the increment of a standard Brownian Motion under the measure $\mathbb{P}^P$. Note that the principal observes the prescribed efforts $\bar{a}_t^A, \bar{a}_t^I$ via the contract Π^P, so the expectation $\mathbb{E}_t^P$ is taken under (i.e., conditional on) the principal's information.

4. The measure $\mathbb{P}$ is induced by the efforts $\hat{a}^A$ and $\hat{a}^I$, so that $d\hat{Z}_t = \frac{dX_t - \hat{a}_t^A dt - \hat{a}_t^I dt}{\sigma}$ is the increment of a standard Brownian Motion under the measure $\mathbb{P}$. Note that the contract Π^A stipulates prescribed efforts $\hat{a}_t^j$ for $j = A, I$, so the expectation $\mathbb{E}_t$ is taken under (i.e., conditional on) public information $\mathcal{F}_t$ and the contract terms Π^A.

We focus on incentive-compatible contracts Π^A and Π^P, in that (in optimum) $a^j = \bar{a}^j = \hat{a}^j$ for $j = A, I$ and $\mathbb{P}$ is the probability measure induced by the efforts a^A and a^I. Note that in optimum, $\mathbb{P}$ coincides with $\mathbb{P}^k$, as prescribed and actual effort levels coincide. In the main text, we do not formally distinguish between prescribed and actual effort levels and carry out the arguments using the equilibrium probability measure $\mathbb{P}$. That is, in the main text, we do not formally distinguish between the different probability measures $(\mathbb{P}, \mathbb{P}^k)$.

Throughout the paper and for all problems, we stipulate that the continuation utility of player $j \in \{A, I\}$, denoted U_t^j and defined in (4.7), and the sensitivities β_t^k for $k \in \{A, I, P\}$, implicitly defined in (4.8) and (4.10), are bounded. That is, $|U_t^j| < M$ and $|\beta_t^k| < M$ for all $t \geq 0$ and $M > 0$. This is merely a regularity condition, used in the verification argument, and, in fact, we can pick $M < \infty$ sufficiently large to ensure that this constraint never binds in optimum.

4.8 Proof of Lemma 20

4.8.1 Part I — Martingale representation

Take player $j's$ continuation utility

$$U_t^j = \mathbb{E}_t \left[\int_t^\infty e^{-r(s-t)} u^j(c_s^j, a_s^j) ds \right],$$

evaluated under the measure $\mathbb{P}$, so the expectation $\mathbb{E}_t$ is taken under (i.e., conditional on) public information $\mathcal{F}_t$ and the contract terms Π^A. Define

$$A_t^j = \mathbb{E}_t \left[\int_0^\infty e^{-rs} u^j(c_s^j, a_s^j) ds \right] = \int_0^t e^{-rs} u^j(c_s^j, a_s^j) ds + e^{-rt} U_t^j. \tag{4.35}$$

By construction, $A^j = \{A_t^j\}$ is a martingale under the probability measure $\mathbb{P}$. By the martingale representation theorem, there exists a stochastic process $\beta^j = \{\beta_t^j\}$ such that

$$e^{rt} dA_t^j = (-\theta^j r U_t^j) \beta_t^j (dX_t - \hat{a}_t^A dt - \hat{a}_t^I dt), \tag{4.36}$$

where $d\hat{Z}_t = \frac{dX_t - \hat{a}_t^A dt - \hat{a}_t^I dt}{\sigma}$ is the increment of a standard Brownian Motion under the probability measure $\mathbb{P}$.

We differentiate (4.35) with respect to time t to obtain an expression for dA_t^j, then plug this expression into (4.36) and solve (4.36) to get

$$dU_t^j = rU_t^j dt - u^j(c_t^j, a_t^j) dt + (-\theta^j r U_t^j) \beta_t^j (dX_t - \hat{a}_t^A dt - \hat{a}_t^I dt), \tag{4.37}$$

which is (4.8).

Next, we consider the principal's continuation payoff

$$v_t^P := \mathbb{E}_t^P \left[\int_t^\infty e^{-r(s-t)} dw_s^P \right], \tag{4.38}$$

evaluated under the measure $\mathbb{P}^P$ that is taken under (i.e., conditional on) the principal's information. Define

$$A_t^P = \mathbb{E}_t^P \left[\int_0^\infty e^{-rs} dw_s^P \right] = \int_0^t e^{-rs} dw_s^P + e^{-rt} v_t^P. \tag{4.39}$$

By construction, $A^P = \{A_t^P\}$ is a martingale under the probability measure $\mathbb{P}^P$. By the martingale representation theorem, there exists a stochastic process $\beta^P = \{\beta_t^P\}$ such that

$$e^{rt} dA_t^P = \beta_t^P (dX_t - \bar{a}_t^A dt - \bar{a}_t^I dt), \tag{4.40}$$

where $dZ_t^P = \frac{dX_t - \bar{a}_t^A dt - \bar{a}_t^I dt}{\sigma}$ is the increment of a standard Brownian motion under the probability measure $\mathbb{P}^P$. We differentiate (4.39) with respect to time t to obtain an expression for dA_t^P, then plug this expression into (4.40) and solve (4.40) to get

$$dv_t^P = rv_t^P dt - dw_t^P + \beta_t^P (dX_t - \bar{a}_t^A dt - \bar{a}_t^I dt),$$

which is (4.10).

4.8.2 Part II — Optimal consumption and effort

Define $\chi^A = \delta$ and $\chi^I = \lambda$. We characterize player j's optimal consumption c_t^j and optimal effort a_t^j. To do so, we verify that proposed effort from Lemma 20, which is denoted $\hat{a}_t^j$ and equal to $\hat{a}_t^j = \beta_t^j / \chi^j$, and proposed consumption from Lemma 20, which is denoted $\hat{c}_t^j$ and satisfies $u^j(\hat{c}^j_t, \hat{a}^j_t) = U^j_t$, yield higher payoff than any other strategy (c^j_t, a^j_t) and therefore are indeed optimal.[15] As a result, we obtain that optimal effort is $a^j_t = \hat{a}^j_t = \beta^j_t/\chi^j$ and optimal consumption satisfies $u^j(c^j_t, a^j_t) = rU^j_t$, as stated by Lemma 20.

For $t < \tau$, define player j's continuation utility

$$U^j_t = \mathbb{E}_t\left[\int_t^\infty e^{-r(s-t)} u^j(\hat{c}^j_s, \hat{a}^j_s)ds\right]$$

under the proposed strategy, when player j exerts effort $\hat{a}^j_s = \beta^j_s/\chi^j$ and consumes according to $u^j(\hat{c}^j_s, \hat{a}^j_s) = rU^j_s$ for $s \geq t$. Under the proposed strategy, player j accumulates savings $\hat{S}^j_s$ according to

$$d\hat{S}^j_s = r\hat{S}^j_s ds + d\hat{w}^j_s - \hat{c}^j_s ds \quad \text{and} \quad \lim_{s\to\infty} \mathbb{E}_0 e^{-rs}\hat{S}^j_s = 0.$$

The expectation $\mathbb{E}_t$ is taken under the probability measure $\mathbb{P}$ induced by the proposed efforts $\hat{a}^A, \hat{a}^I$, so that under this probability measure, $\frac{dX_s - \hat{a}^A_s dt - \hat{a}^I_s dt}{\sigma}$ is the increment of a standard Brownian Motion.

Take a time $t < \tau$. Suppose that player j deviates and follows an alternative strategy up to time t. That is, player j exerts effort a^j_s and chooses consumption c^j_s, for times $s < t$, where the strategy (c^j_s, a^j_s) is (with a slight abuse of notation) not necessarily optimal. Under the strategy (c^j_s, a^j_s), player j accumulates savings S^j_s according to

$$dS^j_s = rS^j_s ds + dw^j_s - c^j_s ds \quad \text{and} \quad \lim_{s\to\infty} \mathbb{E}^j_0 e^{-rs} S^j_s = 0.$$

Recall that the expectation $\mathbb{E}^j_t$ is taken under probability measure $\mathbb{P}^j$ (which is induced by efforts $(a^A, \hat{a}^I)$, if $j = A$, and by efforts $(\hat{a}^A, a^I)$ if $j = I$).

Define

$$\Delta^j_t = \int_0^t e^{r(t-s)}(\hat{c}^j_s - c^j_s)ds,$$

which is the amount of dollars that player j saves by deviating from the proposed consumption policy $\hat{c}^j_s$ for times $s \leq t$. Note that

$$d\Delta^j_t = r\Delta^j_t dt + (\hat{c}^j_t - c^j_t)dt, \tag{4.41}$$

with $\Delta^j_0 = 0$. The fact that $\lim_{s\to\infty} \mathbb{E}_0 e^{-rs}\hat{S}^j_s = 0$ and $\lim_{s\to\infty} \mathbb{E}^j_0 e^{-rs} S^j_s = 0$ hold implies the transversality condition

$$\lim_{s\to\infty} \mathbb{E}^j_0 e^{-rs}\Delta^j_t = \lim_{s\to\infty} \mathbb{E}_0 e^{-rs}\Delta^j_s = 0. \tag{4.42}$$

We stipulate that for times $s \geq t$, player j exerts effort $\hat{a}^j_s$ and chooses consumption $\hat{c}^j_s + r\Delta^j_t$. This consumption policy is feasible in that it does not lead to a violation of the transversality condition for player j's savings, because Δ^j_s satisfies the transversality condition (4.42) and proposed consumption $\hat{c}^j_s$ is feasible.

As a result, player $j's$ continuation utility from time t onward is

$$\begin{aligned}\mathbb{E}^j_t\left[\int_t^\infty e^{-r(s-t)}u^j(\hat{c}^j_s + r\Delta^j_t, \hat{a}^j_s)ds\right] &= e^{-\theta^j r\Delta^j_t}\mathbb{E}^j_t\left[\int_t^\infty e^{-r(s-t)}u^j(\hat{c}^j_s, \hat{a}^j_s)ds\right]\\ =&e^{-\theta^j r\Delta^j_t}\mathbb{E}_t\left[\int_t^\infty e^{-r(s-t)}u^j(\hat{c}^j_s, \hat{a}^j_s)ds\right] = e^{-\theta^j r\Delta^j_t}U^j_t,\end{aligned}$$

[15] As we focus on incentive compatible contracts, proposed effort $\hat{a}^j_t$ must be optimal for player j (at time t), i.e., incentive compatible.

where it was used that (by direct calculation)

$$u^j(\hat{c}^j_s + r\Delta^j_t, \hat{a}^j_s) = e^{-\theta^j r\Delta^j_t} u^j(\hat{c}^j_s, \hat{a}^j_s).$$

Note that because player j exerts prescribed effort $\hat{a}^j_s$ from time t onward, the two expectation operators $\mathbb{E}_t$ and $\mathbb{E}^j_t$ coincide. To show that the proposed strategy is optimal, we demonstrate that following the proposed strategy from time $t = 0$ onward is optimal, in which case $\Delta^j_s = 0$, effort is $\hat{a}^j_s$, and consumption is $\hat{c}^j_s$ for all $s \geq 0$.

Player j's payoff at time zero (under the deviation) can then be written as

$$G^j_t = \int_0^t e^{-rs} u^j(c^j_s, a^j_s)ds + e^{-rt} e^{-\theta^j r\Delta^j_t} U^j_t. \tag{4.43}$$

Using (4.8) (with c^j_t, a^j_t replaced by $\hat{c}^j_t, \hat{a}^j_t$), we differentiate (4.43) w.r.t. time, t, to get

$$\begin{aligned}
&e^{rt} e^{\theta^j r\Delta^j_t} dG^j_t = \\
&= \Big\{ \underbrace{e^{\theta^j r\Delta^j_t} u(c^j_t, a^j_t) - u(\hat{c}^j_t, \hat{a}^j_t) + (-\theta^j r U^j_t)[r\Delta^j_t + (\hat{c}^j_t - c^j_t)] + (-\theta^j r U^j_t)\beta^j_t(a^j_t - \hat{a}^j_t)}_{=\mu^j_t} \Big\} dt \\
&\quad + (-\theta^j r U^j_t)\beta^j_t(dX_t - a^j_t dt - \hat{a}^{-j}_t dt) \\
&=: \mu^j_t dt + (-\theta^j r U^j_t)(dX_t - a^j_t dt - \hat{a}^{-j}_t dt),
\end{aligned}$$

where $a^{-j}_t = a^A_t$, if $j = I$, and $a^{-j}_t = a^I_t$, if $j = A$. Observe that $\frac{dX_t - a^j_t dt - \hat{a}^{-j}_t dt}{\sigma}$ is the increment of a standard Brownian motion under the probability measure $\mathbb{P}^j$ that is taken under player j's information and is induced by efforts a^j and $\hat{a}^{-j}$. Thus,

$$\mathbb{E}^j_0 \left[\int_0^t e^{-rs} e^{-\theta r\Delta^j_s} (-\theta^j r U^j_s)\beta^j_s(dX_s - a^j_s ds - \hat{a}^{-j}_s ds) \right] = 0, \tag{4.44}$$

since the sensitivity β^j_s and continuation utility U^j_s are (by assumption) bounded.

As a next step, we solve the problem

$$\max_{a^j_t \leq A, c^j_t} \mu^j_t. \tag{4.45}$$

Given an interior solution to (4.45) (with $a^j_t < A$), the solution to (4.45) must satisfy the first order conditions $\frac{\partial}{\partial a^j_t}\mu^j_t = \frac{\partial}{\partial c^j_t}\mu^j_t = 0$. Note that $\frac{\partial}{\partial c^j_t}\mu^j_t = 0$ is equivalent to

$$\frac{\partial}{\partial c^j_t} e^{\theta^j r\Delta^j_t} u^j(c^j_t, a^j_t) = -\theta^j r U^j_t \iff \frac{\partial}{\partial c^j_t} u^j(c^j_t - r\Delta^j_t, a^j_t) = -\theta^j r U^j_t,$$

which can be rewritten as

$$u^j(c^j_t - r\Delta^j_t, a^j_t) = rU^j_t, \tag{4.46}$$

since $\frac{\partial}{\partial c} u^j(c,a) = -\theta^j u^j(c,a)$. Due to concave utility, the second order condition is

$$\frac{\partial^2}{\partial (c^j_t)^2} e^{\theta^j r\Delta^j_t} u^j(c^j_t, a^j_t) < 0.$$

Next, note that $\frac{\partial}{\partial a^j_t}\mu^j_t = 0$ is equivalent to

$$\frac{\partial}{\partial a^j_t} e^{\theta^j r\Delta^j_t} u^j(c^j_t, a^j_t) = -(-\theta^j r U^j_t)\beta^j_t \iff \frac{\partial}{\partial a^j_t} u^j(c^j_t - r\Delta^j_t, a^j_t) = -(-\theta^j r U^j_t)\beta^j_t,$$

which can be rewritten as
$$u^j(c_t^j - r\Delta_t^j, a_t^j)\chi^j a_t^j = rU_t^j \beta_t^j. \tag{4.47}$$
Note that the second order condition is
$$\frac{\partial^2}{\partial (a_t^j)^2} e^{\theta^j r \Delta_t^j} u^j(c_t^j, a_t^j) < 0.$$
As the second order conditions are satisfied, the first order conditions (4.46) and (4.47) are sufficient to solve the maximization problem (4.45). We combine (4.46) and (4.47) to get that $a_t^j = \beta_t^j / \chi_t^j = \hat{a}_t^j$ maximizes μ_t^j and solves (4.45). In addition, recall $u^j(\hat{c}_t^j, \hat{a}_t^j) = rU_t^j$, which — combined with (4.46) and $a_t^j = \hat{a}_t^j$ — yields that $c_t^j = \hat{c}_t^j + r\Delta_t^j$ maximizes μ_t^j and solves (4.45).

Evaluating μ_t^j under $a_t^j = \hat{a}_t^j$ and $c_t^j = \hat{c}_t^j + r\Delta_t^j$ and using that (by direct calculation)
$$e^{\theta^j r \Delta_t^j} u^j(c_t^j, a_t^j) = e^{\theta^j r \Delta_t^j} u^j(\hat{c}_t^j + r\Delta_t^j, \hat{a}_t^j) = u^j(\hat{c}_t^j, \hat{a}_t^j)$$
yields $\mu_t^j = 0$ so that $\max_{a_t^j, c_t^j} \mu_t^j = 0$. Thus, $\mu_t^J \leq 0$ under any choice of a_t^j and c_t^j.

Note that (4.44) and $\mu_t^j \leq 0$ imply that G_t^j, with
$$dG_t^j = e^{-rt} e^{-\theta^j r \Delta_t^j} \left(\mu_t^j dt + (-\theta^j r U_t^j) \beta_t^j (dX_t - a_t^j dt - \hat{a}_t^{-j} dt) \right),$$
is a super-martingale (i.e., decreases in expectation) under the measure $\mathbb{P}^j$. As a result,
$$U_0^j = G_0^j \geq \mathbb{E}_0^j G_t^j. \tag{4.48}$$
Observe that the transversality condition (4.42) and the fact that U_t^j is bounded imply
$$\lim_{t \to \infty} \mathbb{E}_0^j e^{-rt} e^{-\theta^j r \Delta_t^j} U_t^j = 0,$$
so that
$$\lim_{t \to \infty} \mathbb{E}_0^j G_t^j = \mathbb{E}_0^j \left[\int_0^\infty e^{-rs} u^j(c_s^j, a_s^j) ds \right]$$
is player j's expected lifetime utility from following the alternative strategy (c^j, a^j) forever. Next, we take the limit $t \to \infty$ on both sides of (4.48) to get
$$U_0^j = \mathbb{E}_0^j \left[\int_0^\infty e^{-rs} u^j(\hat{c}_s^j, \hat{a}_s^j) ds \right] = G_0^j \geq \mathbb{E}_0^j \left[\int_0^\infty e^{-rs} u^j(c_s^j, a_s^j) ds \right]. \tag{4.49}$$
Note that the proposed strategy $(\hat{a}^j, \hat{c}^j)$ (i.e., following the proposed strategy from time $t = 0$ onward) yields lifetime utility U_0^j and higher payoff than any other strategy (c^j, a^j). Therefore, the proposed strategy is optimal, and the above inequality (4.49) holds in equality if $(c_s^j, a_s^j) = (\hat{c}_s^j, \hat{a}_s^j)$ for all $s \geq 0$. Thus, in optimum, $\Delta_s^j = 0$ for all $s \geq 0$, effort is $a_s^j = \hat{a}_s^j = \beta_s^j / \chi^j$ (which is (4.12)), and consumption satisfies $u^j(c_s^j, a_s^j) = u^j(\hat{c}_s^j, \hat{a}_s^j) = rU_s^j$.

4.9 Proof of Lemma 21

The agent's monetary payoff from staying within the firm is $V_t^A = v_t^A + S_t^A$. Due to limited liability, the agent can always leave the firm with her savings S_t^A. As a consequence, the agent finds it optimal to stay within the firm if her monetary payoff from doing so, $v_t^A + S_t^A$, exceeds her monetary payoff from leaving the firm, S_t^A. This leads to the agent's limited liability constraint, $v_t^A \geq 0$.

Likewise, due to limited liability, the intermediary can always liquidate the firm, seize liquidation value R, and leave with his savings S_t^I. Staying with the firm yields monetary payoff $V_t^I = S_t^I + v_t^I$, and thus is optimal if and only $v_t^I \geq R$, i.e., if and only if the intermediary's limited liability constraint holds. Last, the principal's limited liability constraint requires that the principal earns non-negative payoff from time t onward, leading to $v_t^P \geq 0$.

4.10 Proofs of Lemma 22, Lemma 23, and Proposition 15

Lemma 22, Lemma 23, and Proposition 15 jointly describe the solution to the intermediary's problem (4.3). In this section, we characterize the solution to the intermediary's problem (4.3) and thereby prove the claims of Lemma 22, Lemma 23, and Proposition 15. The argument is split into several parts.

Part I puts structure to the intermediary's dynamic optimization problem (4.3) and argues that by the dynamic programming principle, the intermediary's objective function (i.e., value function) solves the Hamilton-Jacobi-Bellman (HJB) equation (4.51), which is a partial differential equation (PDE). Part II simplifies the dynamic optimzation problem by reducing its dimensionality. In particular, Part II shows that the intermediary's objective can be characterized as a solution to an ordinary differential equation (ODE). Part III characterizes the intermediary's incentives and payouts to the principal, establishing the claims of Lemma 22 and the relationship (4.23). Part IV characterizes the agent's optimal incentives, establishing the claims of Lemma 23. Part V characterizes optimal risk-sharing with the principal and, as such, the intermediary's optimal incentives, establishing (4.29). Part VI proves the concavity of the value function and characterizes the payout boundary. Part VII provides the formal verification argument that the contracts Π^A and Π^P from Proposition 15 are indeed optimal and solve the intermediary's problem (4.3). Part VIII characterizes optimal firm liquidation.

4.10.1 Part I

The intermediary chooses effort a^I, consumption c^I, and the contracts Π^A, Π^P to solve (4.3) and to maximize his expected lifetime utility U_0^I. Equivalently, the intermediary maximizes his payoff at time $t = 0$ in monetary terms V_0^I, because $V_0^I = -\frac{\ln(-\theta^I r U_0^I)}{\theta^I r}$ is a monotonic transformation of U_0^I. Recall that by Lemma 20, efforts a_t^A, a_t^I satisfy the incentive condition(s) (4.12) and $u(c_t^j, a_t^j) = rU_t^j$ holds, under any incentive compatible contracts Π^A, Π^P.

In what follows, we consider the intermediary's optimization monetary terms rather than utility terms. That is, to maximize V_0^I, the intermediary chooses incentive-compatible contracts Π^A, Π^P that respect the players' limited liability. Note that as $V_t^I = S_t^I + v_t^I$, the intermediary chooses at time $t = 0$ contracts Π^A, Π^P to (dynamically) maximize

$$V_0^I = S_{0^-}^I + \underbrace{\mathbb{E}_0^I \left[\int_0^\infty e^{-rt} \left(dw_t^I - \frac{1}{2}\theta^I r (\beta_t^I \sigma)^2 ds - g^I(a_t^I) dt \right) \right]}_{=v_0^I}, \tag{4.50}$$

where the integral may include lumpy (i.e., non-infinitesimal) transfers at time $t = 0$, dw_0^P and dw_0^A.[16] Maximizing (4.50) (i.e., V_0^I) is equivalent to maximizing v_0^I, as $S_{0^-}^I$ is exogenous.

The dynamic optimization of v_0^I depends on three state variables: i) the principal's continuation payoff v_t^P, ii) the agent's continuation utility U_t^A or equivalently the agent's continuation payoff in dollars V_t^A, and iii) the agent's savings S_t^A. As a result, we can write the

[16] Recall that

$$dS_t^I = rS_t^I dt + dX_t - dw_t^P - dw_t^A + c_t^I dt.$$

Thus, $S_0^I = dw_0^I - dw_0^A + S_{0^-}^I$, as $S_t^I dt + dX_t - c_t^I dt$ are infinitesimal. When dw_0^P and dw_0^A are infinitesimal, then $S_0^I = S_{0^-}^I$. Otherwise, we include these lumpy transfers at time $t = 0$ in the integral in (4.50).

time-t value of the objective (i.e., the intermediary's deferred compensation)

$$v_t^I = \mathbb{E}_t^I \left[\int_t^\infty e^{-r(s-t)} \left(dX_s - dw_s^P - dw_s^A - \frac{1}{2}\theta^I r(\beta_s^I \sigma)^2 ds - \frac{\lambda (a_s^I)^2}{2} ds \right) \right]$$

as function of V_t^A, S_t^A, v_t^P, in that $v_t^I = \tilde{F}(V_t^A, S_t^A, v_t^P)$. In the following, we suppress time subscripts and the argument of the function $\tilde{F}(\cdot)$, whenever no confusion is likely to arise.

By the dynamic programming principle, the function $\tilde{F}$ solves the following HJB equation

$$r\tilde{F}dt = \max_{\beta^A, \beta^I, \beta^P, dw^P, dw^A \geq 0} \left\{ \mathbb{E}^I[dX - dw^P - dw^A] - \frac{1}{2}\theta^I r(\beta^I \sigma)^2 dt - \frac{\lambda (a^I)^2}{2} dt + \mathbb{E}^I d\tilde{F} \right\}, \quad (4.51)$$

subject to incentive compatibility with respect to effort, (4.12), and the limited liability constraints (see Lemma 21). The term $\mathbb{E}^I d\tilde{F}$ can be expanded using Itô's Lemma, which yields that (4.51) is a partial differential equation (PDE). The expectation $\mathbb{E}^I$ is taken under the measure $\mathbb{P}^I$, induced by efforts $\hat{a}^A$ and a^I. That is, the expectation $\mathbb{E}^I$ is taken under the intermediary's information.

It is beyond the scope of the paper to provide a formal existence and uniqueness proof for a solution to (4.51). Thus, we assume throughout the remainder:

Assumption 1 (Existence, Uniqueness & Smoothness) *The PDE (4.51) admits a unique solution $\tilde{F}$ that is twice continuously differentiable.*

4.10.2 Part II

In this part, we conjecture and verify that $\tilde{F}(V^A, S^A, v^P)$ takes the form $\tilde{F}(V^A, S^A, v^P) = F(v^A) - v^P$ with $v^A = V^A - S^A$ and a (twice continuously differentiable) function $F(v^A)$. To start with, first recall that the expectation $\mathbb{E}^I$ is taken under the probability measure $\mathbb{P}^I$ (induced by efforts $\hat{a}^A$ and a^I) and hence is taken under the intermediary's information. Thus, the following holds

$$\begin{aligned} \mathbb{E}^I[dX - \hat{a}^A dt - a^I dt] &= 0 \\ \mathbb{E}^I[dX - \hat{a}^A dt - \hat{a}^I dt] &= (a^I - \hat{a}^I)dt \\ \mathbb{E}^I[dX - \bar{a}^A dt - \bar{a}^I dt] &= (\hat{a}^A + a^I - \bar{a}^A - \bar{a}^I)dt. \end{aligned} \quad (4.52)$$

Second, because payouts dw^A to the agent i) cannot become negative, ii) reduce v^A by amount dw^A (e.g., see (4.16)), and iii) because the agent's limited liability requires $v^A \geq 0$, it is natural to conjecture that optimal payouts occur at some upper boundary $\overline{v}$ and reflect v^A back into the interior of the state space, like in DeMarzo and Sannikov (2006). That is, we conjecture that as stipulated in Proposition 15, optimal payouts to the agent take the form $dw^A = \max\{v^A - \overline{v}, 0\}$, with endogenous payout boundary $\overline{v}$. At the payout boundary $v^A = \overline{v}$, the smooth pasting condition

$$F'(\overline{v}) = -1.$$

holds, as shown, e.g., in Dumas (1991a) and DeMarzo and Sannikov (2006). We verify the optimality of this payout strategy in Part VII of the proof, where we verify that the contracts Π^A, Π^P from Proposition 15 are indeed optimal.

Third, consider $v^A < \overline{v}$, so $dw^A = 0$. Using the conjecture $\tilde{F}(V^A, S^A, v^P) = F(v^A) - v^P$, Itô's Lemma, (4.16), and (4.10), we get

$$\begin{aligned} \mathbb{E}^I[d\tilde{F}] &= \mathbb{E}^I d[F(v^A) - v^P] \\ &= \left[F'(v^A) \left(\frac{\theta^A r(\beta^A \sigma)^2}{2} + \frac{\delta(\hat{a}^A)^2}{2} \right) + \frac{F''(v^A)(\beta^A \sigma)^2}{2} \right] dt \\ &+ F'(v^A)\beta^A \mathbb{E}^I[dX - \hat{a}^A dt - \hat{a}^I dt] - [rv^P dt - \mathbb{E}^I dw^P] - \beta^P \mathbb{E}^I[dX - \bar{a}^A dt - \bar{a}^I dt]. \end{aligned}$$

Using (4.52), we can rewrite the above as

$$\mathbb{E}^I[d\tilde{F} - dw^P] = \Big\{\beta^P(\bar{a}^A + \bar{a}^I - \hat{a}^A - a^I) + F'(v^A)\beta^A(a^I - \hat{a}^I) - rv^P \\ + F'(v^A)\left(\frac{\theta^A r(\beta^A\sigma)^2}{2} + \frac{\delta(\hat{a}^A)^2}{2}\right) + \frac{F''(v^A)(\beta^A\sigma)^2}{2}\Big\}dt \tag{4.53}$$

Inserting (4.53) and $\tilde{F} = F(v^A) - v^P$ into (4.51), noting $\mathbb{E}^I dX = (\hat{a}^A + a^I)dt$, and simplifying yields

$$rF(v^A) = \max_{\beta^A,\beta^I,\beta^P,dw^P}\Big\{(\hat{a}^A + a^I)(1-\beta^P) + \beta^P(\bar{a}^A + \bar{a}^I) + F'(v^A)\beta^A(a^I - \hat{a}^I) - \frac{\lambda(a^I)^2}{2} \\ - \frac{\theta^I r(\beta^I\sigma)^2}{2} + F'(v^A)\left(\frac{\theta^A r(\beta^A\sigma)^2}{2} + \frac{\delta(\hat{a}^A)^2}{2}\right) + \frac{F''(v^A)(\beta^A\sigma)^2}{2}\Big\}. \tag{4.54}$$

The ordinary differential equation (ODE) (4.54) is solved subject to incentive compatibility with respect to effort, (4.12), and the limited liability constraints. Notably, the right-hand-side of (4.54) depends on V^A and S^A only through the difference $v^A = V^A - S^A$, and so do the left-hand-side and the optimal controls. In addition, the right-hand-side does not depend on v^P or dw^P, and so do the left-hand-side and the optimal controls $\beta^A, \beta^I, \beta^P$. This confirms the conjecture $\tilde{F}(V^A, S^A, v^P) = F(v^A) - v^P$.

Finally, (4.54) illustrates that the intermediary dynamically maximizes the joint monetary payoff $v^I + v^P = F(v^A)$ that is split between intermediary and principal. In other words, maximizing v^I or maximizing $v^I + v^P$ is equivalent. The reason is that the intermediary is residual claimant on total firm value and, as payouts to the principal dw_t^P are not sign-restricted, the intermediary can extract all surplus from the principal.[17]

4.10.3 Part III

This part of the proof characterizes the intermediary's incentives for effort β^I (compare (4.12)) and then characterizes optimal payouts to the principal. First, note that $vol(dw^I + dw^A) = vol(dX) - vol(dw^P) = \sigma - vol(dw^P)$, where $vol(\cdot)$ denotes the volatility of a stochastic process and $dw^I = dX - dw^A - dw^P$. By (4.16), the volatility of $dv^I - dw^A$ is $\beta^I\sigma - vol(dw^I + dw^A)$ which is $\beta^I\sigma - \sigma + vol(dw^P)$.

On the other hand, as $v^I = F(v^A) - v^P$, Itô's Lemma implies that the volatility of $dv^I - dw^A$ is $F'(v^A)\beta^A\sigma - vol(dv^P)$, where $vol(dv^P)$ is the volatility of v^P and $vol(dv^A + dw^A) = \beta^A\sigma$ from (4.16). Note that (4.10) implies $vol(dv^P) = \beta^P\sigma - vol(dw^P)$. Thus,

$$\beta^I\sigma - \sigma + vol(dw^P) = F'(v^A)\beta^A\sigma - vol(dv^P) = F'(v^A)\beta^A\sigma - \beta^P\sigma + vol(dw^P),$$

which can be rewritten as

$$\beta^I\sigma - (1-\beta^P)\sigma = F'(v^A)\beta^A\sigma$$

and is equivalent to (4.23), as stated in Lemma 23.

Second, let us discuss the choice of dw^P and v^P. Note that payouts to the principal dw^P are not sign-restricted. Therefore, it is always possible to stipulate payouts dw^P to the principal, as long as the limited liability constraint $v^P \geq 0$ holds. As payouts to the principal dw^P are always possible and change v^P by amount $-dw^P$ (see (4.10)) but do not change the level of v^A, it follows that controlling dw^P is equivalent to controlling v^P.

Importantly, (4.23) is independent of the choice of dw^P and v^P. Likewise, the right-hand-side of (4.54) is independent of dw^P and v^P, meaning that the choice of dw^P and v^P does not directly affect the intermediary's dynamic optimization in the HJB equation (4.54). As a result, dw^P and v^P affect the intermediary's dynamic optimization only via the limited

[17] As such, an equivalent formulation of the intermediary's dynamic optimization is that the intermediary chooses contracts Π^A, Π^P to dynamically the intermediary's and principal's joint monetary payoff $v_t^I + v_t^P$.

liability constraint $v^I \geq R$, which can be rewritten as $F(v^A) - v^P \geq R$. Because reducing v^P relaxes the limited liability constraint, it follows that setting $v^P = 0$ is optimal. Note that by (4.10), payouts $dw^P = \beta^P(dX - \bar{a}^A dt - \bar{a}^I dt)$ implement the choice $v^P = 0$ (and $dv^P = 0$), as stated in Lemma 22.

4.10.4 Part IV

Part IV of the proof maximizes the HJB equation (4.54) with respect to β^A and hence characterizes the agent's optimal incentives, taking into account the effort incentive constraint (4.12) and the relationship (4.23).

Observe that (4.12) and (4.23) imply

$$\hat{a}^I = a^I = \frac{1 - \beta^P + F'(v^A)\beta^A}{\lambda} = \frac{\beta^I}{\lambda}. \tag{4.55}$$

In addition, one can verify that maximizing the right-hand-side of (4.54) over a^I yields the same expression for optimal effort a^I as in (4.55). Plugging $a^I = \hat{a}^I = \beta^I/\lambda$ and $a^A = \hat{a}^A = \beta^A/\delta$ into (4.54) yields

$$\begin{aligned} rF(v^A) = \max_{\beta^A,\beta^P,\beta^I} \Bigg\{ & \left(\frac{\beta^I}{\lambda} + \frac{\beta^A}{\delta}\right)(1-\beta^P) + \beta^P(\bar{a}^I + \bar{a}^A) - \frac{(\beta^I)^2}{2\lambda} - \frac{\theta^I r(\beta^I\sigma)^2}{2} \\ & + F'(v^A)\left(\frac{\theta^A r(\beta^A\sigma)^2}{2} + \frac{(\beta^A)^2}{2\delta}\right) + \frac{F''(v^A)(\beta^A\sigma)^2}{2}\Bigg\}, \end{aligned} \tag{4.56}$$

which is solved subject to the limited liability constraints and subject to (4.23). Due to (4.23), one of the controls $(\beta^A, \beta^I, \beta^P)$ is redundant and, in particular, maximizing (4.56) over (β^A, β^I) is equivalent to maximizing (4.56) over (β^A, β^P). Also note that the agent's incentives β^A and prescribed effort $\hat{a}^A$ are not observable to the principal and not contractible between principal and intermediary. Accordingly, when maximizing (4.56) with respect to β^A, the intermediary takes the principal's contract Π^P and so effort levels $\bar{a}^A, \bar{a}^I$ and the sensitivity β^P as given.

Taking the first order condition in (4.56) with respect to β^A yields

$$\begin{aligned} & \frac{1-\beta^P}{\delta} + \frac{F'(v^A)(1-\beta^P)}{\lambda} - \left(\frac{1}{\lambda} + \theta^I r\sigma^2\right)(\underbrace{1 - \beta^P + F'(v^A)\beta^A}_{=\beta^I})F'(v^A) \\ & + F'(v^A)\beta^A\left(\frac{1}{\delta} + \theta^A r\sigma^2\right) + F''(v^A)\sigma^2\beta^A = 0. \end{aligned} \tag{4.57}$$

We can solve (4.57) to get

$$\beta^A = (1-\beta^P)\pi^I \tag{4.58}$$

with

$$\pi^I := \frac{1 - \delta\theta^I r\sigma^2 F'(v^A)}{\delta\left(\frac{1}{\lambda} + \theta^I r\sigma^2\right)(F'(v^A))^2 - (\delta\theta^A r\sigma^2 + 1)F'(v^A) - \delta\sigma^2 F''(v^A)}, \tag{4.59}$$

as stated in Lemma 23.

Note that (4.58) and (4.12) jointly imply $a^A = \hat{a}^A = \frac{(1-\beta^P)\pi^I}{\delta}$. That is, in optimum (under incentive compatible contracts), it must hold that $\bar{a}^A = a^A = \hat{a}^A = \frac{(1-\beta^P)\pi^I}{\delta}$. Effectively, (4.58) is the intermediary's "delegation" incentive condition for implementing $\hat{a}^A = \bar{a}^A = a^A = \frac{(1-\beta^P)\pi^I}{\delta}$. As long as $\bar{a}^A = \frac{(1-\beta^P)\pi^I}{\delta}$ holds, the intermediary finds it optimal to implement $\bar{a}^A = \hat{a}^A$.

Finally, we rewrite (4.23) as

$$\beta^A F'(v^A) = \beta^I - (1-\beta^P) \tag{4.60}$$

and multiply both sides of (4.58) by $F'(v^A)$ to get

$$\beta^A F'(v^A) = (1-\beta^P)\pi^I F'(v^A). \tag{4.61}$$

Combining (4.60) and (4.61) yields

$$\beta^I = (1+F'(v^A)\pi^I)(1-\beta^P), \tag{4.62}$$

which is (4.26) (as stated in Lemma 23). Next, we rewrite (4.58) to $1-\beta^P = \frac{\beta^A}{\pi^I}$ and insert this expression into (4.62) to get

$$\beta^A = \beta^I \left(\frac{\pi^I}{1+F'(v^A)\pi^I}\right), \tag{4.63}$$

which is (4.27) (as stated in Lemma 23).

4.10.5 Part V

Part V of the proof maximizes the HJB equation (4.54) with respect to β^I (or equivalently β^P), taking into account the effort incentive constraint (4.12), the relationship (4.23), and the characterization of the agent's incentives (i.e., (4.27) or (4.63)).

Plugging $\bar{a}^A = a^A = \beta^A/\delta$ and $\bar{a}^I = a^I = \beta^I/\lambda$ into (4.56) gives

$$\begin{aligned} rF(v^A) = \max_{\beta^A,\beta^P,\beta^I} \Bigg\{ &\left(\frac{\beta^I}{\lambda}+\frac{\beta^A}{\delta}\right) - \frac{(\beta^I)^2}{2\lambda} - \frac{\theta^I r(\beta^I\sigma)^2}{2} \\ &+ F'(v^A)\left(\frac{\theta^A r(\beta^A\sigma)^2}{2}+\frac{(\beta^A)^2}{2\delta}\right) + \frac{F''(v^A)(\beta^A\sigma)^2}{2}\Bigg\}. \end{aligned} \tag{4.64}$$

The optimization in (4.64) is solved subject to the delegation incentive condition (4.58) (ensuring $\hat{a}^A = \bar{a}^A$) and (4.23). Also recall that (4.58) and (4.23) jointly imply (4.63).

Note that we can use (4.12) to replace sensitivities by efforts and to rewrite (4.64) as

$$\begin{aligned} rF(v^A) = \max_{\beta^A,\beta^P,\beta^I} \Bigg\{ &a^A + a^I - \frac{\lambda(a^I)^2}{2} - \frac{\theta^I r(\beta^I\sigma)^2}{2} \\ &+ F'(v^A)\left(rv^A + \frac{\theta^A r(\beta^A\sigma)^2}{2}+\frac{\delta(a^A)^2}{2}\right) + \frac{F''(v^A)(\beta^A\sigma)^2}{2}\Bigg\}, \end{aligned}$$

which is (4.28) (after dropping control β^P which is possible due to (4.23)). This shows that the intermediary's value function solves under the optimal contracts the HJB equation (4.28), as stipulated by Proposition 15.

Inserting (4.58) and (4.62) into (4.64) changes (4.64) to

$$\begin{aligned} rF(v^A) = \max_{\beta^P} \Bigg\{ &(1-\beta^P)\left(\frac{1+F'(v^A)\pi^I}{\lambda}+\frac{\pi^I}{\delta}\right) \\ &- (1-\beta^P)^2(1+F'(v^A)\pi^I)^2\left(\frac{1}{2\lambda}+\frac{\theta^I r(\sigma)^2}{2}\right) \\ &+ (1-\beta^P)^2 F'(v^A)(\pi^I)^2\left(\frac{1}{2\delta}+\frac{\theta^A r(\sigma)^2}{2}\right) + \frac{(1-\beta^P)^2 F''(v^A)(\pi^I\sigma)^2}{2}\Bigg\}, \end{aligned} \tag{4.65}$$

where we have dropped the redundant controls β^A, β^I. We can solve the optimization in (4.65) to get

$$1-\beta^P = \frac{\frac{1+F'(v^A)\pi^I}{\lambda} + \frac{\pi^I}{\delta}}{\left(\frac{1}{\lambda}+\theta^I r\sigma^2\right)(1+F'(v^A)\pi^I)^2 - F'(v^A)(\pi^I)^2\left(\theta^A r\sigma^2 + \frac{1}{\delta}\right) - F''(v^A)(\pi^I\sigma)^2}, \tag{4.66}$$

as stated in Proposition 15.

Using (4.62), it follows that (4.66) is equivalent to

$$\beta^I = \frac{(1+F'(v^A)\pi^I)\left(\frac{1+F'(v^A)\pi^I}{\lambda} + \frac{\pi^I}{\delta}\right)}{\left(\frac{1}{\lambda}+\theta^I r\sigma^2\right)(1+F'(v^A)\pi^I)^2 - F'(v^A)(\pi^I)^2\left(\frac{1}{\delta} + \theta^A r\sigma^2\right) - F''(v^A)(\pi^I\sigma)^2}, \tag{4.67}$$

as stated in Proposition 15. Due to (4.27) (or (4.63)), the agent's optimal incentives are then

$$\beta^A = \frac{\pi^I\left(\frac{1+F'(v^A)\pi^I}{\lambda} + \frac{\pi^I}{\delta}\right)}{\left(\frac{1}{\lambda}+\theta^I r\sigma^2\right)(1+F'(v^A)\pi^I)^2 - F'(v^A)(\pi^I)^2\left(\frac{1}{\delta} + \theta^A r\sigma^2\right) - F''(v^A)(\pi^I\sigma)^2}. \tag{4.68}$$

To summarize the sensitivities $\beta^A, \beta^I, \beta^P$ that solve the maximization in (4.54) satisfy (4.68), (4.67), and (4.66) respectively. The effort incentive constraints (4.12) map sensitivities β^A, β^I to effort levels a^A, a^I.

4.10.6 Part VI — Properties of the value function

As shown in the previous part of the proof, the value function $F(v^A)$ under the contracts from Proposition 15 solves the HJB equation (4.28), subject to $F'(\overline{v}) = -1$. We demonstrate that $F(v^A)$ is strictly concave, in that $F''(v^A) < 0$ for all $v^A \in [0, \overline{v}]$.

Note that by definition the optimal payout boundary satisfies

$$\overline{v} = \inf\{v^A \geq 0 : F'(v^A) \leq -1\}.$$

Otherwise, payouts would be optimal once v^A reaches some value $v' < \overline{v}$ with $F'(v') \leq -1$, contradicting the fact that $\overline{v}$ is the payout boundary. If now $F''(\overline{v}) > 0$, then the smooth pasting condition, $F'(\overline{v}) = -1$, implies that there exists $v' < \overline{v}$ with $F'(v') < -1$, a contradiction. Thus, $F''(\overline{v}) \leq 0$.

Next, using the envelope theorem, we differentiate the HJB equation (4.28) evaluated under the optimal controls to get

$$F'''(v^A) = \frac{2}{(\beta^A\sigma^2)}F''(v^A)\left(rv^A + \frac{\theta^A r(\beta^A\sigma)^2}{2} + \frac{\delta(a^A)^2}{2}\right). \tag{4.69}$$

Suppose that $F''(\overline{v}) = 0$. Then, $F'''(\overline{v}) = 0$, while $F'(\overline{v}) = -1$. As a result, the solution to (4.28) is affine and takes the form $F(v^A) = K - v^A$ for some constant K. Thus, there exists $v' < \overline{v}$ with $F'(v') = -1$, a contradiction. It follows that $F''(\overline{v}) < 0$. Because (4.69) implies that $F''(v^A) < 0 \iff F'''(v^A) > 0$, it follows that $F''(v^A) < 0$ for all $v^A \in [0, \overline{v}]$.

Finally, we show that the payout boundary satisfies $F(\overline{v}) = R$, so that the intermediary's limited liability constraint binds at the payout boundary. When $F(\overline{v}) > R$, the choice of the payout boundary is not constrained by the intermediary's limited liability, in that it is possible to slightly increase or decrease the payout boundary without violating limited liability. Without constraints, the payout boundary must satisfy the super-contact condition $F''(\overline{v}) = 0$ (as shown, e.g., in Dumas (1991a)). However, we have shown that $F''(\overline{v}) < 0$. As a result, at the payout boundary $\overline{v}$, the limited liability constraint must bind, in that $F(\overline{v}) = R$.

4.10.7 Part VII — Verification

Under the proposed strategy and contracts Π^A, Π^P from Proposition 15, the sensitivities $\beta_t^A, \beta_t^I, \beta_t^P$ solve the maximization in (4.54) or, equivalently, the maximization in (4.56). Payouts to the principal take the form $dw^P = \beta_t^P(dX_t - \bar{a}_t^A dt - \bar{a}_t^I dt)$, and payouts to the agent take the form $dw_t^A = \max\{v_t^A - \overline{v}, 0\}$. As shown in Parts IV and V of the proof, the sensitivities $\beta_t, \beta_t^I, \beta_t^P$ that solve the maximization in (4.56) (or (4.54)) satisfy (4.68), (4.67), and (4.66). The incentive condition (4.12) maps sensitivities β_t^A, β_t^I to effort levels a_t^A, a_t^I. In what follows, we verify that the proposed strategy (i.e., the proposed contracts) from Proposition 15 yields higher payoff than any other strategy (i.e., any other incentive compatible contracts) and thus is indeed optimal.

Take any time $t < \tau$. Suppose that the intermediary deviates from the proposed strategy and follows an alternative strategy up to time t (i.e., for times $s \leq t$), with sensitivities $\beta_s^A, \beta_s^I, \beta_s^P$, payouts to the principal dw_s^P, and payouts to the agent $dw_s^A \geq 0$. After time t (i.e., for times $s \geq t$), the intermediary follows the proposed strategy (contracts) from Proposition 15. Then, the intermediary's payoff at time 0 (under the deviation) can be written as

$$G_t = \int_0^t e^{-rs}\left(dX_s - dw_s^P - dw_s^A - \frac{\lambda(a_s^I)^2}{2}ds - \frac{\theta^I r(\beta_s^I\sigma)^2}{2}ds\right) + e^{-rt}[F(v_t^A) - v_t^P]. \quad (4.70)$$

We use (4.10), (4.16), and Itô's Lemma and differentiate G_t with respect to time, t, to get

$$\begin{aligned} e^{rt}dG_t = \Bigg\{&(\hat{a}_t^A + a_t^I)(1-\beta_t^P) + \beta_t^P(\bar{a}_t^I + \bar{a}_t^A) + F'(v_t^A)\beta_t^A(a_t^I - \hat{a}_t^I) - \frac{\lambda(a_t^I)^2}{2} \\ &- \frac{\theta^I r(\beta_t^I\sigma)^2}{2} + F'(v_t^A)\left(\frac{\theta^A r(\beta_t^A\sigma)^2}{2} + \frac{\delta(\hat{a}_t^A)^2}{2}\right) + \frac{F''(v_t^A)(\beta_t^A\sigma)^2}{2} - rF(v_t^A)\Bigg\}dt \\ &+ \underbrace{(1-\beta_t^P + F'(v_t^A)\beta_t^A)}_{=\beta_t^I}[dX_t - \hat{a}_t^A dt - a_t^I dt] - dw_t^A(F'(v_t^A)+1) \\ &= \mu_t^G dt + \beta_t^I[dX_t - \hat{a}_t^A dt - a_t^I dt] + dw_t^A(F'(v_t^A) - 1). \end{aligned}$$

By (4.56) (or equivalently (4.54)), the drift term in curly brackets, μ_t^G is zero when β_t^A, β_t^I, and β_t^P solve the optimization in (4.56) (or equivalently (4.54)) subject to incentive compatibility and limited liability constraints.[18] As shown in Parts IV and V of this proof, this is the case if and only if β_t^A, β_t^I, and β_t^P satisfy (4.68), (4.67), and (4.66). Any other choice of β_t^A, β_t^I, and β_t^P makes the drift term μ_t^G weakly negative, in that $\mu_t^G \leq 0$.

In addition, recall that because $F(v^A)$ is strictly concave and because $F'(\overline{v}) = 1$, it follows that $F'(v_t^A) \geq -1$ with equality if and only if $v_t^A = \overline{v}$. As $dw_t^A \geq 0$ and $F'(v_t^A) \geq -1$, the term $-dw_t^A(F'(v_t^A)+1)$ is weakly negative under any payout policy $dw_t^A \geq 0$ and zero under the proposed payout policy $dw_t^A = \max\{v_t^A - \overline{v}, 0\}$.

Next, recall that $\frac{dX_t - \hat{a}_t^A dt - a_t^I dt}{\sigma}$ is the increment of a standard Brownian motion under the probability measure $\mathbb{P}^I$ that is taken under the intermediary's information. Because the sensitivity β_t^I is by assumption bounded, it follows that

$$\mathbb{E}_0^I\left[\int_0^t e^{-rs}\beta_s^I\left(dX_s - \hat{a}_s^A ds - a_s^I ds]\right)\right] = 0.$$

Thus, G_t, with

$$dG_t = e^{-rt}\left(\underbrace{\mu_t^G dt}_{\leq 0}\underbrace{-dw_t^A(F'(v_t^A)+1)}_{\leq 0}\right) + e^{-rt}\beta_t^I[dX_t - \hat{a}_t^A dt - a_t^I dt],$$

[18] As shown in Part III of this proof, it is optimal to set $dw_t^P = \beta_t^P(dX_t - \bar{a}_t^A dt - \bar{a}_t^I dt)$. As a result, payouts to the principal affect the intermediary's optimization only via the sensitivity β_t^P and the limited liability constraint $F(v_t^A) - v_t^P \geq R$.

follows a super-martingale (i.e., decreases in expectation) under the measure $\mathbb{P}^I$. Thus,

$$G_0 \geq \mathbb{E}_0^I G_t \tag{4.71}$$

Because $a_t^j \in [0, A]$ with $A < \infty$ for $j = A, I$, it follows that total surplus is bounded from above by $2A/r$. As such, $F(v_t^A) - v_t^P$ is bounded from above by $2A/r$ and bounded from below by 0 (due to limited liability) so that

$$0 \leq \lim_{t\to\infty} e^{-rt}[F(v_t^A) - v_t^P] \leq \lim_{t\to\infty} e^{-rt}\frac{2A}{r} = 0. \tag{4.72}$$

Thus, (4.70) implies

$$\lim_{t\to\infty} \mathbb{E}_0^I G_t = \mathbb{E}_0^I \left[\int_0^\tau e^{-rs}\left(dX_s - dw_s^P - dw_s^A - \frac{\lambda(a_s^I)^2}{2}ds - \frac{\theta^I r(\beta_s^I\sigma)^2}{2}ds\right)\right] = \mathbb{E}_0^I G_\tau.$$

Next, we take the limit $t \to \infty$ in (4.71) to obtain

$$\underbrace{G_0}_{=F(v_0^A)-v_0^P} \geq \mathbb{E}_0^I G_\tau = \mathbb{E}_0^I \left[\int_0^\tau e^{-rs}\left(dX_s - dw_s^P - dw_s^A - \frac{\lambda(a_s^I)^2}{2}ds - \frac{\theta^I r(\beta_s^I\sigma)^2}{2}ds\right)\right]. \tag{4.73}$$

The inequality (4.73) implies that the proposed contracts Π^A, Π^P from Proposition 15 are indeed optimal and solve the intermediary's problem (4.3), as (4.73) holds in equality under the proposed contracts and the intermediary's expected payoff at time 0 under the proposed contracts is $G_0 = F(v_0^A) - v_0^P$.

4.10.8 Part VIII — Liquidation

We demonstrate under what circumstances firm liquidation is indeed optimal when the agent's contract is terminated. Note that the agent's contract is terminated at time τ when $v_\tau^A = 0$, leading to $a_t^A = \beta_t^A = v_t^A = 0$ for $t \geq \tau$. If the intermediary does not liquidate the firm, the HJB equation (4.28) implies that (after inserting $a^A = \beta^A = v^A = 0$):

$$F(0) = \max_{\beta^I} \frac{1}{r}\left(a^I - \frac{\lambda(a^I)^2}{2} - \frac{\theta^I r(\beta^I\sigma)^2}{2}\right), \tag{4.74}$$

which is solved subject to the effort incentive condition $a^I = \frac{\beta^I}{\lambda}$. Note that as the intermediary has deep pockets, she can run the firm forever without the agent after the agent's contract is terminated at time τ and doing so yields payoff characterized in (4.74).

The solution to the optimization problem in (4.74) satisfies

$$\beta^I = \frac{1}{1 + \lambda\theta^I r\sigma^2} \quad \text{and} \quad a^I = \frac{\beta^I}{\lambda} = \frac{1}{\lambda + \lambda^2\theta^I r\sigma^2}. \tag{4.75}$$

Inserting the expressions (4.75) into (4.74) yields

$$F(0) = \frac{1}{r}\left(a^I - \frac{\lambda(a^I)^2}{2} - \frac{\theta^I r(a^I\lambda\sigma)^2}{2}\right)$$

with a^I from (4.75), which is the intermediary's payoff when she continues running the firm without the agent after time τ. It follows that when

$$R \geq R_L := \frac{1}{r}\left(a^I - \frac{\lambda(a^I)^2}{2} - \frac{\theta^I r(a^I\lambda\sigma)^2}{2}\right) \quad \text{with} \quad a^I = \frac{1}{\lambda + \lambda^2\theta^I r\sigma^2}, \tag{4.76}$$

it is optimal for the intermediary to liquidate the firm when $v^A = 0$ and the agent's contract is terminated (instead of running the firm without the agent). Recall that we have made the assumption $R \geq R_L$. Thus, the intermediary optimally liquidates the firm when $v^A = 0$ and seizes the liquidation value R, so the boundary condition $F(0) = R$ applies.

4.11 Proof of Corollary 6

We prove both claims of the corollary separately. All claims of the corollary are proven in the limit $\theta^I, 1/\lambda \to 0$, and by continuity hold when θ^I and $1/\lambda$ are sufficiently small.

4.11.1 Claim 1

First, we take the expression for π^I in (4.25) and take the limit $\theta^I \to 0$ and $1/\lambda \to 0$ to obtain

$$\pi^I = \frac{1}{-(\delta\theta^A r\sigma^2 + 1)F'(v^A) - \delta\sigma^2 F''(v^A)} \tag{4.77}$$

Second, we take the expression for $1 - \beta^P$ in (4.29) and take the limit $\theta^I \to 0$ and $1/\lambda \to 0$ to obtain

$$\begin{aligned} 1 - \beta^P &= \frac{\frac{\pi^I}{\delta}}{-F'(v^A)(\pi^I)^2\left(\frac{1}{\delta} + \theta^A r\sigma^2\right) - F''(v^A)(\pi^I\sigma)^2} \\ &= \frac{\frac{1}{\delta}}{-F'(v^A)\pi^I\left(\frac{1}{\delta} + \theta^A r\sigma^2\right) - \pi^I F''(v^A)\sigma^2}, \end{aligned} \tag{4.78}$$

where the second equality follows from dividing through $\pi^I \neq 0$.

Third, inspecting (4.77) and (4.78), we observe that

$$(1 - \beta^P)\pi^I = \pi^I \iff 1 - \beta^P = 1 \iff \beta^P = 0.$$

Note that Lemma 23 implies $\beta^A = \pi^I(1 - \beta^P) = \pi^I$ and

$$\beta^I = (1 + F'(v^A)\pi^I)(1 - \beta^P) = 1 + F'(v^A)\pi^I,$$

which was to show.

4.11.2 Claim 2

Suppose that $v^A < v^*$, so that $F'(v^A) > 1$ as $F'(v^*) = 0$ and $F''(v^A) < 0$ for all $v^A \in [0, \overline{v}]$. Because the agent's effort is by assumption interior (i.e., $a^A < A$), it must be that $\pi^I > 0$, so that $\beta^I = 1 + F'(v^A)\pi^I > 1$ and $a^I = \beta^I/\lambda > 1/\lambda$ (see (4.12)) where $a^I_{FB} = 1/\lambda$. Conversely, when $v^A > v^*$, then $F'(v^A) < 0$, so that $\beta^I = 1 + F'(v^A)\pi^I < 1$ and $a^I = \beta^I/\lambda < 1/\lambda$.

4.12 Proof of Proposition 16

The proof of Proposition 16 is analogous to the proof of Proposition of 15. To point out the difference and to highlight comparability, we structure the proof of Proposition 16 along the lines of the proof of Proposition of 15. As a result, the proof of Proposition 16 is split into several parts that correspond to the respective parts in the proof of Proposition 15.

4.12.1 Parts I-III

Identical to Parts I-III from the proof of Proposition 15. As such, the intermediary's value function can be expressed as a function of v^A, in that $v^I = F(v^A) - v^P$. The function $F(v^A)$ solves the HJB equation (4.54). Payouts to the agent take the form $dw^A = \max\{v - \overline{v}, 0\}$ with endogenous payout boundary $\overline{v}$. Optimal payouts to the principal take the form $dw^P =$

$\beta^P(dX - \bar{a}^A dt - \bar{a}^I dt)$ and implement the choice $v^P = 0$, where $\bar{a}^j = \hat{a}^j$ holds because the contract Π^A is observed by the principal and contractible between principal and intermediary. In addition, the incentive conditions for effort, (4.12), and the relationship (4.23) hold.

4.12.2 Part IV and part V

Recall that because Π^A is observable for the principal and contractible between principal and intermediary, it follows that $\hat{a}^j = \bar{a}^j$ for $j = A, I$. As a result, (4.12) implies

$$\hat{a}^A = \bar{a}^A = a^A = \frac{\beta^A}{\delta} \quad \text{and} \quad \hat{a}^I = \bar{a}^I = a^I = \frac{\beta^I}{\lambda}. \tag{4.79}$$

In addition, one can verify that maximizing the right-hand-side of (4.54) over a^I yields the same expression for optimal effort a^I as in (4.79). We insert $a^I = \hat{a}^I = \bar{a}^I = \beta^I/\lambda$ and $a^A = \hat{a}^A = \bar{a}^I = \beta^A/\delta$ as well as $\beta^I = (1-\beta^P) + F'(v^A)\beta^A$ — i.e., the relationship (4.23) — into (4.54) to obtain

$$\begin{aligned} rF(v^A) = \max_{\beta^A,\beta^I} \Bigg\{ & \left(\frac{\beta^I}{\lambda} + \frac{\beta^A}{\delta}\right) - \frac{(\beta^I)^2}{2\lambda} - \frac{\theta^I r(\beta^I\sigma)^2}{2} \\ & + F'(v^A)\left(\frac{\theta^A r(\beta^A\sigma)^2}{2} + \frac{(\beta^A)^2}{2\delta}\right) + \frac{F''(v^A)(\beta^A\sigma)^2}{2}\Bigg\}, \end{aligned} \tag{4.80}$$

where we drop the controls dw^P and β^P (which is possible due to (4.23)).

Also observe that we can use (4.12) to rewrite (4.80) as

$$\begin{aligned} rF(v^A) = \max_{\beta^A,\beta^I} \Bigg\{ & a^A + a^I - \frac{\lambda(a^I)^2}{2} - \frac{\theta^I r(\beta^I\sigma)^2}{2} \\ & + F'(v^A)\left(rv^A + \frac{\theta^A r(\beta^A\sigma)^2}{2} + \frac{\delta(a^A)^2}{2}\right) + \frac{F''(v^A)(\beta^A\sigma)^2}{2}\Bigg\}, \end{aligned}$$

which is (4.28). This shows that the intermediary's value function solves under the optimal contracts the HJB equation (4.28), as stipulated by Proposition 16.

The optimal values β^A and β^I, solving the maximization in (4.80), must solve the first order conditions

$$\begin{aligned} \frac{\partial F(v^A)}{\partial \beta^A} &\propto \frac{1}{\delta} + F'(v^A)\left(\beta^A\theta^A r\sigma^2 + \frac{\beta^A}{\delta}\right) + \beta^A F''(v^A)\sigma^2 = 0 \\ \frac{\partial F(v^A)}{\partial \beta^I} &\propto \frac{1}{\lambda} - \frac{\beta^I}{\lambda} - \beta^I\theta^I r\sigma^2 = 0. \end{aligned}$$

We can solve these two equations to get

$$\beta^A = \frac{1}{-(\delta\theta^A r\sigma^2 + 1)F'(v^A) - \delta\sigma^2 F''(v^A)}, \tag{4.81}$$

which is (4.30), and

$$\beta^I = \frac{1}{1 + \lambda\theta^I r\sigma^2}, \tag{4.82}$$

which is (4.31). Using (4.23), we obtain

$$1 - \beta^P = \beta^I - F'(v^A)\beta^A = \frac{1}{1 + \lambda\theta^I r\sigma^2} - \frac{F'(v^A)}{-(\delta\theta^A r\sigma^2 + 1)F'(v^A) - \delta\sigma^2 F''(v^A)}, \tag{4.83}$$

which can be solved for β^P.

4.12.3 Part VI — Properties of the value function

Identical to Part VI of the proof of Proposition 15. That is, the payout boundary satisfies $F(\overline{v}) = R$ and the value function is strictly concave, in that $F''(\overline{v}) < 0$ for $v \in [0, \overline{v}]$.

4.12.4 Part VII — Verification

We provide the formal verification argument that the proposed strategy and contracts Π^A, Π^P from Proposition 16 are indeed optimal. Under the proposed contracts Π^A, Π^P, the sensitivities $\beta_t^A, \beta_t^I, \beta_t^P$ solve the maximization in (4.80), while $\bar{a}_t^j = \hat{a}_t^j$ as the contract Π^A is publicly observable and contractible between principal and intermediary. Payouts to the principal are $dw_t^P = \beta_t^P(dX_t - \bar{a}_t^A dt - \bar{a}_t^I dt)$, and payouts to the agent are $dw_t^A = \max\{v_t^A - \overline{v}, 0\}$, with endogenous payout boundary $\overline{v}$. As shown in Parts IV and V of the proof, the sensitivities $\beta_t^A, \beta_t^I, \beta_t^P$ that solve the maximization in (4.80) satisfy (4.81), (4.82) and (4.83). The incentive condition (4.12) map sensitivities β_t^A, β_t^I to effort levels a_t^A, a_t^I. In what follows, we verify that the proposed strategy (i.e., the proposed contracts) from Proposition 16 yields higher payoff than any other strategy (i.e., any other incentive compatible contracts) and thus is indeed optimal.

Take any time $t < \tau$. Suppose that the intermediary deviates from the proposed strategy and follows an alternative strategy up to time t (i.e., for times $s \leq t$), with sensitivities $\beta_s^A, \beta_s^I, \beta_s^P$, payouts to the principal dw_s^P, and payouts to the agent $dw_s^A \geq 0$. After time t (i.e., for times $s \geq t$), the intermediary follows the proposed strategy (contracts) from Proposition 16. Then, the intermediary's payoff at time 0 (under the deviation) can be written as

$$G_t = \int_0^t e^{-rs}\left(dX_s - dw_s^P - dw_s^A - \frac{\lambda(a_s^I)^2}{2}ds - \frac{\theta^I r(\beta_s^I\sigma)^2}{2}ds\right) + e^{-rt}[F(v_t^A) - v_t^P]. \quad (4.84)$$

We use (4.10), (4.16) and Itô's Lemma and differentiate G_t with respect to time, t, to get

$$\begin{aligned} e^{rt}dG_t = &\Bigg\{(\hat{a}_t^A + a_t^I) - \frac{\theta^I r(\beta_t^I\sigma)^2}{2} - \frac{\lambda(a_t^I)^2}{2} \\ &+ F'(v_t^A)\left(\frac{\theta^A r(\beta_t^A\sigma)^2}{2} + \frac{\delta(\hat{a}_t^A)^2}{2}\right) + \frac{F''(v_t^A)(\beta_t^A\sigma)^2}{2} - rF(v_t^A)\Bigg\}dt \\ &+ \underbrace{(1 - \beta_t^P + F'(v_t^A)\beta_t^A)}_{=\beta_t^I}[dX_t - \hat{a}_t^A dt - a_t^I dt] - dw_t^A(F'(v_t^A) + 1) \\ = &\ \mu_t^G dt + \beta_t^I[dX_t - \hat{a}_t^A dt - a_t^I dt] + dw_t^A(F'(v_t^A) - 1). \end{aligned}$$

Note that the incentive condition (4.12) and the observability (and contractibility) of Π^A imply $a_t^I = \hat{a}_t^I = \bar{a}_t^I = \beta_t^I/\lambda$ and $a_t^A = \hat{a}_t^A = \bar{a}_t^I = \beta_t^A/\delta$.

By (4.80), the drift term in curly brackets, μ_t^G is zero when β_t^A, β_t^I, and β_t^P solve the optimizaton in (4.80) subject to the incentive compatibility constraint (4.12) and the limited liability constraints.[19] As shown in Parts IV and V of this proof, this is the case if and only β_t^A, β_t^I, and β_t^P satisfy (4.81), (4.82) and (4.83). Any other choice of β_t^A, β_t^I, and β_t^P makes the drift term μ_t^G weakly negative, in that $\mu_t^G \leq 0$. In addition, observe that because $F'(v_t^A) \geq -1$ with equality if and only $v_t^A = \overline{v}$, the term $-dw_t^A(F'(v_t^A) + 1)$ is weakly negative under any payout policy $dw_t^A \geq 0$ and zero under the proposed payout policy $dw_t^A = \max\{v_t^A - \overline{v}, 0\}$.

Next, recall that $\frac{dX_t - \hat{a}_t^A dt - a_t^I dt}{\sigma}$ is the increment of a standard Brownian motion under the probability measure $\mathbb{P}^I$ that is taken under the intermediary's information. Because the sensitivity β_t^I is by assumption bounded, it follows that

$$\mathbb{E}_0^I\left[\int_0^t e^{-rs}\beta_s^I\left(dX_s - \hat{a}_s^A ds - a_s^I ds]\right)\right] = 0.$$

[19] Recall that it is optimal to set $dw_t^P = \beta_t^P(dX_t - \bar{a}_t^A dt - \bar{a}_t^I dt)$. As a result, payouts to the principal affect the intermediary's optimization only via the sensitivity β_t^P.

Thus, G_t, with

$$dG_t = e^{-rt}\left(\underbrace{\mu_t^G dt}_{\leq 0} \underbrace{-dw_t^A(F'(v_t^A)+1)}_{\leq 0}\right) + e^{-rt}\beta_t^I[dX_t - \hat{a}_t^A dt - a_t^I dt],$$

follows a super-martingale (i.e., decreases in expectation) under the measure $\mathbb{P}^I$. Thus,

$$G_0 \geq \mathbb{E}_0^I G_t \tag{4.85}$$

Because $a_t^j \in [0, A]$ with $A < \infty$ for $j = A, I$, it follows that total surplus is bounded from above by $2A/r$. As such, $F(v_t^A) - v_t^P$ is bounded from above by $2A/r$ and bounded from below by 0 (due to limited liability) so that

$$0 \leq \lim_{t\to\infty} e^{-rt}[F(v_t^A) - v_y^P] \leq \lim_{t\to\infty} e^{-rt}\frac{2A}{r} = 0. \tag{4.86}$$

Thus, (4.84) implies

$$\lim_{t\to\infty} \mathbb{E}_0^I G_t = \mathbb{E}_0^I\left[\int_0^\tau e^{-rs}\left(dX_s - dw_s^P - dw_s^A - \frac{\lambda(a_s^I)^2}{2}ds - \frac{\theta^I r(\beta_s^I\sigma)^2}{2}ds\right)\right] = \mathbb{E}^I G_\tau$$

Next, we take the limit $t \to \infty$ in (4.85) to obtain

$$\underbrace{G_0}_{=F(v_0^A)-v_0^P} \geq \mathbb{E}_0^I G_\tau = \mathbb{E}_0^I\left[\int_0^\tau e^{-rs}\left(dX_s - dw_s^P - dw_s^A - \frac{\lambda(a_s^I)^2}{2}ds - \frac{\theta^I r(\beta_s^I\sigma)^2}{2}ds\right)\right]. \tag{4.87}$$

The inequality (4.87) implies that the proposed contracts Π^A, Π^P from Proposition 16 are indeed optimal and solve the intermediary's problem, as (4.87) holds in equality under the proposed contracts and the intermediary's expected payoff at time 0 under the proposed contracts is $G_0 = F(v_0^A) - v_0^P$.

4.12.5 Part VIII — Liquidation

Identical to Part VII of the proof of Proposition 15. That is, the intermediary optimally liquidates the firm at $v^A = 0$ when the agent's contract is terminated, if $R \geq R_L$ with R_L defined in (4.76).

4.13 Robustness and extensions

4.13.1 Dynamic delegation

This section derives the model solution and HJB equation under dynamic delegation. In what follows, we denote (with a slight abuse of notation) the value function by $F(v^A)$ too. By the dynamic programming principle, incentives β^A and β^I are chosen to maximize the sum of expected cash flows $a^A + a^I$ and the expected change in value $\frac{\mathbb{E}dF(v^A)}{dt}$, adjusted for risk and effort costs. This leads to the maximization problem

$$\max_{\beta^A,\beta^I}\left\{a^A + a^I - \frac{\lambda(a^I)^2}{2} - \frac{\theta^I r(\beta^I\sigma)^2}{2} + \frac{\mathbb{E}dF(v^A)}{dt}\right\} \tag{4.88}$$

that is subject to the effort incentive constraints (4.12) and additionally to the constraint (4.27), if and only if the agent's prescribed effort and incentives are not observable to the principal.

Thus, direct contracting is optimal in state v^A if and only if $B^*(v^A) - \eta \geq A^*(v^A)$ where

$$A^*(v^A) \equiv \max_{\beta^A,\beta^I} \left\{ a^A + a^I - \frac{\lambda(a^I)^2}{2} - \frac{\theta^I r(\beta^I\sigma)^2}{2} + \frac{\mathbb{E}dF(v^A)}{dt} \right\} \text{ s.t. (4.12) and (4.27)}$$

$$B^*(v^A) \equiv \max_{\beta^A,\beta^I} \left\{ a^A + a^I - \frac{\lambda(a^I)^2}{2} - \frac{\theta^I r(\beta^I\sigma)^2}{2} + \frac{\mathbb{E}dF(v^A)}{dt} \right\} \text{ s.t. (4.12).}$$

We can expand the term $\frac{\mathbb{E}dF(v^A)}{dt}$, using Itô's Lemma, and solve for the intermediary's and agent's optimal incentives,

$$\beta^I = \begin{cases} \frac{(1+F'(v^A)\pi^I)\left(\frac{1+F'(v^A)\pi^I}{\lambda}+\frac{\pi^I}{\delta}\right)}{\left(\frac{1}{\lambda}+\theta^I r\sigma^2\right)(1+F'(v^A)\pi^I)^2-F'(v^A)(\pi^I)^2\left(\frac{1}{\delta}+\theta^A r\sigma^2\right)-F''(v^A)(\pi^I\sigma)^2} & \text{if } B^*(v^A)-\eta < A^*(v^A) \\ \frac{1}{1+\lambda\theta^I r\sigma^2} & \text{if } B^*(v^A)-\eta \geq A^*(v^A) \end{cases} \tag{4.89}$$

and

$$\beta^A = \begin{cases} \frac{\pi^I\left(\frac{1+F'(v^A)\pi^I}{\lambda}+\frac{\pi^I}{\delta}\right)}{\left(\frac{1}{\lambda}+\theta^I r\sigma^2\right)(1+F'(v^A)\pi^I)^2-F'(v^A)(\pi^I)^2\left(\frac{1}{\delta}+\theta^A r\sigma^2\right)-F''(v^A)(\pi^I\sigma)^2} & \text{if } B^*(v^A)-\eta < A^*(v^A) \\ \frac{1}{-(\delta\theta^A r\sigma^2+1)F'(v^A)-\delta\sigma^2 F''(v^A)} & \text{if } B^*(v^A)-\eta \geq A^*(v^A). \end{cases} \tag{4.90}$$

Having characterized the optimal controls (incentives) β^A, β^I, the value function $F(v^A)$ solves the ODE (under the optimal controls)

$$\begin{aligned} rF(v^A) = \Bigg\{ & a^A + a^I - \frac{\lambda(a^I)^2}{2} - \frac{\theta^I r(\beta^I\sigma)^2}{2} - \eta\mathbb{I}\{B^*(v^A) - \eta \geq A^*(v^A)\} \\ & + F'(v^A)\left(rv^A + \frac{\theta^A r(\beta^A\sigma)^2}{2} + \frac{\delta(a^A)^2}{2}\right) + \frac{F''(v^A)(\beta^A\sigma)^2}{2}\Bigg\}, \end{aligned} \tag{4.91}$$

where $\mathbb{I}\{\cdot\}$ denotes the indicator function which equals one if the argument $\{\cdot\}$ is true and equals zero otherwise. The optimal values of β^A and β^I are characterized in (4.89) and (4.90). The effort incentive constraint (4.12) maps the sensitivities β^A, β^I to effort levels a^A, a^I. As in the baseline version of the model, the boundary conditions $F(0) = R$ (liquidation), $F'(\overline{v}) = -1$ (smooth pasting), and $F(\overline{v}) = R$ (limited liability) apply.

4.13.2 Liquidation value

We discuss the model solution without the assumption $R \geq R_L$. For this purpose, we consider that $R < R_L$. Note that the intermediary derives expected payoff $F(0) = R_L$ from operating the firm without the agent (for details see Appendix 4.10.8) from time $\tau = \inf\{t \geq 0 : v_t^A = 0\}$. Also observe that at time τ, the agent's contract is terminated, so that $\beta_t^A = dw_t^A = a_t^A = v_t^A = 0$ for all $t \geq \tau$. The reason is that when $v_t^A = 0$, then the intermediary must eliminate the volatility of v_t^A and set $\beta_t^A = 0$ to ensure that the limited liability constraint $v_t^A \geq 0$ is met. As such, when v_t^A hits zero at time $t = \tau$, it remains forever at this point, in that $v^A = 0$ is an absorbing state and the agent's contract is terminated when $v^A = 0$.

As $R < R_L$, $F(0) = R_L > R$ and the intermediary does not liquidate the firm at time τ. Instead, the intermediary continues to operate the firm without the agent, yielding payoff R_L. As a result, the only element that changes in the solution relative to the baseline with $R \geq R_L$ is the boundary condition of $F(v^A)$ at $v^A = 0$, which becomes $F(0) = R_L$ when $R < R_L$. If, on the other hand, $R \geq R_L$, the intermediary prefers liquidation over running the firm without the agent, leading to $F(0) = R$. As a result, we can write compactly

$$F(0) = \max\{R, R_L\},$$

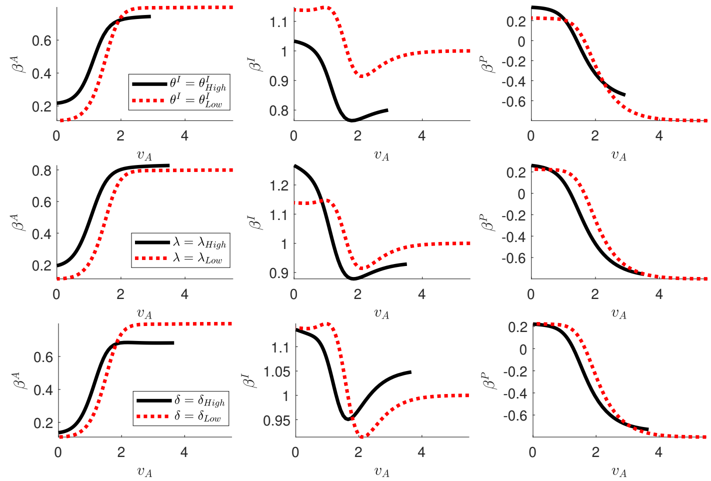

FIGURE 4.5: Model solution for different levels of the intermediary's risk-aversion θ^I, the intermediary's cost of effort λ, and the agent's cost of effort δ. We use $\theta^I_{Low} = 0$, as in the baseline, and $\theta^I_{High} = \theta^A = 5$; $\lambda_{High} = 1.2 > \lambda_{Low} = 1$; $\delta_{High} = 1.2 > \delta_{Low} = 1$.

which nests the baseline model (with $R \geq R_L$) as a special case. Because the exact value of $F(0)$ is not crucial for the properties of the model solution, the value function, and the optimal controls, we do not expect any qualitative changes in model outcomes and implications when $R < R_L$.

4.13.3 Intermediary risk aversion and agency costs

Figure 4.5 plots the optimal incentives against v^A for different levels of the intermediary's risk-aversion θ^I, the intermediary's cost of effort λ, and the agent's cost of effort δ. The figure suggests that the dynamics of β^A, β^I, and β^P described in the previous section are robust to changes in these parameters. That is, the agent's incentives β^A increase in v^A and the intermediary's incentives β^I as well as the risk-sharing with the principal β^P are high (low) when v^A is low (high) and the trickle-up effect generates additional incentives (disincentives). Nevertheless, a few notable observations are in order.

First, although the intermediary is risk-averse, there is no perfect risk-sharing (i.e., $\beta^P < 1$) and the intermediary even takes additional risk via $\beta^P < 0$ to upkeep his incentives when v^A is large. In addition, for low values of v^A, an increase in the intermediary's risk aversion $\theta^I > 0$ strengthens the risk-sharing motive and increases β^P, decreasing the intermediary's incentives β^I but increasing the agent's incentives β^A. The reason is that an increase in θ^I reduces the slope of the value function $F'(v^A)$ for small values of v^A, which reduces trickle-up incentives. As trickle-up incentives undermine trickle-down incentives (see term D in (4.25)), trickle-down incentives (i.e., π^I) increase, enhancing the agent's incentives.

Second, an increase in λ reduces the negative impact of the trickle-up effect on trickle-down incentives (see term D in (4.25)), thereby improving the agent's incentives β^A especially for low values of v^A. In addition, higher λ leads to stronger incentives to the intermediary for low values of v^A and to weaker incentives for high values of v^A.

Third, an increase in δ has relatively little impact on the agent's and the intermediary's incentives for low values of v^A yet increases the intermediary's incentives and decreases the agent's incentives for high values of v^A.

4.13.4 Financially constrained intermediary

In this section, we solve the model for the case when the intermediary is financially constrained, in a sense that $dw_t^I = dX_t - dw_t^A - dw_t^P \geq 0$ at all times $t > 0$. It is possible for the intermediary to inject funds at time $t = 0$ at the initial round of financing (i.e., $dw_0^I < 0$ is possible).

Recall that the intermediary is residual claimant on total firm value and, as payouts to the principal dw_t^P are not sign-restricted, the intermediary can extract all surplus from the principal (representing competitive outside investors). As such, it is equivalent whether the intermediary chooses contracts Π^A, Π^P to dynamically maximize his own payoff v_t^I or the joint payoff $v_t^I + v_t^P$. Like in the baseline model, we can express all model quantities as functions of v^A, so that v^A is the only payoff-relevant state variable, and we omit time subscripts whenever no confusion is likely to arise. Thus, the optimal contracts Π^A, Π^P dynamically maximize the value function $v^I + v^P = F(v^A)$.

To solve the model with the constraint $dw_t^I \geq 0$ for $t > 0$, we first solve the relaxed problem without this constraint, which is akin to solving the baseline model. Thus, the value function $F(v^A)$ solves the HJB equation (4.28), subject to all relevant incentive constraints. The maximization in (4.28) yields the same values for the optimal controls $\beta^A, \beta^I, \beta^P$ (and a^A, a^I) as in the baseline. When $v^A = 0$, the agent's contract is terminated and the firm is liquidated, leading to $F(0) = R$. The agent is paid at the payout boundary $\overline{v}$, so that $dw^A = \max\{v^A - \overline{v}, 0\}$. At the payout boundary, the smooth pasting condition $F'(\overline{v}) = -1$ applies. While the limited liability constraint in the baseline requires that $F(v^A) - v^P \geq R$, the limited liability constraint under the alternative formulation requires $F(v^A) \geq R$. Like in the baseline, the limited liability constraint is (optimally) tight at the payout boundary, leading to $F(\overline{v}) = R$.

Note that while it is optimal in the baseline to set $v^P = 0$ to relax the limited liability constraint $F(v^A) - v^P \geq R$, the limited liability constraint is unaffected by the exact choice of $v^P = 0$ under the alternative specification. In addition, the HJB equation (4.28) is independent of v^P or dw^P, so that the choice of v^P and hence the choice of dw^P is not payoff-relevant. In what follows, we determine v^P or equivalently $v^I = \hat{F}(v^A) = F(v^A) - v^P$ to achieve $dw^I \geq 0$ after time zero, without affecting payoff $F(v^A)$. To achieve $dw_t^I \geq 0$ at all times $t > 0$, two requirements must be satisfied at all times $t > 0$: i) the volatility of dw^I, denoted $vol(dw^I) = \sigma_w(v^A)$, must be zero, and ii) the drift of dw^I, denoted $\mu_w = \mu_w(v^A)$, must be positive.

To move forward, we postulate that

$$dw^I = \mu_w(v^A)dt + \sigma_w(v^A)dZ + \xi_w dw^A,$$

where $\sigma_w(v^A) = 0$, $\mu_w(v^A) \geq 0$, and $\xi_w \geq 0$. As $dw^I = dX - dw^A - dw^P$, we can solve

$$dX - dw^A - dw^P = \mu_w(v^A)dt + \sigma_w(v^A)dZ + \xi_w dw^A$$

and obtain

$$dw^P = dX - dw^A(\xi_w + 1) - \mu_w(v^A)dt - \sigma_w(v^A)dZ, \tag{4.92}$$

determining the principal's payouts, dw^P.

First, we characterize under what conditions the volatility of dw^I, i.e., $\sigma_w(v^A)$, is zero. For this sake, it suffices to consider that $v^A < \overline{v}$ (the behavior at the boundary and ξ_w are discussed below), so that $dw^A = 0$. Recall that by (4.16), v^I evolves according to

$$dv^I = rv^I dt + \frac{\lambda}{2}(a^I)^2 dt + \frac{\theta^I r}{2}\lambda(\beta^I\sigma)^2 dt - dw^I + \beta^I \sigma dZ, \tag{4.93}$$

where we already impose that in optimum prescribed and actual effort coincide so that $\sigma dZ = dX - \hat{a}^A dt - \hat{a}^I dt$. Note that a^I and β^I are determined by the HJB equation (4.28), and the incentive conditions. As $vol(dw^I) = \sigma_w(v^A) = 0$, it follows from (4.93) that the volatility of v^I, denoted $vol(dv^I)$, is $vol(dv^I) = \beta^I\sigma$. On the other hand, Itô's Lemma implies that the volatility of $v^I = \hat{F}(v^A)$ satisfies

$$vol(dv^I) = \hat{F}'(v^A)vol(dv^A) = \hat{F}'(v^A)\beta^A\sigma. \tag{4.94}$$

Note that (4.16) implies that $vol(dv^A) = \beta^A\sigma$. As a result, we obtain that

$$\hat{F}'(v^A) = \frac{\beta^I}{\beta^A} = \frac{\beta^I(v^A)}{\beta^A(v^A)}, \tag{4.95}$$

where the choice of $\beta^A = \beta^A(v^A), \beta^I = \beta^I(v^A)$ is determined by (4.28) and depends on the state v^A.

Note that the intermediary receives a lumpy payout when the agent is paid, that is, if and only if $v^A \geq \overline{v}$. At the payout boundary $\overline{v}$, $\hat{F}'(\overline{v}) = \beta^I(\overline{v})/\beta^A(\overline{v})$. Thus, for each dollar the agent is paid, the intermediary is paid $\xi_w = \beta^I(\overline{v})/\beta^A(\overline{v})$ dollars, in that $dw^I/dw^A = \xi_w = \beta^I(\overline{v})/\beta^A(\overline{v})$.

Using (4.95), one obtains that the function $\hat{F}(v^A)$ takes the form

$$\hat{F}(v^A) = K + \hat{F}_0(v^A), \tag{4.96}$$

for some constant K. The function $\hat{F}_0(v^A)$ solves (4.95) subject to $\hat{F}_0(0) = 0$ and is therefore unique. Also note that this implies $v^P = F(v^A) - K - \hat{F}_0(v^A)$.

Second, the drift of dw^I, denoted $\mu_w = \mu_w(v^A)$, must be positive, $\mu_w(v^A) \geq 0$. As $\mu_w \geq 0$, it follows by (4.93) that the drift of $v^I = \hat{F}(v^A)$, denoted μ_v, satisfies

$$\mu_v(v^A) \geq rv^I + \frac{\lambda}{2}(a^I)^2 + \frac{\theta^I r}{2}\lambda(\beta^I\sigma)^2. \tag{4.97}$$

On the other hand, Itô's Lemma (and (4.16), imply that the drift of $v^I = \hat{F}(v^A)$ satisfies

$$\mu_v(v^A) = \hat{F}'(v^A)\left(rv^A + \frac{\theta^A r(\beta^A\sigma)^2}{2} + \frac{\delta(a^A)^2}{2}\right) + \frac{\hat{F}''(v^A)(\beta^A\sigma)^2}{2}, \tag{4.98}$$

where a^A and β^A are determined by the HJB equation (4.28). We combine (4.97) and (4.98) to get

$$rv^I + \frac{\lambda}{2}(a^I)^2 + \frac{\theta^I r}{2}\lambda(\beta^I\sigma)^2 \geq \hat{F}'(v^A)\left(rv^A + \frac{\theta^A r(\beta^A\sigma)^2}{2} + \frac{\delta(a^A)^2}{2}\right) + \frac{\hat{F}''(v^A)(\beta^A\sigma)^2}{2}. \tag{4.99}$$

As a next step, we use (4.96) and $\hat{F}'(v^A) = \hat{F}_0'(v^A)$ and $\hat{F}''(v^A) = \hat{F}_0''(v^A)$ and rewrite (4.99) as

$$\begin{aligned} &r(\hat{F}_0(v^A) + K) + \frac{\lambda}{2}(a^I)^2 + \frac{\theta^I r}{2}\lambda(\beta^I\sigma)^2 \\ \geq &\hat{F}_0'(v^A)\left(rv^A + \frac{\theta^A r(\beta^A\sigma)^2}{2} + \frac{\delta(a^A)^2}{2}\right) + \frac{\hat{F}_0''(v^A)(\beta^A\sigma)^2}{2}. \end{aligned} \tag{4.100}$$

Recall that $\hat{F}_0(v^A)$ solves (4.95) subject to $\hat{F}_0(0) = 0$ and is therefore unique.

Observe that the left-hand-side of (4.100) increases with K, while the right-hand-side of (4.100) is independent of K. Thus, we can always find a constant K such that (4.100) holds for all $v^A \in [0, \overline{v}]$ and therefore $dw^I \geq 0, \mu_w \geq 0 = \sigma_w$ hold after time zero. For instance, one can set $K = K^*$ with

$$K^* = \inf\{K \geq 0 : \text{(4.100) holds for all } v^A \in [0, \overline{v}]\}.$$

Optimal payouts to the principal dw^P are then defined as the residual that implements $v^P = F(v^A) - \hat{F}(v^A)$, taking into account the law of motion (4.10). Specifically, optimal payouts to the principal are characterized in (4.92).

Importantly, when $v_0^P < 0$ at time zero, then the intermediary makes an initial lump-sum payment of $-dw_0^I > 0$ dollars to the principal (i.e., $dw_0^I < 0$), so that the principal breaks even. After time zero, the intermediary's payouts $dw^I \geq 0$ are no more negative. The intermediary receives β^I/β^A dollars per dollar the agent is paid. As desired, we have determined the intermediary's and the principal's compensation $(\hat{F}(v^A), v^P)$ and as such payouts to the intermediary and principal (dw^I, dw^P) to ensure $dw^I \geq 0$ after time zero, without affecting the optimal contracts and the value function $F(v^A)$.

Chapter 5

Optimal Financing with Tokens[1]

5.1 Introduction

Initial coin offerings (ICOs) have become an important source of financing for firms that develop digital platforms (Howell et al., 2020). By the end of 2018, over 5500 firms had attempted to raise funds using an ICO, raising over 30 billion dollars (Lyandres et al., 2019) and with at least 20 ICOs taking in more than 100 million dollars (Howell et al., 2020). In an ICO, a firm raises funds by issuing cryptographically secured tokens. Because these tokens serve as the means of payment on a platform or offer access to the firm's services, they possess *utility features* and are therefore often called *utility tokens*. Despite the popularity of ICOs and the considerable growth of the academic literature on this new form of financing, a number of key questions remain open. Chief among these is whether an ICO should be preferred to alternative ways of financing, such as financing with equity or with tokens other than utility tokens.

Tokens indeed come in many different forms. Many tokens only possess utility features and do not have any security features, such as cash flow or dividend rights. This is the case for example for the tokens issued in the ICOs of Filecoin or Golem. Symmetrically, several tokens—such as the LDC Crypto token or the BCAP token—do not possess utility features and resemble traditional securities, except that they are recorded and exchanged on a blockchain. Tokens with security features are classified by the so-called Howey Test as securities.[2] These tokens are called *security tokens* and sold in security token offerings (STOs). Remarkably, many tokens exhibit both utility and security features. For instance, multiple crypto-exchanges—such as Binance, BitMax, or KuCoin—feature tokens that are used to trade on the exchange and additionally allow token holders to earn income related to the overall transaction volume.[3] Digital banking platforms—such as Nexo or Bankera—have issued tokens of a similar type. Likewise, cryptocurrencies with proof-of-stake consensus algorithms—such as NEO, Cardano, or Ethereum after its Casper Protocol—both facilitate transactions and generate income to token holders.[4]

This paper develops a unifying model that nests these different types of tokens and studies the optimal token design in the presence of frictions that generally prevail in firms developing digital platforms, such as the need to raise outside funds to finance platform development and the ensuing agency conflicts between insiders (platform developers) and outsiders. Specifically, we develop a model in which a startup firm, owned by penniless developers, builds a platform that facilitates P2P transactions among users. As in Cong, Li, and Wang (2020a,b), the platform features network effects that imply complementarities in

[1]This Chapter is based on Gryglewicz, Mayer, and Morellec (2021).

[2]According to the Howey test, an investment contract is a security if the following four conditions hold: 1) it is an investment of money, 2) in a common enterprise, 3) with an expectation of profit, 4) with profit generated by a third party. Conditions 1 and 2 are typically satisfied for any type of token offering. Conditions 3 and 4 are satisfied for example if the token distributes dividends.

[3]While Binance distributes profits to token holders through buybacks (i.e., token burning), KuCoin and BitMax explicitly pay dividends to token holders. In addition, transacting with the native exchange token offers fee discounts.

[4]Token holders are rewarded for staking (i.e., holding) tokens. Ethereum will switch to a proof-of-stake consensus algorithm after the so-called Casper Protocol (Buterin and Griffith, 2017) is implemented.

users' endogenous adoption and transaction decisions. In addition, the platform generates cash flows that increase in the level of platform adoption (i.e., the platform transaction volume) and arise, e.g., from transaction fees, advertisement proceeds, and/or from utilizing transaction data.

Entities conducting token offerings tend to have unproven business models and are most often in the pre-product stage (Howell et al., 2020). To capture these key features, we consider that the platform is initially not fully developed and that the startup firm has financing needs in that developers lack the funds to finance platform development. To raise the necessary funds, the startup can issue equity and/or tokens that may serve as the transaction medium on the platform and thus may exhibit *utility* features. These tokens may also exhibit *security* features, in that they may pay dividends in relation to platform cash flows. In addition to financing needs, platform development is subject to moral hazard. Specifically, platform success depends on developers' hidden effort, which comes at a cost to developers.

In the model, developers' revenues stem from selling tokens to platform users and from the ownership of the startup equity, which is a claim on the cash flows that the platform generates. Users' motive to hold tokens and the pricing of tokens reflect both the token utility and security features, with an equilibrium token price that increases with the level of platform adoption. Token security features affect users' platform adoption decisions and, therefore, the value of the platform and of its native tokens. Because they grant cash flow rights to token holders, security features also reduce the value of developers' equity in the startup firm and undermine their incentives, which are determined by their equity ownership and by the tokens they retain. Crucially, equity and token incentives not only differ in their strength but also in their relationship with the token design. The paper solves for the optimal token design in this environment characterized by financing needs and moral hazard and derives the following main findings.

First, considering the financing problem of a platform that uses tokens as transaction medium, we demonstrate that issuing tokens to finance the startup generally maximizes both developers' payoff and the value of the platform. We also show that while token security features trigger endogenous network effects and spur platform adoption, their provision is affected by moral hazard and financing needs. Specifically, granting cash flow rights to token holders improves the return to holding tokens and therefore boosts the platform transaction volume. This, in turn, raises the platform's cash flows, which implies even more transactions and dividends. However, token security features dilute developers' equity ownership in the startup firm. Because the incentives generated by each dollar of equity ownership are stronger than the incentives from a dollar of token ownership, token security features undermine incentives. As a result, the optimal level of cash flow rights granted to token holders decreases in the extent of moral hazard. Since the under-provision of security features reduces platform adoption and value, moral hazard intensifies financing constraints. Symmetrically, larger financing needs imply that developers retain fewer tokens, thereby exacerbating the moral hazard problem. Financing needs and moral hazard thus reinforce each other, leading to low levels of token security features and token retention. We also show that moral hazard is more severe when network effects are low or the platform development phase is long, which induces low levels of security features and token retention.

Second, we analyze when issuing a utility token without security features is optimal. That is, we analyze when developers prefer an initial coin offering over a security token offering. An ICO is the optimal funding model if the platform value derives from facilitating transactions rather than from generating cash flows. An ICO is also preferable to a security token offering if financing needs, agency frictions, or the platform development phase are large. Thus, while the ICO funding model is often criticized on the basis that many firms have not yet delivered on their product, our analysis suggests to the contrary that projects with a long development phase are particularly suitable for conducting an ICO. Moreover, our model implies that startups with innovative business models, which are particularly prone to moral hazard, optimally raise funds via ICOs, consistent with Fahlenbrach and Frattaroli (2019) or Howell et al. (2020).

Third, we examine when using fiat money as the platform transaction medium and issuing equity to finance platform development is optimal. Ceteris paribus, the ability to transact with fiat money reduces the cost of transacting for users and increases both the transaction volume and platform earnings. Intuitively, users are more willing to transact with fiat money as they do not bear crypto-related transaction costs. However, issuing tokens without utility features (or security tokens that resemble conventional equity) may constrain developers' ability to raise funds and harm platform success, notably, when platform value mostly comes from facilitating transactions among users. Financing platform development with equity is therefore only optimal if platform cash flows are expected to be large or if network effects are strong. For firms without very high cash flows (or very strong network effects), the platform is generally optimally financed with tokens, unless moral hazard is severe or financing needs are large.

Fourth, we study the asset pricing implications of token utility and security features. We show that while token security features spur platform adoption, they also amplify token price volatility. The reason is that security features generate endogenous network effects that increase the sensitivity of platform adoption to productivity shocks. This boosts the token price volatility because the token derives its value from the level of platform adoption. The effects of security features on the token price volatility are larger when the token possesses more utility features or when network effects are stronger. Thus, according to our model, the combination of token utility and security features should cause particularly volatile token prices.

Finally, we study various extensions of the model, in particular the relation between optimal platform financing and adverse selection. We demonstrate that adverse selection has ambiguous effects on the provision of token security features, depending on whether a separating or pooling equilibrium prevails. In a separating equilibrium, in which different types of platforms are financed with different types of tokens and ICOs and STOs coexist, adverse selection increases the provision of token security features by high-quality platforms, implying a positive relation between the provision of security features and the ex post value of platforms or the likelihood of platform success. In a pooling equilibrium in which all platforms are financed with the same tokens, adverse selection decreases the provision of token security features.

Our work is related to the literature on blockchain economics, tokenomics, and cryptocurrencies. Notable contributions include Athey, Parashkevov, Sarukkai, and Xia (2016), Abadi and Brunnermeier (2018), Makarov and Schoar (2019), Liu and Tsyvinski (2019), Cong and He (2019), Cao, Cong, and Yang (2019), Huberman, Leshno, and Moallemi (2019), Biais, Bisiere, Bouvard, and Casamatta (2019), Cong, He, and Li (2020), Prat and Walter (2019), Saleh (2019), Pagnotta (2020), Easley, O'Hara, and Basu (2019), and Hinzen, John, and Saleh (2019). A review of this rapidly evolving research area is provided by Chen, Cong, and Xiao (2019).

A large subset of this literature focuses on ICOs with many empirical papers studying determinants of ICO success or documenting post-ICO patterns. Important contributions include Howell et al. (2020), Fahlenbrach and Frattaroli (2019), and Lyandres et al. (2019). Many firms issuing tokens develop a decentralized platform that promises network effects. Much of the theoretical literature on ICOs highlights the coordination benefits inherent to utility tokens; see e.g. Li and Mann (2020), Sockin and Xiong (2020), and Catalini and Gans (2018). Further theories on ICOs include Chod and Lyandres (2019), Chod, Trichakis, and Yang (2019), Goldstein, Gupta, and Sverchkov (2019), Holden and Malani (2019), Lee and Parlour (2018), Lyandres (2019), Malinova and Park (2017), and Mayer (2020b). In contrast to these papers, our model is not limited to utility tokens but encompasses a richer class of tokens. In addition, we study the effects of financing needs and moral hazard on token design, while most research to date takes the token and platform design as exogenously given. Li and Mann (2019) provide a review of the early literature on ICOs.

Our paper also advances the literature on the economics of platforms. Early contributions in this literature, such as Rochet and Tirole (2003), do not consider tokens. More recently, important progress has been made on platform finance with tokens. Notably, Cong et al. (2020b) analyze the pricing implications of users' inter-temporal adoption decisions. Cong et al. (2020a) connect tokenomics to corporate finance, with a focus on optimal token-supply policy to finance investment in platform quality. While we employ similar modeling of users' platform adoption decisions, our paper differs from Cong et al. (2020a,b) in several important dimensions. First, Cong et al. (2020a,b) do not consider tokens with dividend rights and security features. Second, while Cong et al. (2020a) features conflicts of interest between users and developers, they abstract from moral hazard and platform financing needs which are the key frictions we model in this paper.

Finally, our paper also relates to the literature on the optimal design of securities. Seminal contributions include Townsend (1979), Gale and Hellwig (1985), and Bolton and Scharfstein (1990), or, in dynamic settings, DeMarzo and Sannikov (2006) and DeMarzo and Fishman (2007b). Our focus is on the design of tokens and the comparison of tokens with equity financing. A distinguishing feature of our framework is that platform financing (i.e., the design of tokens) affects endogenous platform adoption, cash-flows, and firm value, even if there are no frictions. By contrast, in standard models of security design, such a link between firm value and financing requires frictions (such as adverse selection, moral hazard, or taxes).

Section 5.2 presents the model. Section 5.3 solves for the optimal token design when the platform uses tokens as transaction medium. Sections 5.4 analyze the model implications. Section 5.5 derives conditions under which equity financing with fiat money as transaction medium is optimal. Section 5.6 examines the asset pricing implications of token utility and security features. Section 5.7 investigates the robustness of our findings to various model extensions. Section 5.8 summarizes our main testable predictions. Section 5.9 concludes. All proofs are gathered in the Appendix.

5.2 Baseline model

Time is continuous and defined over $[0, \infty)$. There are two types of agents: developers and a unit mass of platform users indexed by $i \in [0,1]$. All individuals are risk neutral and discount future payoffs at rate $r > 0$. Developers run a startup firm that launches a digital platform but lack the capital to develop it. They obtain funds at time zero by issuing tokens. Tokens serve as the transaction medium on the platform. They are in fixed unit supply and possess equilibrium price P_t. In addition, they are perfectly divisible, reflecting the fact that crypto tokens can generally be traded in fractional amounts. We conjecture and verify that token-based financing always dominates equity financing *when tokens serve as transaction medium*. In particular, developers (optimally) do not issue outside equity and always own 100% of the startup's equity.

Platform transactions. The platform allows users to conduct peer-to-peer transactions. As in Cong et al. (2020a,b), any user i has transaction needs and derives a utility flow

$$A_t N_t^{\chi} \frac{x_{it}^{\eta}}{\eta} \tag{5.1}$$

from a transaction of x_{it} dollars on the platform where $\eta \in (0,1)$.[5] The coefficient A_t is the platform productivity, which characterizes the usefulness of the platform. The specification in (6.1) captures network effects in that any user's utility from transacting increases in the volume of platform transactions N_t. That is, the higher the transaction volume, the easier it is to find a transaction counter-party and the more valuable it becomes to join the platform. The parameter $\chi \in [0, 1-\eta)$ characterizes the strength of these network effects.

Transacting on the platform is costly. First, any user has to hold x_{it} dollars in tokens (or x_{it}/P_t tokens) for vdt units of time in order to transact.[6] Holding tokens is therefore costly because it implies a foregone opportunity to invest and earn interest for vdt units of time. The parameter $v > 0$ captures potential delays in settlements, in acquiring tokens, or

[5] This utility flow can be micro-founded by a random search and matching protocol; see Cong et al. (2020a,b).

[6] Appendix 5.13 provides a micro-foundation for this holding period. Cong et al. (2020a,b) assume that $v = 1$. When $v = 0$, security tokens have no transaction value and resemble conventional equity.

in finding an appropriate counter-party.[7] Second, in addition to these holdings costs, users incur direct costs $\phi x_{it} dt$ for a transaction of size x_{it} on the platform, where $\phi > 0$. This direct cost captures for instance transaction fees charged by miners or crypto-exchanges or a physical cost of platform operation that is charged to users. This direct cost may also be related to the effort and attention required for transacting on the platform, as in Cong et al. (2020b).

Cash flows. Once developed, the platform generates cash flows

$$dD_t = \mu(A_t) N_t dt, \tag{5.2}$$

where $\mu(A_t)$ with $\frac{\partial \mu(A)}{\partial A} \geq 0$ is the platform cash flow rate. In practice, platforms may generate cash flows with advertisement proceeds, transaction fees, and/or by selling/using user data. Naturally, cash flows increase with the transaction volume N_t and platform productivity A_t as a more useful platform implies a higher user activity on both the extensive and intensive margins, which in turn raises the profits that platform operators extract, e.g., by setting (per-transaction) fees or selling user data. For analytical tractability, we assume that cash flows are linear in the transaction volume N_t and that there is no direct link between $\mu(A_t)$ and ϕ in that $\frac{\partial \mu(A)}{\partial \phi} = 0$. Under this assumption, the cost ϕ is a dead-weight loss as in, e.g., Cong et al. (2020b). Section 5.7 incorporates endogenous transaction fees charged by platform developers to users and analyzes their effects on token design and platform adoption and value.

Platform development: Moral hazard and financing. Firms conducting token offerings are young and most often in the pre-product stage (Howell et al., 2020). To capture this feature, we consider that the platform is developed over some time period $[0, \tau)$ and launched at time τ once a milestone has been reached. The arrival time of the milestone τ is governed by a Poisson process M_t with constant intensity Λ, so that over each time interval of length dt there is a probability Λdt that the platform development is complete and the expected time to development is $\frac{1}{\Lambda}$.

Platform development is subject to moral hazard and financing needs. Moral hazard arises because platform success depends on developers' hidden effort $a_t \in \{0, 1\}$ which comes against a flow cost κa_t to developers, with $\kappa \geq 0$.[8] Specifically, in case the milestone is reached over the time interval $[t, t + dt)$, the platform is successful only if developers exert effort over $[t, t + dt)$. Formally, we have that $A_s = 0$ for $s < \tau$ and

$$A_s = A_L + (A_H - A_L) \mathbf{1}_{\{a_\tau = 1\}}$$

for $s \geq \tau$, where developers have to choose effort a_t before the random event $dM_t \in \{0, 1\}$ realizes over $[t, t + dt)$. This modelling of productivity shocks is also employed in, e.g., Board and Meyer-ter Vehn (2013) and Hoffmann and Pfeil (2021). It follows that moral hazard is severe when the cost of effort κ or the expected time to development $1/\Lambda$ is large. Define $\mu_j = \mu(A_j)$ for $j \in \{H, L\}$. Fig. 5.1 shows the timing of events over a time interval $[t, t + dt)$. For simplicity, platform productivity is constant after time τ. We study the implications of productivity shocks arising after time τ in Section 5.6 and show that this assumption has no bearing on our key findings.

In addition to moral hazard, the startup firm faces financing needs in that platform development requires investing $I > 0$ and developers do not have the capital to cover these needs. At inception, developers thus sell $1 - \beta_0$ tokens to the market and raise $(1 - \beta_0) P_0$ dollars. Funds raised by issuing tokens must be sufficient to cover the financing needs of the

[7]Because, in practice, blockchain protocol and settlement latency (Easley et al., 2019; Hautsch, Scheuch, and Voigt, 2019) limit the influence that developers have on v, we treat it as an exogenous parameter. For example, transactions on the Bitcoin blockchain cannot occur instantaneously since a new block has to be created for the transaction settlement which takes, on average, 10 minutes.

[8]Section 5.7 shows that this setup is isomorphic to a model with cash diversion. We thank the referee for encouraging us to generalize the model in this direction.

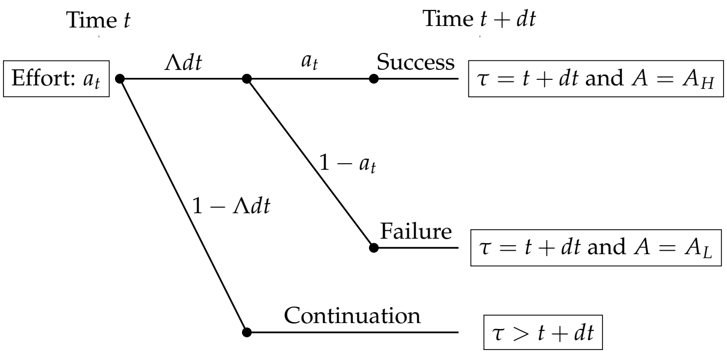

FIGURE 5.1: Heuristic Timing over $[t, t + dt)$. With probability Λdt platform development is complete. Platform success depends on developers' hidden effort $a_t \in \{0, 1\}$.

firm, leading to the constraint:

$$(1 - \beta_0)P_0 \geq I. \tag{5.3}$$

Funds raised at time zero are optimally invested in platform development or paid out as dividends.

Developers have incentives to exert effort because they hold tokens and own the firm's equity. While providing the funds to finance platform development, token issuance also leads to a potential dilution of developers' stake in the firm, triggering moral hazard. Notably, developers initially retain $\beta_0 \in [0, 1]$ tokens that are only optimally sold when the milestone is reached at time τ.[9] That is, developers sell $1 - \beta_0$ tokens at time zero and β_0 tokens at time τ. We emphasize that we do not restrict developers to this particular token trading behavior. Because the only productivity shock realizes at time τ and developers and users discount at the same rate r, there is simply no reason to trade at any other time $t \notin \{0, \tau\}$. We therefore denote the developers' token holdings β_t over $[0, \tau)$ by β.

Security features. Besides having utility features by serving as the platform transaction medium, tokens may also have security features in that they may pay a fraction $\alpha \in [0, 1]$ of total cash flows dD_t to token holders, with the balance $(1 - \alpha)dD_t$ being paid out as a dividend to the startup equity holders. Therefore, even though developers own 100% of the startup's equity, a token with $\alpha > 0$ dilutes their cash flow rights and the value of their equity ownership in the startup.

In summary, token utility features are represented by the convenience yield in (6.1). Token security features are captured by the token dividends αdD_t. In practice, the Howey test would classify any token with cash-flow rights as security, so that we refer to tokens with $\alpha > 0$ as security tokens. Conversely, when $\alpha = 0$, the token is a utility token and does not possess security features. That is, our model encompasses initial coin offerings as a special case in which $\alpha = 0$. There is an ongoing debate on whether utility tokens are securities. In classifying tokens as securities, our paper follows recent practice. In particular, tokens without cash flow rights (i.e., with $\alpha = 0$) are typically not classified as securities and, therefore, referred to as utility tokens.

Users' adoption decisions. Before the platform is developed at time τ, tokens do not offer transaction benefits. Consequently, tokens are fairly priced by ordinary risk-neutral

[9]Section 5.7 introduces speculators in the model and convex effort costs and shows that, in this alternative setup, developers may continuously trade between inception and the milestone. It also shows that this extension has no other bearings on our results. We thank the referee for encouraging us to generalize the model in this direction.

investors/users implying that expected capital gains, $\mathbb{E}[dP_t]$, and dividends, $\mathbb{E}[\alpha\, dD_t]$, alone offer investors the required return, $rP_t\, dt$:

$$\mathbb{E}[dP_t + \alpha\, dD_t] = rP_t\, dt \quad \text{for} \quad t < \tau. \tag{5.4}$$

After time τ, the platform is developed and holding x_{it}/P_t tokens over a time period of length $v\, dt$ generates additional transaction benefits and costs:

$$dR_{it} := \underbrace{A_t N_t^{\chi} \frac{x_{it}^{\eta}}{\eta} dt}_{\text{Convenience yield}} - \underbrace{x_{it}\phi\, dt}_{\text{Transaction cost}} + v x_{it}\left(\underbrace{\frac{dP_t}{P_t}}_{\text{Capital gains}} + \underbrace{\frac{\alpha\, dD_t}{P_t}}_{\text{Dividend yield}} - \underbrace{r\, dt}_{\text{Funding cost}} \right). \tag{5.5}$$

Eq. (5.5) shows that by holding tokens and transacting on the platform, users realize both a convenience yield and capital gains. A transaction of size x_{it} comes at an effective cost $(vr + \phi)x_{it}$ that consists of the (funding) costs of holding tokens and the direct transaction costs.

The optimal transaction volume x_{it} for user i maximizes the expected utility flow at each point in time:

$$\max_{x_{it} \geq 0} \mathbb{E}[dR_{it}].$$

This yields an optimal transaction volume given by

$$x_{it}^{1-\eta} = \frac{A_t P_t N_t^{\chi} dt}{\phi P_t dt + v\big(rP_t dt - \mathbb{E}[dP_t + \alpha\, dD_t]\big)}. \tag{5.6}$$

All users $i \in [0,1]$ face the same trade-off when determining their optimal transaction volume. We thus have $N_t = \int_0^1 x_{it} di = x_{it}$ so that the transaction volume at time $t \geq \tau$ satisfies

$$N_t = \left(\frac{A_t}{vr - v\mathbb{E}[dP_t + \alpha dD_t]/(P_t dt) + \phi} \right)^{\frac{1}{1-\xi}}, \tag{5.7}$$

where we define for convenience $\xi := \chi + \eta$ as the transformed network effects parameter. A higher value of N_t means that each user is more active on the platform. As a result, the transaction volume N_t captures the degree of *platform adoption* at time t.

Developers' problem. Developers choose effort a_t, their token holdings β_t, and the cash flow rights α attached to tokens. When tokens possess cash flow rights, developers receive $1 - \alpha + \beta_t \alpha$ dollars for each dollar of cash flows produced by the firm. Developers can sell their initial allocation of tokens at the prevailing market price. Accordingly, their optimization problem can be written as

$$V_0 = \max_{\alpha, \{\beta_t\}, \{a_t\}} \mathbb{E}\left[\int_0^{\infty} e^{-rt}(-P_t\, d\beta_t + (1 - \alpha + \alpha\beta_t) dD_t - \kappa a_t dt) \right] - I, \tag{5.8}$$

subject to the financing constraint (5.3).

5.3 Equilibrium and model solution

We study a *Markov Perfect Equilibrium*.

Definition 1 *In a Markov Perfect Equilibrium, the following conditions must be satisfied:*

1. *All individuals act optimally: Users maximize*

$$w_i := \max_{\{x_{it}\}} \mathbb{E}\left[\int_0^{\infty} e^{-rt} dR_{it} \right] \tag{5.9}$$

and developers solve (5.8).

2. *The token market clears before the milestone in that (5.4) is satisfied for all $t < \tau$.*
3. *The token market clears after the milestone in that:*

$$\frac{v}{P_t}\left(\int_0^1 x_{it}di\right) = \frac{vN_t}{P_t} \leq 1 - \beta_t, \tag{5.10}$$

for all $t \geq \tau$. If and only if the inequality (5.10) is strict, (5.4) holds for $t \geq \tau$.

The left-hand side in the market clearing condition (5.10) represents the token demand for transaction reasons. The right hand side represents the token supply. The token demand for transaction reasons is the product of the transaction volume N_t/P_t measured in units of tokens and the duration of the token holding period v. Intuitively, if v is large, users from previous transaction periods need to hold tokens in the current period, thereby increasing demand. (Appendix 5.13 discusses market clearing in more detail.) Lastly, if the token demand for transaction reasons is below the token supply, tokens must be held solely for their dividend rights and their price is determined by (5.4).

In the following, we solve the model before and after the milestone separately for any choice of $\alpha \in [0,1]$ and $\beta \in [0,1]$. Based on the model solution, we then determine the optimal level of security features α and the optimal level of token retention β.

5.3.1 Model solution after the milestone

We solve the model for any outcome $j \in \{H, L\}$. Since all uncertainty is resolved after time τ, it follows that all quantities remain constant at levels $X_j = X_t$ for all $t \geq \tau$ for $X \in \{P, N\}$ in that $dP_t = 0$ for all $t \geq \tau$. We incorporate uncertainty after time τ in Section 5.6, where we discuss the asset pricing implications of token utility and security features.

Because of their utility benefits, tokens are more valuable for users than for developers after time τ. In addition, there is no moral hazard problem once the platform has been launched. As a result, there is no value for developers in retaining tokens after time τ. Thus, developers sell all retained tokens at time τ so that $\beta_t = 0$ for all $t \geq \tau$. This implies that the value of developers' stake in the startup firm at time τ is equal to the value $\frac{(1-\alpha)\mu_j N_j}{r}$ of the startup's equity, where the price and transaction levels remain constant at levels P_j and N_j respectively.

Next, we derive the token price. Users may hold tokens for transaction purposes and/or because of their dividends. If dividends $\alpha\mu N_j$ exceed the funding cost rP_j, users hold tokens purely for investment motives and the token price is given by the present value of its dividend rights:

$$P_j = \frac{\alpha\mu_j N_j}{r}.$$

In this case, the token is priced according to its *security features*. Otherwise, the token is held for transaction purposes and priced according to its *utility features*. In this case, N_j/P_j tokens are held over a period of length vdt and the effective token demand over a short period of time $[t, t+dt)$ is then given by vN_τ/P_τ. Token supply for $t \geq \tau$ is given by $1 - \beta_t = 1$. Market clearing therefore implies that:

$$P_j = vN_j.$$

Combining the two cases, we obtain that the user base in Eq. (6.10) simplifies to

$$N_t = N_j(\alpha) = N_j = \left(\frac{A_j}{\max\{0, vr - \alpha\mu_j\} + \phi}\right)^{\frac{1}{1-\xi}} \tag{5.11}$$

and the token price is given by

$$P_t = P_j = \begin{cases} v\left(\frac{A_j}{vr-\alpha\mu_j+\phi}\right)^{\frac{1}{1-\xi}} & \text{if } vr > \alpha\mu_j \\ \frac{\alpha\mu_j}{r}\left(\frac{A_j}{\phi}\right)^{\frac{1}{1-\xi}} & \text{if } vr \le \alpha\mu_j. \end{cases} \tag{5.12}$$

Equations (5.11) and (5.12) reveal that utility features determine the token price if and only if the opportunity cost vr of holding tokens exceed the token dividend yield $\alpha\mu_j$. Thus, in our framework, v and α determine the users' underlying motive to hold tokens. When v is relatively large compared to α, users hold tokens over an extended time period mainly for transaction purposes and, therefore, because of their utility features. By contrast, if v is low compared to α—for instance, when fiat money can be used as transaction medium on the platform and $v = 0$—tokens are only held for their cash flow rights and their price increases with α.

Last, the token price and platform adoption are closely related to the *token velocity*, defined as the ratio of the platform's real transaction value over the token market capitalization. In our model, it is given by:

$$velocity := \frac{N_t}{P_t} = \min\left\{\frac{1}{v}, \frac{r}{\alpha\mu(A_t)}\right\}.$$

This equation shows that if the token is priced according to its utility features, token velocity equals the inverse of the holding period v. Remarkably, security features $\alpha > 0$ bound the token velocity from above and so can be useful to address problems associated with high token velocity.[10]

5.3.2 Model solution before the milestone

Incentive compatibility

Consider first developers' incentives to maximize the platform's transaction value through their effort choice a_t. Suppose developers exert effort, so that $a_t = 1$. With probability Λdt the milestone arrives over the next time interval and future platform productivity equals A_H so that developers' payoff at time τ equals:

$$\beta P_H + \frac{(1-\alpha)\mu_H N_H}{r},$$

which is the sum of the value of the tokens they retain and the present value of future cash flows. By contrast, if developers shirk and choose $a_t = 0$, future platform productivity becomes A_L and their payoff at time τ equals:

$$\beta P_L + \frac{(1-\alpha)\mu_L N_L}{r}.$$

Hence, developers exert effort at any time $t < \tau$ (i.e., $a_t = 1$) if and only if:

$$IC(\alpha) := \underbrace{\Lambda\left(\beta P_H + \frac{(1-\alpha)\mu_H N_H}{r}\right)}_{\text{Payoff under } a_t=1} - \kappa - \underbrace{\Lambda\left(\beta P_L + \frac{(1-\alpha)\mu_L N_L}{r}\right)}_{\text{Payoff under } a_t=0} \ge 0. \tag{5.13}$$

Developers' incentives to exert effort are driven by the tokens they retain and their equity stake in the startup firm. Token-based incentives are captured by the retention level β. Equity incentives are captured by the fraction of the platform cash flows $1-\alpha$ accruing to the startup's owners.

[10] The token velocity problem is widely discussed among crypto-practicioners (see e.g. https://www.coindesk.com/blockchain-token-velocity-problem).

Developers' problem and initial token issuance

Consider next the platform development phase $[0, \tau)$. Unless otherwise mentioned, we assume that platform development costs I are not prohibitively large and that full effort is optimal. In addition, we set $\mu_L = 0$ for analytical convenience, so that the platform produces cash flows if and only if developers exert effort. These assumptions are gathered in the following:

Assumption 2 *Exerting effort is efficient in that the project produces cash flows and has positive net present value (NPV) if and only if $a_t = 1$ for all $t < \tau$. Formally,* (5.27), (5.28), *and* (5.29) *in Appendix 5.10 have to be met.*

When Assumption 1 is satisfied and developers exert effort, we have $P_t = P_H$ after the milestone has been reached. The fair price of the token for risk-neutral users over $[0, \tau)$ is then given by:

$$P_t = P_0 = \frac{\Lambda P_H}{r + \Lambda}. \tag{5.14}$$

Notably, absent further constraints, developers and users value tokens equally before time τ as they both apply the same discount rate. However, because a higher retention level β relaxes condition (5.13), developers issue the minimal amount of tokens needed to finance platform development. We thus have for $\beta = \beta_0$:

$$(1 - \beta)P_0 = I \Longleftrightarrow \beta = 1 - \frac{I}{P_0}. \tag{5.15}$$

Developers optimally do not sell tokens over $(0, \tau)$ as there are simply no gains from trade, so that $\beta_t = \beta$ for $t \in [0, \tau)$.

Upon reaching the milestone, developers sell all retained tokens β at price P_H and further enjoy the perpetual dividend stream $(1 - \alpha)N_H\mu_H$. Hence, their continuation value over $(0, \tau)$ conditional on full effort is:

$$V(\alpha) = \frac{1}{r + \Lambda}\left[\Lambda\left(\beta P_H + \frac{(1 - \alpha)\mu_H N_H}{r}\right) - \kappa\right] \tag{5.16}$$

with $\beta = 1 - \frac{I}{P_0}$. We can rewrite the value function as:

$$V(\alpha) = \frac{\Lambda S(\alpha) - \kappa}{r + \Lambda} - I, \tag{5.17}$$

where

$$S(\alpha) = P_H + \frac{(1 - \alpha)\mu_H N_H}{r}. \tag{5.18}$$

Eq. (5.17) is the net present value of the project to developers, which is given by the value of the platform net of the investment cost. In this equation, $S(\alpha)$ is the sum of the value of all tokens in circulation and the value of the startup equity after τ. Therefore, $S(\alpha)$ captures the monetary *platform value* after time τ, i.e., the overall surplus in dollar terms. In Eq. (5.18), P_H is the value of all tokens (i.e., the token market capitalization) while $(1 - \alpha)\mu_H N_H$ is the dividend flow.

At time zero, developers design the token and choose the optimal level of dividend rights α to maximize the value they extract from the platform. That is, developers solve

$$\max_{\alpha \in [0,1]} V(\alpha) \text{ s.t. } (5.13) \text{ and } (5.3). \tag{5.19}$$

Using Eq. (5.17), we thus have that developers maximize $S(\alpha)$ subject to the incentive constraint (5.13) and the financing constraint (5.3). We conclude the section by establishing the existence of an equilibrium with positive adoption.[11]

[11] Note that there are other degenerate equilibria in which no user adopts the platform and the platform and tokens are worthlesss. Throughout the paper, we do not direct our attention to these degenerate, less interesting equilibria.

Proposition 17 (Equilibrium existence) *There exists a Markov-Perfect Equilibrium with positive, maximal adoption after the milestone in that $N_t = N_H$, $\forall t \geq \tau$. In this equilibrium:*

1. *Developers' value function is given by* (5.17) *for $t < \tau$ and equals zero for $t \geq \tau$.*
2. *Developers sell tokens only at times 0 and τ and the retention level is given by* (5.15) *and $\beta_t = 0$, $\forall t \geq \tau$. The optimal level of security features α is characterized by* (5.8).
3. *The token price is characterized by* (5.12) *for $t \geq \tau$ and by* (5.14) *for $t < \tau$.*

5.4 Analysis

5.4.1 The frictionless benchmark

We start by studying the model without moral hazard. In this frictionless benchmark, the incentive compatibility constraint (5.13) becomes irrelevant, and developers choose α to maximize the platform value $S(\alpha)$. This holds true even if $I > 0$. The following Proposition demonstrates that, absent agency conflicts and transaction costs (i.e., for low ϕ), full dividend rights $\alpha = 1$ are optimal when token are priced according to utility features (i.e., when $vr \geq \mu_H$). It also shows that an increase in transaction costs generally reduces the optimal amount of security features.

Proposition 18 (Frictionless benchmark) *Define $\hat{\alpha} = \min\left\{1, \frac{vr}{\mu_H}\right\}$. When there is no moral hazard ($\kappa = 0$), developers maximize $S(\alpha)$ and choose $\alpha = \bar{\alpha}$ with $\bar{\alpha} = \arg\max_\alpha S(\alpha)$ satisfying:*

$$\bar{\alpha} = \begin{cases} x \in [\hat{\alpha}, 1] & \text{if } \max\{vr - \mu_H, 0\} \geq \frac{(1-\xi)\phi - \mu_H}{\xi} \\ 0 & \text{if } vr \leq \frac{(1-\xi)\phi - \mu_H}{\xi} \\ \frac{vr}{\mu_H} + \frac{1}{\xi} - \frac{\phi(1-\xi)}{\xi \mu_H} & \text{otherwise.} \end{cases}$$

When $\bar{\alpha} \geq \hat{\alpha}$, it holds that $S'(\alpha) \geq 0$ for $\alpha \in [0, \hat{\alpha})$.

Because developers and users discount at the same rate r, they also value dividends—ceteris paribus—the same. However, dividends paid to users rather than to developers increase the returns to holding tokens and spur transaction volume and adoption. This in turn boosts cash flows and, as a result, dividends to token holders and adoption. That is, security features induce *endogenous* network effects via the cash flow channel. Therefore, absent frictions it is optimal to allocate full cash flow rights to users when token are priced according to utility features and the cost of transacting on the platform is small and does not represent an impediment to platform development. Unlike α, the parameter v has ambiguous effects on platform adoption and token prices. An increase in v raises the cost of transacting for users, hampering platform adoption in that $\frac{\partial N_H}{\partial v} < 0$. At the same time, an increase in v may boost the token price due to the market clearing condition $P_H = vN_H$, which holds when tokens are priced according to utility features (i.e., when $vr \geq \mu_H$).

Throughout, we focus on environments in which the cost of transacting ϕ is low and the token is priced according to its utility features. That is, unless otherwise mentioned, we assume that:

Assumption 3 *Parameters satisfy*

1. $vr > \mu_H$.
2. $vr > \frac{(1-\xi)(\phi - \mu_H)}{\xi}$.

The first condition ensures that the token is priced according to its utility features (i.e., $P_H = vN_H$). The second condition implies that $\bar{\alpha} = 1$ (see Proposition 18). As we show below, all frictions drive the level of security features below $\bar{\alpha} = 1$, so this choice can be viewed as a normalization.

5.4.2 Moral hazard and financing needs

As shown by (5.19), developers maximize $S(\alpha)$ subject to the incentive constraint (5.13) and the financing constraint (5.3). Since $\bar{\alpha} = 1$, $S'(\alpha) \geq 0$ for all $\alpha \in [0,1]$ (see Proposition 18 with $\bar{\alpha} = \hat{\alpha} = 1$), and (5.3) is optimally tight as in (5.15), developers choose the maximal value α that satisfies the incentive constraint (5.13). Therefore, the optimal level of security features α in the tokens issued by the startup firm is given by:

$$\max_{\alpha \in [0,1]} \alpha \;\text{ s.t. }\; IC(\alpha) \geq 0.$$

We can now examine how the optimal level of security features α and the token retention level β depend on moral hazard and financing needs. In the absence of financing needs, i.e., when $I = 0$, developers retain all tokens and are therefore able to capture all the monetary proceeds that the platform generates. As a result, even if $\kappa > 0$, there are no agency conflicts in that developers maximize $S(\alpha)$ and choose $\alpha = \bar{\alpha} = 1$. Conversely, financing needs $I > 0$ lead to a lower token retention level $\beta < 1$ and give rise to agency conflicts between developers (insiders) and users (outsiders). These agency conflicts affect the optimal design of tokens and therefore platform value, which in turn determines the severity of the financing frictions. In the following, we analyze how moral hazard and financing needs jointly shape the design of tokens and the provision of incentives.

When $\alpha < 1$ and $\beta > 0$, developers have both equity-based incentives and token-based incentives. Equity-based incentives primarily relate to platform cash flows. Token-based incentives primarily relate to platform adoption. Because a higher platform adoption also leads to higher cash flows, equity-based incentives de facto generate payoff sensitivity to both platform adoption and cash flows. More formally, observe that for any given α, the value of developers' equity before time τ satisfies

$$E(A) = \frac{\Lambda}{r+\Lambda}\frac{(1-\alpha)N(A)\mu(A)}{r}$$

where the second term on the right hand side of this equation represents the value of equity at time τ. Here, $N(A)$ is the level of platform adoption as a function of A and $\mu(A)N(A)$ is the platform's cash-flow (also written as function of A). This implies that the incentives (i.e., the sensitivity with respect to productivity A) generated by a dollar of equity ownership are equal to

$$\frac{dE/dA}{E} = \underbrace{\frac{d\mu/dA}{\mu}}_{\substack{\text{Cash flow}\\ \text{sensitivity}}} + \underbrace{\frac{dN/dA}{N}}_{\substack{\text{Sensitivity}\\ \text{to platform}\\ \text{adoption}}}, \tag{5.20}$$

whereas the incentives from a dollar token ownership — owing to $P = vN$ — are equal to

$$\frac{dP/dA}{P} = \frac{dN/dA}{N}. \tag{5.21}$$

Equations (5.20) and (5.21) show that, as long as cash flows increase with platform productivity, equity incentives are stronger than token-based incentives. Because of their greater strength, equity incentives are particularly important in firms characterized by severe moral hazard. Therefore, incentives optimally become more equity-based and less token-based if the cost of effort (i.e., κ), the expected time to platform development (i.e., $1/\Lambda$), or financing needs (i.e., I) increase. That is, financing and agency frictions or a long platform development phase lead to an under-provision of token security features. The provision of equity incentives reduces the token price P_H and thus requires developers to sell more tokens at inception to cover financing costs I, thereby reducing the token retention level β.

Fig. 5.2 illustrates these findings by plotting the optimal level of token security features α and developers' retention level β as functions of financing needs I, the expected time to platform development $1/\Lambda$, and agency frictions κ. Input parameter values for this figure are described in Appendix 5.10. They follow from prior contributions in the literature and

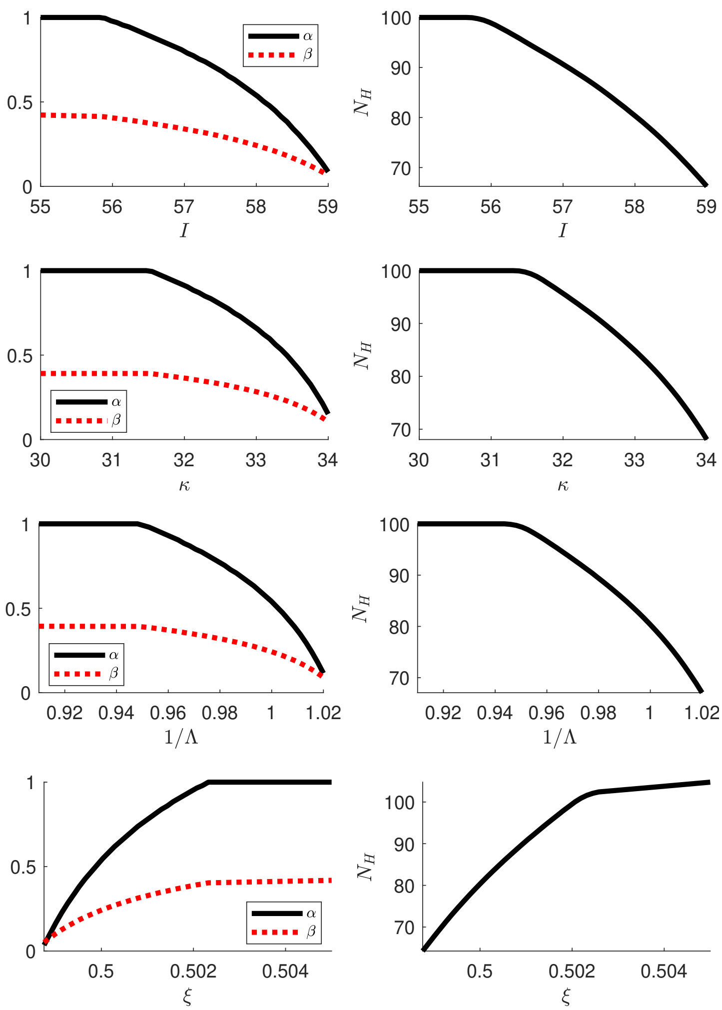

FIGURE 5.2: The effects of financing needs I, moral hazard κ, expected time to platform development $1/\Lambda$, and network effects ξ on token design and platform adoption.

imply an optimal retention level of $\beta = 39\%$ in our base case environment, in line with the average retention level reported in Fahlenbrach and Frattaroli (2019). The right panels of Fig. 5.2 demonstrate the effects of agency and financing frictions on platform adoption when the token is optimally designed; as discussed above a decrease in security features leads to a decrease in platform adoption.

Remarkably, network effects ξ relax the incentive condition (5.13). The intuition is that strong network effects make developers' revenues more contingent on platform adoption, thereby aligning users' and developers' incentives. In addition, stronger network effects make it more valuable to grant dividend rights to token holders as security features lead to higher cash flows to token holders and boost adoption, which triggers even higher cash flows and adoption. These endogenous network effects arising from the cash flow channel are amplified by the exogenous network effects ξ. As a result, stronger network effects imply more token-based incentives, i.e., a higher retention level β, and less equity incentives $1-\alpha$ to developers as illustrated by Fig. 5.2.

Finally, we also demonstrate that it is strictly sub-optimal for developers to raise funds by issuing equity next to (transaction) tokens. This is for two reasons. First, because equity incentives are stronger than token incentives, selling equity to outside investors exacerbates moral hazard, which is costly when either $\kappa, 1/\Lambda$, or I is sufficiently large. Second, while the high cash-flow rights α attached to tokens spur platform adoption, they also reduce the startup firm's equity value and preclude a financing via equity. These two mechanisms make it optimal to bundle transaction benefits and cash-flow rights in (i.e., attach utility and security features to) *one* security rather than offering *two* securities that deliver dividends and transaction benefits separately.

The following proposition gathers our analytical results.

Proposition 19 (Optimal token financing) *The following holds:*

1. *Optimal token security features are given by $\alpha = \bar{\alpha}$ if either κ, $1/\Lambda$, or I is sufficiently small.*
2. *The optimal level of security features satisfies $\frac{d\alpha}{dI} \leq 0$, $\frac{d\alpha}{d(1/\Lambda)} \leq 0$ and $\frac{d\alpha}{d\kappa} \leq 0$, where the inequalities are strict only if the incentive condition* (5.13) *is tight.*
3. *The optimal token retention level satisfies $\frac{d\beta}{dI} < 0$, $\frac{d\beta}{d(1/\Lambda)} > 0$ and $\frac{d\beta}{d\kappa} \leq 0$, where the latter inequality is strict only if the incentive condition* (5.13) *is tight.*
4. *If $A_H > vr + \phi$, then $\frac{\partial IC(\alpha)}{\partial \xi} > 0$ for any $\alpha \in [0,1]$, $\frac{d\alpha}{d\xi} \geq 0$, and $\frac{d\beta}{d\xi} > 0$. The former inequality is strictly only if the incentive condition* (5.13) *is tight.*
5. *Raising funds by issuing equity next to tokens is strictly sub-optimal.*

5.4.3 ICO vs. STO: When to include token security features?

In our framework, the token does not exhibit security features when $\alpha = 0$. In this case, the token derives its value only from its transaction benefits. Such tokens are generally referred to as utility tokens and issued in a lightly regulated way by means of an initial coin offering (see Howell et al., 2020; Fahlenbrach and Frattaroli, 2019). The following proposition establishes that whether an ICO is preferred to a security token offering (STO) depends on platform characteristics.

Proposition 20 (ICOs vs. STOs) *An initial coin offering (i.e., $\alpha = 0$) is optimal if $\mu_H \leq (1-\xi)\phi - \xi vr$. A security token offering (i.e., $\alpha > 0$) is optimal if*

1. *$\mu_H > (1-\xi)\phi - \xi vr$,*
2. *either κ, $1/\Lambda$, or I is sufficiently small.*

The comparison between an ICO and STO can be conducted for a fixed platform value since the statements in Proposition 20 do not explicitly involve A_H. The inequality conditions in Proposition 20 imply that when the platform is expected to generate low (or even negative) cash flows (i.e., for low μ_H), the ICO financing model is optimal. In this case,

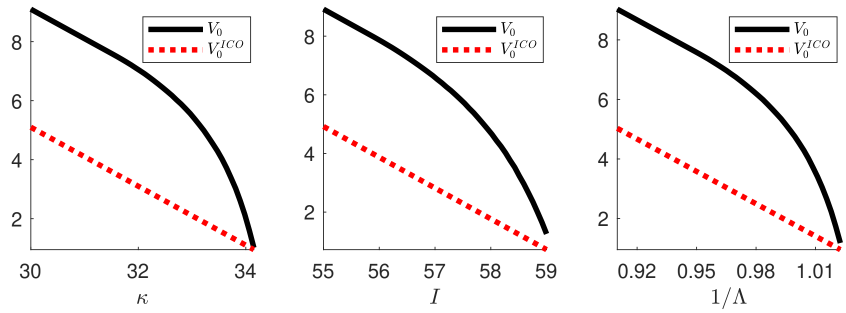

FIGURE 5.3: ICO versus STO. V_0 and V_0^{ICO} represent the developers value under the optimal security token ($\alpha > 0$) and utility token ($\alpha = 0$), respectively.

the platform essentially derives its value from facilitating transactions among users. By contrast, the platform's ability to generate cash flows adds value to security token offerings even though the issuance of a security token dilutes developers' cash-flow rights. The economic mechanism behind this result is that granting cash-flow rights to users spurs platform adoption, which is particularly valuable for large μ_H. Proposition 20 also implies that stronger network effects (i.e., high ξ) favor STOs. This is because cash-flow rights embedded in security tokens magnify network effects and spur platform adoption even more.

Proposition 20 also demonstrates that financing and agency frictions make STOs less attractive. This is because an increase in frictions renders equity incentives more valuable, thereby reducing the value of security tokens relative to utility tokens. Likewise, projects with long expected times to development are subject to more severe moral hazard and, therefore, are more suitable for ICO financing. Interestingly, several empirical studies (see, e.g., Howell et al., 2020; Fahlenbrach and Frattaroli, 2019) report that many ICO financed projects have not yet delivered their promised product. While this fact is often interpreted as evidence for the failure of the ICO financing model, our analysis suggests the opposite in that projects with longer expected times to completion especially benefit from ICO financing.

Fig. 5.3 illustrates these results. Our baseline parameter values are such that STOs dominate ICOs. Under severe financing and agency frictions, it may become optimal to issue a pure utility token via an ICO to avoid diluting developers' equity stake.

5.5 Equity vs. tokens: When to include token utility features?

We have worked so far under the assumption that tokens serve as the platform transaction medium. Instead, platform developers can decide to use fiat money as transaction medium. Doing so removes utility features from tokens. It also eliminates the token holding period vdt, thereby reducing users' effective transaction costs. When $v = 0$, the adoption level and token price satisfy

$$N_H = \left(\frac{A_H}{\phi}\right)^{\frac{1}{1-\xi}} \quad \text{and} \quad P_H = \frac{\alpha \mu_H N_H}{r},$$

which shows that using fiat money as transaction medium potentially spurs adoption and that this effect is stronger when platform network effects are stronger.

Without token utility features, the token price is the present value of the dividend stream to token holders and the token essentially represents an equity claim. This implies that token and equity incentives are equivalent so that the choice of α becomes irrelevant. The overall surplus is then given by:

$$S(\alpha) = P_H + \frac{(1-\alpha)\mu_H N_H}{r} = \frac{\mu_H N_H}{r},$$

which is just the platform expected dividend stream and is independent of the choice of α. In the analysis below, we examine when equity financing (in combination with fiat money as transaction medium) dominates token financing (in combination with tokens as transaction medium).

In general, platform developers can choose between attaching utility features to tokens, i.e., setting $v^* = v$, or omitting them, i.e., setting $v^* = 0$, where the parameter v is given and exogenously fixed, e.g., due to technological constraints. As a result, developers' optimization problem reads

$$V_0 = \max_{\alpha, v^* \in \{0,v\}, \{\beta_t\}, \{a_t\}} \mathbb{E}\left[\int_0^\infty e^{-rt}(-P_t\, d\beta_t + (1-\alpha+\alpha\beta_t)dD_t - \kappa a_t dt)\right] - I. \tag{5.22}$$

Proposition 21 derives conditions under which equity financing is optimal.

Proposition 21 (Optimality of equity financing) *Issuing equity to finance platform development and using fiat money instead of tokens as platform transaction medium is optimal if and only if*

$$\frac{\mu_H}{vr} \geq \left(\frac{\phi}{vr - \mu_H + \phi}\right)^{\frac{1}{1-\xi}}, \tag{5.23}$$

and always leads to a higher level of adoption. Otherwise, a token-based platform is optimal. Condition (5.23) *is satisfied if v or ξ are sufficiently large or if $\mu_H \in [\phi(1-\xi), vr)$ is sufficiently large. A token-based platform is optimal if*

1. *Condition* (5.23) *is not satisfied,*
2. *either κ, $1/\Lambda$, or I is sufficiently small.*

Proposition 21 shows that token financing is optimal if network effects or platform cash flows are not very high. In these instances, the issuance of a token that serves as a platform transaction medium allows developers to raise more funds, which mitigates financing frictions and contributes to platform success. Conversely, a token without utility features is optimal only if the platform cash flows are high. In this case, the startup uses fiat money as a transaction medium and is financed with equity. If, in addition, network effects are strong, reducing transaction costs by allowing users to transact with fiat money boosts adoption. As expected, using fiat money as a transaction medium becomes also optimal if the cost v of transacting with tokens is large.[12]

Issuing equity to finance platform development and using fiat money as a transaction medium rather than tokens with utility features can be optimal for firms with lower cash flows if financing needs are large and agency frictions are severe. To understand this finding, note that a fiat-based platform implies that developers' value fully stems from their equity ownership in the startup firm. Hence, developers' incentives are equity-based and therefore stronger, which is particularly valuable if financing needs are large and moral hazard is severe. In line with this reasoning, equity financing (or a token without utility features) is preferred for large values of κ, $1/\Lambda$, or I.

Fig. 5.4 illustrates these findings by plotting the optimal financing choice of the platform for different levels of cash flows and frictions. In both panels, the platform has negative NPV for combinations of parameter values below the solid black line, so that it cannot be financed. In both panels equity financing is always preferred when cash flows are very high (area above the dashed red line). For firms without high cash flows, the platform is generally

[12] Blockchain-based tokens may offer some transaction benefits over fiat money. These could be related to security, privacy, or reliability. Remarkably, in our model, tokens can become optimal even under the assumption that cash carries strictly lower transaction costs.

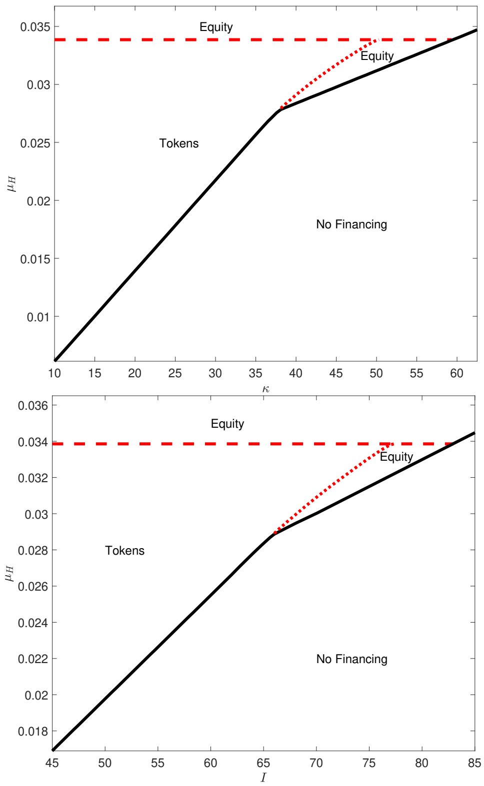

FIGURE 5.4: Token-financing versus equity financing. The figure plots the optimal financing strategy for platform developers for different combinations of cash flows μ_H and agency frictions κ (left panel) or financing needs I (right panel).

financed with tokens unless frictions (moral hazard κ or financing needs I) are very high, in which case it is financed with equity (top right corner). As frictions decrease, financing with tokens becomes optimal. Financing with tokens can even be optimal for platforms that do not generate cash flows (or very low cash flows) if the value of the transactions conducted by users is sufficiently large (bottom left corner).

Finally, according to Proposition 19, it is not optimal to raise funds by issuing equity next to tokens when the latter are used as transaction medium. According to Proposition 21, it can be optimal to issue equity instead of tokens to finance platform development and use fiat money as platform transaction medium. As a result, financing with a mix of equity and tokens is not optimal, in that the startup firm optimally finances platform development by issuing either tokens or equity.

5.6 Productivity shocks and token price volatility

We now allow for uncertainty after the milestone by introducing persistent productivity shocks that are not affected by developers' actions. Importantly, the introduction of productivity shocks after time τ does not qualitatively affect developers' decisions before time τ, in that the results derived above continue to hold (see Appendix 5.12.2). Our focus in this

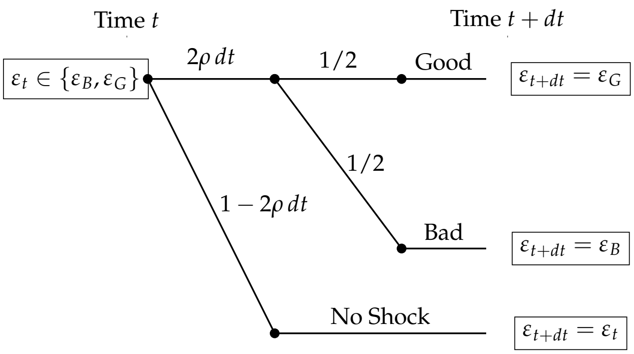

FIGURE 5.5: Productivity shocks over $[t, t + dt)$. The branches of the tree contain the probabilities of the respective random event over $[t, t + dt)$.

section is therefore on the asset pricing implications of token utility and security features rather than on developers' problem.

We introduce productivity shocks by assuming that for $t \geq \tau$ and $\bar{A} \in \{A_L, A_H\}$, platform productivity is given by

$$A_t = \bar{A} + \varepsilon_t, \text{ with } \varepsilon_t \in \{\varepsilon_B, \varepsilon_G\} \text{ and } \bar{A} + \varepsilon_B \geq 0 \text{ and } \varepsilon_G \geq \varepsilon_B.$$

Productivity shocks are as follows. If $\varepsilon = \varepsilon_G$, the platform is subject to a negative productivity shocks $dA = \varepsilon_B - \varepsilon_G$ over dt with probability ρdt. Likewise, if $\varepsilon = \varepsilon_B$, the platform experiences a positive shock $dA = \varepsilon_G - \varepsilon_B$ with probability ρdt. Consequently, the volatility of the productivity shocks, i.e. fundamental volatility, is given by $\varepsilon_G - \varepsilon_B$. We emphasize that productivity shocks, unlike $\bar{A}$, are purely random and not affected by developers' actions.

In general, there are many benefits to having a stable transaction medium (Doepke and Schneider, 2017). For instance, price fluctuations expose transacting users to risks during the transaction settlement period and lead to a drop in users' transaction activities. Excessive price volatility is thus likely to hamper platform adoption. This implies that platform projects should aim for a relatively stable token price and so should try to limit price fluctuations and therefore volatility.

5.6.1 Solution

We characterize the equilibrium token pricing after time τ for a given $A_t = \bar{A}$. Formally, we have to derive the state-dependent adoption levels N_G and N_B and token prices P_G and P_B. With productivity shocks, the platform produces state contingent cash-flows $dD_i = \mu(\bar{A} + \varepsilon_i)N_i dt$ for $i = G, B$, with $\mu'(\cdot) \geq 0$. Using the same steps as above shows that adoption satisfies at time $t \geq \tau$:

$$N_t = N_i = \left(\frac{\bar{A} + \varepsilon_i}{\phi + v \max\{0, r - \mathbb{E}[dP_i + \alpha dD_i]/[P_i dt]\}} \right)^{\frac{1}{1-\xi}}, \text{ for } i = G, B. \tag{5.24}$$

Let us next solve for the token price. Assume first that utility features price the token in both states $i = G, B$ so that $P_i = vN_i$. This is the case when $\mathbb{E}[dP_i + \alpha dD_i] < rP_i dt$ in both states $i = G, B$, i.e., when the expected returns to holding tokens are lower than r. Using Eq. (5.24) and $\mathbb{E}dP_G = \rho(P_B - P_G)dt$, $\mathbb{E}dP_B = \rho(P_G - P_B)dt$ as well as $dD_i = \mu(\bar{A} + \varepsilon_i)N_i dt$, we can solve for

$$P_t = \begin{cases} P_G = v\left(\frac{\bar{A}+\varepsilon_G}{\phi+vr-\alpha\mu(\bar{A}+\varepsilon_G)-v\rho(P_B/P_G-1)}\right)^{\frac{1}{1-\xi}}, \text{if } \varepsilon_t = \varepsilon_G \\ P_B = v\left(\frac{\bar{A}+\varepsilon_B}{\phi+vr-\alpha\mu(\bar{A}+\varepsilon_B)-v\rho(P_G/P_B-1)}\right)^{\frac{1}{1-\xi}}, \text{if } \varepsilon_t = \varepsilon_B. \end{cases} \tag{5.25}$$

Assume next that token security features pin down the token price in both states G, B. In this case, $\mathbb{E}[dP_i + \alpha dD_i] = rP_i dt$ for $i = G, B$ and we can use $N_i = \left(\frac{\bar{A}+\varepsilon_i}{\phi}\right)^{\frac{1}{1-\xi}}$ and $dD_i = \mu(\bar{A}+\varepsilon_i)N_i dt$ for $i = G, B$ to solve for the token price as

$$P_t = \begin{cases} P_G = \frac{1}{r+\rho}\left(\alpha\mu(\bar{A}+\varepsilon_G)\left(\frac{\bar{A}+\varepsilon_G}{\phi}\right)^{\frac{1}{1-\xi}} + \rho P_B\right), \text{if } \varepsilon_t = \varepsilon_G \\ P_B = \frac{1}{r+\rho}\left(\alpha\mu(\bar{A}+\varepsilon_B)\left(\frac{\bar{A}+\varepsilon_B}{\phi}\right)^{\frac{1}{1-\xi}} + \rho P_G\right), \text{if } \varepsilon_t = \varepsilon_B. \end{cases} \tag{5.26}$$

In general, the token prices P_G and P_B are not available in closed form, unless one considers the limit case $\max_A |\mu'(A)| \to 0$ and $\xi \to 0$; see Proposition 22 below.[13] For parsimony, we do not discuss the case in which the token is priced according to its utility features in one state and its security features in another state.

5.6.2 Token price volatility: The role of utility and security features

Our objective is to characterize the effects of utility and security features on token price volatility. In the following, we analyze token price volatility both in absolute terms, that is, $\sigma := P_G - P_B$, and scaled by the average (steady-state) token price $\overline{P} = \frac{P_H+P_L}{2}$, that is $\bar{\sigma} := \frac{\sigma}{\overline{P}}$.[14]

First, consider that token utility features pin down the token price in that the token derives its price from the level of adoption and $N_t = vP_t$. For $\max_A |\mu'(A)| \to 0$ and $\xi \to 0$, we have that

$$\sigma \simeq \frac{v(\varepsilon_G - \varepsilon_B)}{v(r+2\rho) - \alpha\mu_H + \phi},$$
$$\bar{\sigma} \simeq \frac{2\,(\varepsilon_G - \varepsilon_B)\,(vr - \alpha\mu_H + \phi)}{(2A_H + \varepsilon_G + \varepsilon_B)\,(v(r+2\rho) - \alpha\mu_H + \phi)}$$

Naturally, volatility increases with fundamental volatility $\varepsilon_G - \varepsilon_B$. More interestingly, security features α amplify rather than curb the volatility. The reason is that higher security features imply stronger endogenous network effects. These network effects increase the sensitivity of platform adoption to productivity shocks. This boosts the token price volatility because the token derives its value from the level of platform adoption. Due to the endogenous network effects, volatility σ and scaled volatility $\bar{\sigma}$ are increasing and convex in α. In sum, network effects induced by token security features spur adoption at the cost of an increased price volatility. While closed-form expressions for the token price volatility are only available for $\max_A |\mu'(A)| \to 0$ and $\xi \to 0$, Fig. 5.6 numerically shows that the above findings also hold for our baseline environment for which $\mu'(\cdot) > 0$ and $\xi > 0$.

Second, consider that the token is priced according to its security features in both states G and B, which is the case when v is sufficiently small. In this case, using (5.26), one can

[13] While $\xi \to 0$ precludes network effects arising from specification (6.1), our model still features endogenous network effects in that a higher adoption level N_t leads to higher cash flows and dividends, which in turn increases N_t.

[14] Note that by our specification of productivity shocks, the system spends on average equal time in both states, so that the expected long-run price is just the equally-weighted average of P_G and P_B.

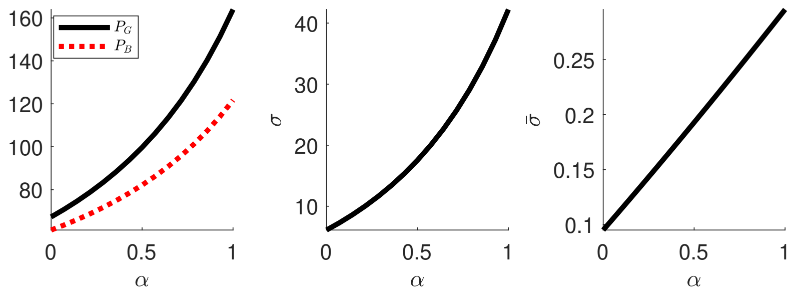

FIGURE 5.6: Token price and volatility as functions of α in our base case environment with $\xi > 0$ and $\mu'(A) > 0$ and the baseline parameters. We pick $\mu(A) = \bar{\mu}A^{\omega}$, with $\bar{\mu} = 0.025$ and $\omega = 10$. $\bar{\sigma}$ is the volatility σ, divided by the steady-state token price $0.5(P_G + P_B)$. The shocks are characterized by $\varepsilon_G = -\varepsilon_B = 0.1$ and $\rho = 0.1$.

calculate that

$$\sigma = \frac{\alpha}{r+2\rho}\left(\mu(\bar{A}+\varepsilon_G)\left(\frac{\bar{A}+\varepsilon_G}{\phi}\right)^{\frac{1}{1-\xi}} - \mu(\bar{A}+\varepsilon_L)\left(\frac{\bar{A}+\varepsilon_L}{\phi}\right)^{\frac{1}{1-\xi}}\right),$$

$$\bar{\sigma} = \frac{2r}{r+2\rho}\left(\frac{\mu(\bar{A}+\varepsilon_G)\left(\frac{\bar{A}+\varepsilon_G}{\phi}\right)^{\frac{1}{1-\xi}} - \mu(\bar{A}+\varepsilon_L)\left(\frac{\bar{A}+\varepsilon_L}{\phi}\right)^{\frac{1}{1-\xi}}}{\mu(\bar{A}+\varepsilon_G)\left(\frac{\bar{A}+\varepsilon_G}{\phi}\right)^{\frac{1}{1-\xi}} + \mu(\bar{A}+\varepsilon_L)\left(\frac{\bar{A}+\varepsilon_L}{\phi}\right)^{\frac{1}{1-\xi}}}\right).$$

That is, when the token is priced according to its security features, token price volatility σ is linear in α, and therefore less sensitive to the provision of security features, while scaled volatility $\bar{\sigma}$ is independent of α. This holds also true for $\xi > 0$ and $\mu'(A) > 0$. The reason for this lower sensitivity is that dividends do not generate network effects if the token is priced according to its security features. Overall, our results highlight that the combination of token utility and security features leads to especially high token price volatility.

The following proposition summarizes our analytical results.

Proposition 22 (Token price volatility) *The following holds:*

1. *Consider the limiting case,* $\max_A |\mu'(A)| \to 0$ *and* $\xi \to 0$. *If the token is priced according to its utility features in both states G and B, then* σ *and* $\bar{\sigma}$ *are increasing and convex in* α.
2. *If the token is priced according to its security features in both states G and B, then* σ *is linearly increasing in* α *and* $\bar{\sigma}$ *is independent of* α.

5.7 Model extensions

5.7.1 Cash diversion

Appendix 5.14 modifies our baseline model by considering that developers can secretly divert cash and receive per dollar diverted $\lambda \in [0,1]$ dollars (in this extension effort choice is thus replaced by diversion). As in the baseline model, platform cash flows are observable after time τ and there is no moral hazard problem once the milestone is reached. We show in this Appendix that the incentive compatibility constraint ensuring that developers do not divert funds is similar to that of the baseline model. For $\lambda \equiv \frac{\kappa(r+\Lambda)}{I}$, this model variant is in fact isomorphic to the baseline model.

5.7.2 Adverse selection

Appendix 5.15 extends our baseline model to incorporate adverse selection. We consider in this Appendix that there exist two types of firms (platforms): a good platform, as described in the baseline version of the model, and a bad platform whose productivity after the milestone equals $A_t = A_L$ with $\mu_t = \mu_L$ with certainty. Both platforms require an initial investment of I. The platform is good with exogenous probability $\pi \in [0, 1]$. Developers are privately informed about platform quality. Token investors only know the probability π that a platform is good. We demonstrate that because the bad type firm has negative NPV in the baseline model (under Assumption 2), there does not exist a separating equilibrium in our baseline environment where Assumption 2 is satisfied. The reason is that in a separating equilibrium, the bad type firm would not receive financing and thus would realize zero payoff, while mimicking the good type yields positive payoff.

Appendix 5.15 then studies the (unique) pooling equilibrium and shows that introducing adverse selection has no qualitative bearing on the model predictions regarding the effects of financing needs (I), cost of effort (κ), or expected time to platform development ($1/\Lambda$) on the optimal level of retention (β) and token security features (α). The main effect of adverse selection is to quantitatively reduce the level of security features attached to tokens (of a good type platform). Indeed, in a pooling equilibrium, developers of a good type firm have to sell more tokens to cover initial financing needs, due to the decrease in the token price relative to the perfect information case, leading to lower token retention β and to greater moral hazard. To maintain incentive compatibility, developers must in turn possess more equity incentives, which requires granting lower cash-flow rights to token holders.

To make the analysis complete, Appendix 5.15 relaxes Assumption 2 by considering environments in which bad type platforms have positive NPV. In this case, there may exist a separating equilibrium, in which token security features signal platform quality. Indeed, according to Proposition 18, attaching security features to tokens is only optimal if platform productivity and cash flows are sufficiently large. Thus, attaching security features to tokens is optimal for good type firms yet costly for bad type firms, facilitating a separating equilibrium. Such a separating equilibrium exists when financing needs (I) are sufficiently low, network effects (ξ) are large, when transaction frictions (v) are high, or when the platform cash flow rate is high. By contrast, sufficiently high costs of effort κ or a sufficiently long time to project completion $1/\Lambda$ preclude the existence of the separating equilibrium.[15] That is, token security features signal good platform quality but their ability to do so crucially depends on platform characteristics and the severity of moral hazard.

Moreover, in a separating equilibrium, adverse selection may boost the provision of token security features, while increasing initial token retention by developers. The reason is that a good type firm signals platform quality by attaching more security features to tokens, thereby increasing the token price at time zero. Consequently, the developers of a good type firm sell fewer tokens to cover initial financing needs I.

In summary, the model implies that adverse selection has an ambiguous effects on the provision of token security features, depending on whether a separating or pooling equilibrium prevails. In a separating equilibrium, in which different types of platforms are financed with different types of tokens and ICOs and STOs coexist, adverse selection increases the provision of token security features. In a pooling equilibrium in which all platforms are financed with the same tokens, adverse selection decreases the provision of token security features.

5.7.3 Endogenous transaction fees

Appendix 5.16 extends the model by allowing developers to charge an endogenous fee $f > 0$ to users for transacting on the platform. This fee increases users' direct cost of transacting to $f + \phi$ and changes platform cash flows $(\mu(A_t) + f)N_t$ directly via f and indirectly via N_t. We consider two cases depending on developers' ability to commit to a fee structure. In the main case discussed here, developers cannot commit.

[15] When κ or $1/\Lambda$ is large, exerting effort is no longer efficient and the good type prefers to mimic the bad type.

Without commitment, the optimal dynamic fee f maximizes at each point in time the dividends accruing to developers $(1-\alpha+\beta\alpha)(\mu_H+f)N_H$ and therefore maximizes platform cash flows $(\mu_H+f)N_H$. The optimal dynamic fee depends on whether the token utility or security features pin down the token price. Moreover, the optimal fee follows a hump-shaped pattern in α. Thus optimal fees are the lowest when tokens have either minimum or maximum utility features with $\alpha=0$ or $\alpha=1$, respectively. This has a bearing on the optimal level of security features. As we show, the optimal tokens with endogenous transaction fees have either low or maximum security features, depending on some platform characteristics, but intermediate values of α are always suboptimal. In this context, the issuance of a utility token (i.e., $\alpha=0$) or a token with heavy cash flow rights can be viewed as a commitment device not to charge high fees in the future, which is particularly useful in the presence of commitment problems to future fees.

Interestingly, the optimal transaction fee f can be negative. In this case, the startup firm *subsidizes* the user base to accelerate platform adoption. In practice, such subsidies are commonly employed by large technology firms. For instance, Alibaba implemented in 2019 a reward scheme providing subsidies to attract developers to its various platforms (Chod et al., 2019). Similarly, Uber is planning to offer financial services, including loans to drivers at favorable rates. We show in the appendix that subsidies to the user base are more likely if the platform is financed with utility tokens or if network effects ξ are strong. In addition, subsidies are only optimal if the platform generates enough revenues μ_H to finance these subsidies. Finally, subsidies are more likely if the blockchain technology facilitates commitment. The reason is that with commitment, developers set fees with more focus on platform adoption instead of only on instantaneous cash flows.

5.7.4 Dynamic trading

Appendix 5.17 extends the model by considering the role of speculators. Because speculators are financially less constrained or more diversified than users and developers, their presence creates gains from trade, so that developers benefit from selling tokens to speculators. As developers cannot commit their token trading strategy, trading opportunities can potentially undermine developers incentives for platform development. To ensure smooth trading patterns, we introduce convex costs of effort for developers.

As in the baseline model, retained tokens β provide incentives. However, the presence of gains from trade makes developers to gradually sell their tokens throughout the development phase so that β smoothly decreases. As developers cannot commit to keeping tokens, they sell tokens and decrease the token price up to the point that they become marginally indifferent between buying and selling tokens. Consequently, all gains from trade are in equilibrium dissipated by the subsequent rise in agency costs. Appendix 5.17 shows how dynamic trading affects the amount of initial retained tokens β and the rate of security features α. We also demonstrate that the main predictions of the baseline model are robust to this extension.

5.7.5 Flow costs of platform development

Appendix 5.18 presents a model variant in which platform development requires operating (monetary) flow costs instead of an initial lump sum cost I at time zero. To raise funds to cover these flow costs, developers dynamically sell tokens to the market, reducing their token retention level and incentives.[16] As a result, the model variant of Appendix 5.18 features similar forces at work and trade-offs as the model variant of Appendix 5.17.

[16]We assume in this Appendix that there are no exogenous costs of selling tokens and raising funds. Introducing fixed costs of raising funds (similar to those in, e.g., Bolton et al., 2011; Hugonnier, Malamud, and Morellec, 2015a) would lead firms to retain cash and would add an additional state variable to the dynamic optimization problem of platform developers.

5.8 Predictions

Our paper provides several new empirical predictions related to platform financing and token design. In the following, we summarize our main predictions.

Prediction 1: *Using fiat money as the platform transaction medium and equity financing is only optimal for platforms that expect high cash flows or strong network effects. For firms without high cash flows and strong network effects (i.e., for which transaction benefits are more important as a source of platform value), using tokens as transaction medium and token financing is optimal, unless moral hazard is severe or financing needs are large.*

This first prediction follows from Proposition 21 and relates to the optimal form of financing. According to this prediction, only platforms where expected cash flows are large (as a fraction of total platform value) should finance platform development with equity issues. Two additional key determinants of optimal financing are moral hazard, which is positively related to the expected time to platform completion, and the cost of developing the platform.

Prediction 2: *For firms relying on token financing, ICOs are expected to be more prevalent for platforms whose value comes mostly from facilitating transactions among users while STOs are expected to be more prevalent for platforms whose value comes mostly from generating cash flows.*

This second prediction follows from Proposition 20 and shows that when using token financing the relative importance of cash flows versus transaction benefits is a key driver of token design.

Prediction 3: *For firms relying on token financing, token security features and developers' retention levels should decrease with the severity of moral hazard and the level of financing needs.*

The third prediction follows from Proposition 19 and underlines the importance of frictions in token design. Notably, because the incentives generated by each dollar of equity ownership are stronger than the incentives from a dollar of token ownership (and equity incentives are undermined by token security features), the level of security features in tokens should decrease with financing needs and moral hazard.

Prediction 4: *When different types of platforms are financed with different types of tokens and ICOs and STOs coexist, adverse selection increases the provision of token security features and the likelihood of platform success increases with token security features.*

The fourth prediction follows the analysis of the effects of adverse selection on token design. Our model predicts that in environments characterized by more informational asymmetries, e.g., with less developed white papers or without code development on open source platforms, token security features should be more prevalent. It also predicts a positive relation between the ex post value of platforms or the likelihood of platform success and the level of token security features.

Prediction 5: *Token price volatility is increasing in security features.*

This last prediction follows from Proposition 22 and derives from the fact that security features generate endogenous network effects that increase the sensitivity of platform adoption to productivity shocks, thereby increasing volatility.

5.9 Conclusion

We study a model in which a startup firm run by developers launches a digital platform. To finance platform development, developers issue tokens that serve as the transaction medium on the platform and thus possess utility features. Tokens may additionally possess cash flow rights and, thus, security features. In the model, platform development is subject to financing needs and moral hazard. This unified model allows us to identify the costs and benefits of various token designs used in practice to finance startup firms.

We show that dividend rights granted to token holders spur platform adoption but dilute developers' equity stake and therefore undermine incentives. As a result, an increase in financing needs or in agency frictions leads to a decrease in token security features. The model also derives conditions under which different types of financing modes are optimal. Specifically, a security token offering or an initial coin offering always dominates traditional equity financing when tokens serve as the transaction medium on the platform. By contrast, whether a security token offering is preferred to an initial coin offering crucially depends on platform and startup characteristics, notably the ability to generate cash flows in addition to facilitating transactions among users.

We also examine when using fiat money as the platform transaction medium and issuing equity to finance platform development is optimal. We find that financing platform development with equity is only optimal if platform cash flows are expected to be large or if network effects are strong. For firms without very high cash flows (or very strong network effects), the platform is generally optimally financed with tokens, unless moral hazard is severe (due e.g., to a long development phase) or financing needs are large.

Finally, we derive additional results by studying various extensions of the model. For instance, we consider the relation between optimal platform financing and platform transaction fees or adverse selection. Notably, we show that the issuance of a pure utility token can be viewed as a commitment device not to charge high transaction fees in the future. In addition, in environments characterized by informational asymmetries, token security features may signal good platform quality and thus may help to distinguish good token offerings from bad ones.

5.10 Discussion of parametric assumptions

Parameter conditions for the analytical solution. We give explicit parameter conditions for Assumption 2, which hold throughout unless otherwise mentioned.

1. We assume that platform development costs I are not excessive and that developers can raise I dollars by issuing tokens, in that

$$I < v\left(\frac{\Lambda}{r+\Lambda}\right)\left(\frac{A_H}{vr+\phi}\right)^{\frac{1}{1-\xi}}. \tag{5.27}$$

   Condition (5.27) ensures that (5.15) admits a positive solution, $\beta_0 > 0$, and thus facilitates financing.

2. We assume the project has positive net present value (NPV) when full effort is exerted and token security features are chosen optimally, in that there exists $\alpha \in [0,1]$ with

$$\Lambda\left(P_H + \frac{(1-\alpha)\mu_H N_H}{r}\right) > (r+\Lambda)I + \kappa. \tag{5.28}$$

3. We assume that the project has negative NPV when no effort is exerted, in that

$$\Lambda\left(P_L + \frac{(1-\alpha)\mu_L N_L}{r}\right) \leq (r+\Lambda)I \tag{5.29}$$

   for any $\alpha \in [0,1]$.

Note that conditions (5.28) and (5.29) jointly imply that

$$\Lambda\left(P_H + \frac{(1-\alpha)\mu_H N_H}{r} - P_L - \frac{(1-\alpha)\mu_L N_L}{r}\right) > \kappa$$

for some α, meaning that exerting effort is efficient. Also recall that $\mu_L = 0$.

Parameters for the numerical analysis. As in Cong et al. (2020a), we set the discount rate to $r = 0.05$, the velocity parameters to $v = 1$, and the network effects parameter to $\chi = 0.125$. The parameter η is set to $\eta = 0.375$ implying that $\xi = 0.5$. Interpreting one unit of time as one year, we set $\Lambda = 1$ implying that developers retain tokens for about one year (because the average time to milestone equals $1/\Lambda$). This is consistent with the findings of Fahlenbrach and Frattaroli (2019) who report that the weighted-average lock-up period for tokens is about one year. We normalize $A_H = 1$. In fact, the absolute value of A_H is not particularly important; instead, its relationship with A_L matters. The value A_L is set to $A_L = 0.55$.

The function $\mu(A)$ is such that $\mu_L = \mu(A_L) = 0$ and $\mu_H = \mu(A_H) = 0.025$, ensuring that $\mu_H < vr$ as stipulated by Assumption 2. We pick $\phi = 0.075$ in order to normalize N_H in the frictionless case to $N_H = 100$. This is convenient because any value N can be interpreted in percentage terms of the adoption level N_H in the frictionless benchmark. Notably, $\phi = 0.075$ also satisfies Assumption 2.

We choose I in order to match the sample average of token retention levels for ICOs. Specifically, we set $I = 58$, which implies in the frictionless benchmark the retention level $\beta = 39\%$, the average token retention level reported for ICOs by Fahlenbrach and Frattaroli (2019). The effort cost κ is varied and chosen so as to generate the desired tensions. We set $\kappa = 33.33$, which is 33.33% of the token price in the frictionless benchmark. This way we capture the high degree of agency problems and agency costs prevailing in this market (Howell et al., 2020; Fahlenbrach and Frattaroli, 2019).

When we vary κ and I, we make sure that Assumption 2 is satisfied. This implies $\kappa < 34.15$ and $I < 59.5$. A similar constraint applies to ξ and v. When we vary v, we employ a lower level of $\phi = 0.0735$ in order to satisfy Assumption 2 across the whole range of values of

v considered. We emphasize that our results are robust across various choices of parameter values.

5.11 Omitted proofs

5.11.1 Auxiliary results

We state the following two auxiliary lemmata.

Lemma 24 *Define* $\hat{\alpha} = \min\{1, \frac{vr}{\mu_H}\}$. *It holds that*

$$\bar{\alpha} = \arg\max_{\alpha} S(\alpha) = \begin{cases} x \in [\hat{\alpha}, 1] & \text{if } \max\{vr - \mu_H, 0\} \geq \frac{(1-\xi)\phi - \mu_H}{\xi} \\ 0 & \text{if } vr \leq \frac{(1-\xi)\phi-\mu_H}{\xi} \\ \frac{vr}{\mu_H} + \frac{1}{\xi} - \frac{\phi(1-\xi)}{\xi\mu_H} & \text{otherwise.} \end{cases}$$

with $S(\alpha) = P_H + \frac{(1-\alpha)\mu_H N_H}{r}$ *and* $N_H = \left(\frac{A_H}{vr-\alpha\mu_H+\phi}\right)^{\frac{1}{1-\xi}}$. *When* $\bar{\alpha} \geq \hat{\alpha}$, $S'(\alpha) > 0$ *for* $\alpha \in [0, \hat{\alpha})$.

Recall that $\hat{\alpha} = \min\{1, \frac{vr}{\mu_H}\}$ and $N_H = \left(\frac{A_H}{\max\{0, vr-\alpha\mu_H\}+\phi}\right)^{\frac{1}{1-\xi}}$. First, note that for all $\alpha \in (\hat{\alpha}, 1]$, we have by (5.12) that $P_H = \frac{\alpha\mu_H N_H}{r}$ and thus $S(\alpha) = \frac{\mu_H N_H}{r}$ with $N_H = \left(\frac{A_H}{\phi}\right)^{\frac{1}{1-\xi}}$. Hence, $S(\alpha)$ does not depend on α for $\alpha > \hat{\alpha}$ (i.e., $S'(\alpha) = 0$ for $\alpha > \hat{\alpha}$).

Second, define $\varepsilon := 1/(1-\xi) \geq 1$ and calculate for $\alpha < \hat{\alpha}$ (in which case $P_H = vN_H$)

$$\begin{aligned} S'(\alpha) &= \varepsilon\left(v + \frac{(1-\alpha)\mu_H}{r}\right) N_H \frac{\mu_H}{vr - \mu_H\alpha + \phi} - \frac{\mu_H}{r} N_H \\ &\propto \varepsilon\left(vr + (1-\alpha)\mu_H\right) \frac{\mu_H}{vr - \mu_H\alpha + \phi} - \mu_H \propto \varepsilon vr + \varepsilon(1-\alpha)\mu_H - vr + \mu_H\alpha - \phi \\ &= v(\varepsilon-1)r + \varepsilon\mu_H - \alpha(\varepsilon-1)\mu_H - \phi \propto vr + \frac{\mu_H}{\xi} - \alpha\mu_H - \frac{(1-\xi)\phi}{\xi}. \end{aligned}$$

It follows that $S'(\alpha)$ has at most one root on $(0, \hat{\alpha})$. If $\alpha = \bar{\alpha} \in (0, \hat{\alpha})$, then the optimal $\alpha = \bar{\alpha}$ solves the FOC $S'(\bar{\alpha}) = 0$, so $\bar{\alpha} = \frac{vr}{\mu_H} + \frac{1}{\xi} - \frac{\phi(1-\xi)}{\xi\mu_H}$.

Next, if $S'(0) \leq 0$, then $S'(\alpha) < 0$ for all $\alpha < 0$, hence $\alpha = \bar{\alpha} = 0$ is optimal if $S'(0) \leq 0$. Observe that

$$S'(0) \leq 0 \iff \frac{vr}{\mu_H} + \frac{1}{\xi} - \frac{(1-\xi)\phi}{\xi\mu_H} \leq 0.$$

The above inequality condition is equivalent to $vr \leq \frac{(1-\xi)\phi-\mu_H}{\xi}$.

Last, if $S'(\hat{\alpha}) \geq 0$, then $S'(\alpha) > 0$ for all $\alpha < \hat{\alpha}$, so $\alpha = \bar{\alpha} \in [\hat{\alpha}, 1]$ is optimal. Observe that

$$S'(\hat{\alpha}) \geq 0 \iff vr - \hat{\alpha}\mu_H + \frac{\mu_H}{\xi} - \frac{(1-\xi)\phi}{\xi} \geq 0. \tag{5.30}$$

The above inequality can be compactly rewritten as $\max\{vr - \mu_H, 0\} \geq \frac{(1-\xi)\phi-\mu_H}{\xi}$.

Lemma 25 *It holds that*

$$\arg\max_{\alpha} S(\alpha) = \arg\max_{\alpha} \left(\beta P_H + \frac{(1-\alpha)\mu_H N_H}{r}\right) - \frac{\kappa}{\Lambda},$$

with $S(\alpha) = P_H + \frac{(1-\alpha)\mu_H N_H}{r}$, $N_H = \left(\frac{A_H}{\max\{0, vr-\alpha\mu_H\}+\phi}\right)^{\frac{1}{1-\xi}}$, *and* $\beta = \beta_0$ *satisfying* (5.15).

Because $\beta = \beta_0$ satisfies (5.15), it holds that $\beta = 1 - \frac{I}{P_H}\frac{r+\Lambda}{\Lambda}$, implying

$$\beta P_H + \frac{(1-\alpha)\mu_H N_H}{r} - \frac{\kappa}{\Lambda} = S(\alpha) - \frac{\kappa}{\Lambda} - \frac{r+\Lambda}{\Lambda} I.$$

Because $\frac{\kappa}{\Lambda}$ and $\frac{(r+\Lambda)I}{\Lambda}$ do not depend on α, the claim follows.

5.11.2 Proof of Proposition 17

The assertion follows directly from the developments in the main text.

5.11.3 Proof of Proposition 18

Lemma 24 derives the expression of $\bar{\alpha} = \arg\max_{\alpha \in [0,1]} S(\alpha)$. When $\kappa = 0$, there is no moral hazard problem and the incentive condition (5.13) (i.e., $IC(\alpha) \geq 0$) is not relevant for the developers' optimization problem (5.8). As a result, the developers solve

$$\max_{\alpha} \left(\beta P_H + \frac{(1-\alpha)\mu_H N_H}{r} \right) - \frac{\kappa}{\Lambda},$$

where $\beta = \beta_0$ satisfies (5.15). By Lemma 25, the developers choose the level of α to maximize $S(\alpha)$, so that $\alpha = \bar{\alpha}$.

5.11.4 Proof of Proposition 19

Claim 1

When $\kappa/\Lambda \to 0$, the incentive condition $IC(\alpha) \geq 0$ is always satisfied for any α and thus not relevant for the developers' optimization problem. As a result, the developers solve $\max_{\alpha} \left(\beta P_H + \frac{(1-\alpha)\mu_H N_H}{r} \right) - \frac{\kappa}{\Lambda}$, where $\beta = \beta_0$ satisfies (5.15). By Lemma 25, the developers choose the level of α to maximize $S(\alpha)$, so that $\alpha = \bar{\alpha}$. By continuity, it holds that $\alpha = \bar{\alpha}$ for κ/Λ sufficiently small, i.e., for κ or $1/\Lambda$ sufficiently small.

Consider the limit case $I \to 0$, denoted by $I = 0$. When $I = 0$, the incentive condition $IC(\alpha) \geq 0$ becomes

$$\Lambda \left(P_H + \frac{(1-\alpha)\mu_H N_H}{r} - P_L \right) = \Lambda \left(S(\alpha) - P_L \right) \geq \kappa.$$

As by Assumption 3, $\alpha = \bar{\alpha} = 1$ maximizes $S(\alpha)$, it follows by means of parameter condition (5.28) (i.e., Assumption 2) that $\Lambda S(1) > (r+\Lambda)I + \kappa$. At the same time, parameter condition (5.29) implies that $P_L \leq (r+\Lambda)I$ for any $\alpha \in [0,1]$. As a result, $IC(\alpha) > 0$ for $\alpha = 1$. Because $\alpha = \bar{\alpha} = 1$ is the developers' optimal choice absent frictions (see Proposition 18 and Lemma 24), it follows that the incentive condition $IC(\alpha)$ is not relevant for the developers' optimization problem and loose in optimum, when $I = 0$. That is, when $I = 0$, the developers maximize $S(\alpha)$ over α and choose $\alpha = \bar{\alpha} = 1$. By continuity, it follows that $\alpha = \bar{\alpha}$, provided $I > 0$ is sufficiently small.

Claims 2 and 3

Take the financing constraint (5.3) (which binds in optimum) or equivalently (5.15), that is, $\beta = 1 - \frac{I}{P_H}\frac{r+\Lambda}{\Lambda}$. It follows that β strictly decreases in I when P_H decreases in I. Note that by the token pricing equation (5.12), P_H does not depend on I directly but only through the optimal choice of α. Therefore, if β increases in I, then α must strictly increase in I, because P_H strictly increases in α.

Take the surplus (i.e., overall platform value) $S(\alpha)$ and observe that $S(\alpha)$ does not depend on I directly but only through the optimal choice of α. Also note that an increase in financing frictions/needs, as captured by I, in optimum cannot cause more efficient provision of token security features α, i.e., cannot increase surplus $S(\alpha)$. Because $\alpha = \bar{\alpha} = 1$ is efficient and optimal absent frictions and when $I = 0$ and because $S'(\alpha) > 0$ for all $\alpha \in [0, 1)$, this means that α cannot strictly increase in I. That is, α decreases in I and so does P_H, implying that β strictly decreases in I.

Also note that only if the incentive condition $IC(\alpha) \geq 0$ is tight, then α strictly decreases in I, as I can affect the choice of optimal α only via the incentive condition (5.13). Analogously, it follows that β strictly increases in Λ and that α (strictly) increases in Λ (only if the incentive condition $IC(\alpha) \geq 0$ is tight).

Finally, note that κ affects P_H and thus β only through the optimal choice of α. If P_H increases in κ, then α must increase in κ, as P_H increases in α. However, it is clear that an increase in agency frictions, as captured by κ, in optimum cannot trigger more efficient efficient provision of token security features α, i.e., cannot increase $S(\alpha)$. Because $\alpha = \bar{\alpha} = 1$ is efficient and optimal absent agency frictions (i.e., when $\kappa = 0$) and because $S'(\alpha) > 0$ for $\alpha \in [0, 1)$, optimal α must decrease in κ. Thus, $d\beta/d\kappa \leq 0$ and $d\alpha/d\kappa \leq 0$, where the inequalities are strict only if the incentive condition $IC(\alpha) \geq 0$ is tight.

Claim 4

Note that $N_H = \left(\frac{A_H}{vr+\phi-\alpha\mu_H}\right)^{\frac{1}{1-\xi}} \geq \left(\frac{A_H}{vr+\phi}\right)^{\frac{1}{1-\xi}} > 1$, where the second inequality uses the parameter assumption $A_H > vr + \phi$. In addition, due to $N_H > 1$, it follows that N_H strictly increases in ξ and so does $P_H = vN_H$, thus $\frac{\partial\beta}{\partial\xi} > 0$ by means of (5.15) with $\beta = \beta_0$. Next, note that:

$$\frac{\partial N_H}{\partial \xi} = N_H \ln(N_H) \frac{1}{(1-\xi)^2} > N_L \ln(N_L) \frac{1}{(1-\xi)^2} = \frac{\partial N_L}{\partial \xi}. \tag{5.31}$$

Taking $IC(\alpha) = \Lambda\left(\beta v + \frac{(1-\alpha)\mu_H}{r}\right) N_H - \kappa - \Lambda\beta v N_L$, it follows that

$$\begin{aligned}
\frac{\partial IC(\alpha)}{\partial \xi} &= \Lambda v \frac{\partial \beta}{\partial \xi}(N_H - N_L) + \Lambda\beta v\left(\frac{\partial N_H}{\partial \xi} - \frac{\partial N_L}{\partial \xi}\right) + \frac{\Lambda(1-\alpha)\mu_H}{r}\frac{\partial N_H}{\partial \xi} \\
&> \Lambda v \frac{\partial \beta}{\partial \xi}(N_H - N_L) + \Lambda\beta v\left(\frac{\partial N_H}{\partial \xi} - \frac{\partial N_L}{\partial \xi}\right) > 0,
\end{aligned}$$

where the first inequality uses $\mu_H > 0$ and the second inequality uses (5.31). Because $\alpha = \bar{\alpha} = 1$ is efficient absent frictions, because $S'(\alpha) > 0$ for all $\alpha \in [0, 1)$, and because an increase in ξ relaxes incentive compatibility (i.e., $\frac{\partial IC(\alpha)}{\partial\xi} > 0$), it follows that α increases in ξ and β strictly increases in ξ. Note that α strictly increases in ξ only if the incentive condition $IC(\alpha) \geq 0$ is tight in optimum.

Claim 5

We show that developers do not find it optimal to issue equity next to tokens and start with some additional notation that allows for equity issuance (next to tokens). The value of equity derives from the dividends received by shareholders. As a result, it is given by the discounted stream of expected future dividends

$$E_{j,\tau} = \int_t^{\infty} e^{-r(s-t)} \mu(A_s) N_s (1-\alpha) ds = \frac{(1-\alpha)N_j\mu_j}{r}$$

after time τ (i.e., for $t \geq \tau$) for $j = H, L$ and by

$$E_j = \mathbb{E}_t\left[\int_t^{\infty} e^{-r(s-t)} \mu(A_s) N_s (1-\alpha) ds\right] = \frac{\Lambda E_{j,\tau}}{r + \Lambda}$$

before time τ (i.e., for $t < \tau$) for $j = H, L$. We denote by γ the developers' equity retention level after time zero. In our model, there is no reason to issue equity after time zero. Further, it suffices to focus on instances in which $\alpha < 1$, as otherwise the equity value trivially equals zero.

When the startup firm can issue equity at time 0, the financing constraint (which binds in optimum) becomes

$$\frac{\Lambda}{r+\Lambda}\left((1-\beta)P_H + \frac{(1-\gamma)(1-\alpha)N_H\mu_H}{r}\right) = I, \tag{5.32}$$

as the startup firm can cover the cost I of developing the platform by raising equity and/or by selling tokens. Selling equity, like granting dividend rights to token holders, implies a dilution of the developers' stake in the firm. With equity financing, the incentive constraint becomes

$$IC_E(\alpha) := \Lambda\left(\beta P_H + \frac{\gamma(1-\alpha)\mu_H N_H}{r}\right) - \kappa - \Lambda\beta P_L \geq 0. \tag{5.33}$$

We can then derive developers' (continuation) payoff before time τ as

$$V_E(\alpha) = \frac{\Lambda(v\beta + \gamma(1-\alpha)\mu_H/r)N_H - \kappa}{r+\Lambda} = \frac{\Lambda S(\alpha) - \kappa}{r+\Lambda} - I, \tag{5.34}$$

where the subscript "E" denotes quantities under external equity financing. Next, using (5.32), we formulate the developers' optimization problem as

$$\max_{\alpha,\beta,\gamma\in[0,1]} V_E(\alpha) \text{ s.t. } (5.32).$$

That is, by (5.34), developers maximize $S(\alpha)$ subject to the incentive constraint (5.33) and the financing constraint (5.32) (which binds in optimum) over $\alpha, \beta, \gamma \in [0,1]$. Because equity and tokens are fairly priced, developers can extract all the surplus from the platform so that $V_E(\alpha)$ does not directly depend on (β, γ). Thus, for any α, it is optimal for developers to choose (β, γ) to maximize $IC_E(\alpha)$. Due to (5.34), for given (β, γ), it is optimal for developers to choose α to maximize $S(\alpha)$ subject to $IC_E(\alpha) \geq 0$. Because of $\bar{\alpha} = 1$ and $S'(\alpha) > 0$ for all $\alpha < 1$, the developers choose in optimum the maximum level of α that satisfies $IC_E(\alpha) \geq 0$.

Let α_{NE} denote the optimal level of security features without equity financing (i.e., with $\gamma = 1$ and $IC(\alpha_{NE}) \geq 0$). First consider that $\alpha_{NE} = 1$. Then the equity value is zero. Thus, raising funds by means of equity requires to set $\alpha < 1$. However, setting $\alpha < 1$ is sub-optimal because $\alpha = \bar{\alpha} = 1$ maximizes $S(\alpha)$ (see Assumption 2 and Lemma 24) and therefore $V_E(\alpha)$. Therefore, let us consider in the following that $\alpha_{NE} < 1$. By Assumption 2 and Lemma 24, absent frictions, the optimal choice of α is equal to $\bar{\alpha} = 1$. Thus, $\alpha_{NE} < 1$ implies that $IC(\alpha_{NE}) = 0$.

Next, take any α and the financing constraint $(1-\beta)P_H + \frac{(1-\gamma)(1-\alpha)N_H\mu_H}{r} = \frac{(r+\Lambda)I}{\Lambda}$ and implicitly differentiate w.r.t. γ to obtain

$$0 = -P_H\frac{d\beta}{d\gamma} - \frac{(1-\alpha)N_H\mu_H}{r} \Longrightarrow \frac{d\beta}{d\gamma} = -\frac{(1-\alpha)N_H\mu_H}{P_H r} = -\frac{(1-\alpha)\mu_H}{vr},$$

where we used the pricing relationship $P_H = N_H v$, implied by $vr > \mu_H$ (see Assumption 3). We look at the incentive condition:

$$IC_E(\alpha) = IC_E(\alpha|\gamma) := \Lambda\left(\beta v + \frac{\gamma(1-\alpha)\mu_H}{r}\right)N_H - \kappa - \Lambda\beta v N_L \geq 0$$

and calculate:

$$\frac{dIC_E(\alpha)}{d\gamma} \propto N_H\left(\frac{d\beta}{d\gamma} + \frac{(1-\alpha)\mu_H}{vr}\right) - N_L\frac{d\beta}{d\gamma} = \frac{(1-\alpha)N_L}{vr}\mu_H.$$

Thus, $dIC_E(\alpha_{NE})/d\gamma > 0$.

Also note that $V_E(\alpha)$ does not directly on (β, γ). Choosing $\gamma = 1$ and β to satisfy the financing constraint (5.3) yields payoff $V_E(\alpha_{NE})$, with $IC_E(\alpha_{NE}|\gamma = 1) = 0$ and $IC_E(\alpha|\gamma =$

$1) < 0$ for $\alpha \in (\alpha_{NE}, 1]$. If there existed $\alpha > \alpha_{NE}$ with $IC_E(\alpha|\gamma = 1) \geq 0$, then it would be optimal to increase token security features up to α as $S'(\alpha) > 0$ for all $\alpha \in [0,1)$ (owing to Assumption 3 and Lemma 24), contradicting the optimality of α_{NE}. Here, $IC_E(\alpha|\gamma = x)$ explicitly denotes the function $IC_E(\cdot)$, when tokens possess cash-flow rights α and developers retain fraction γ of equity.

Consider now that the developers choose $\gamma = \hat{\gamma} < 1$. Owing to $dIC_E(\alpha)/d\gamma > 0$ for $\alpha < 1$, it follows that $IC_E(\alpha|\gamma = \hat{\gamma}) < 0$ for $\alpha \geq \alpha_{NE}$. Thus, developers realize payoff bounded by $V_E(\alpha_E)$ with some $\alpha_E < \alpha_{NE}$. Because of Assumption 3 and Lemma 24 (i.e., $S'(\alpha) > 0$ for $\alpha \in [0,1)$) and (5.34), it holds that $V_E(\alpha_E) < V_E(\alpha_{NE})$, so that setting $\gamma < 1$ is sub-optimal and raising funds by issuing equity next to tokens is sub-optimal.

5.11.5 Proof of Proposition 20

By Proposition 18, $\alpha = 0$ maximizes the overall surplus if and only if $vr\xi \leq (1-\xi)\phi - \mu_H$. By Lemma 18, $\alpha = 0$ then also maximizes $\beta P_H + \frac{(1-\alpha)\mu_H N_H}{r}$, with $\beta = \beta_0$ satisfying (5.15).

Recall that the developers' incentive constraint reads

$$IC(\alpha) = \Lambda\left(\beta P_H + \frac{(1-\alpha)\mu_H N_H}{r} - \beta P_L\right) - \kappa \geq 0.$$

Because in optimum $\beta = \beta_0$ solves (5.15) and P_H strictly increases in α (see the token pricing equation (5.12)), it follows that β increases in α, while P_L does not depend on $\alpha = 0$, due to $\mu_L = 0$ (see the token pricing equation (5.12)). Therefore, βP_L increases in α and hence is minimized for $\alpha = 0$. As $\alpha = 0$ maximizes $\beta P_H + \frac{(1-\alpha)\mu_H N_H}{r}$, it follows that $\alpha = 0$ maximizes $IC(\alpha)$. As a result, when $vr\xi \leq (1-\xi)\phi - \mu_H$, $\alpha = 0$ maximizes both overall platform value (surplus) $S(\alpha)$ and incentives $IC(\alpha)$ and thus is optimal.

Next, a STO with $\alpha > 0$ maximizes the surplus $S(\alpha)$ if (see Proposition 18)

$$\bar{\alpha} > 0 \iff vr\xi > (1-\xi)\phi - \mu_H.$$

By Lemma 18, $\alpha > 0$ then also maximizes $\beta P_H + \frac{(1-\alpha)\mu_H N_H}{r}$, with $\beta = \beta_0$ satisfying (5.15). Then, $\alpha > 0$ is optimal when (in optimum) $IC(\alpha) > 0$ and the incentive constraint (5.13) does not affect the developers' optimization problem and so does not constrain the choice of α relative to the frictionless benchmark. This is the case if either κ, $1/\Lambda$ or I is sufficiently small (for details: see proof of Claim 1 in Proposition 19).

5.11.6 Proof of Proposition 21

Recall that by Assumption 3 and Lemma 24, the value of a token-based platform is maximized for $\bar{\alpha} = 1$. In addition, we have that $vr > \mu_H$. The adoption level of a fiat-based platform equals $N^F := \left(\frac{A_H}{\phi}\right)^{\frac{1}{1-\xi}}$ and is — due to $vr > \mu_H$ — always larger than the adoption level of a token-based platform, $N^T := \left(\frac{A_H}{\phi + vr - \alpha\mu_H}\right)^{\frac{1}{1-\xi}}$, for any $\alpha \in [0,1]$. That is, $N^F > N^T$.

Consider the problem $\max_{v^* \in \{0,v\}} S(\bar{\alpha})$, where v is an exogenous parameter. This maximization problem can be solved by comparing the surplus (i.e., overall platform value) under a fiat-based platform, given by $\mathcal{A} := \frac{\mu_H}{r}\left(\frac{A_H}{\phi}\right)^{\frac{1}{1-\xi}}$, with the surplus under a token-based platform when $\alpha = \bar{\alpha} = 1$, given by $\mathcal{B} := v\left(\frac{A_H}{\phi + vr - \mu_H}\right)^{\frac{1}{1-\xi}}$. Then, a fiat-based platform yields higher surplus than a token-based platform if and only if $\mathcal{A} \geq \mathcal{B}$, that is, if and only if $\frac{\mu_H}{vr} \geq \left(\frac{\phi}{\phi + vr - \mu_H}\right)^{\frac{1}{1-\xi}}$, which yields (5.23).

Recall that the developers' incentive constraint reads

$$IC(\alpha) = \Lambda\left(\beta P_H + \frac{(1-\alpha)\mu_H N_H}{r} - \beta P_L\right) - \kappa \geq 0.$$

The arguments presented in the proof of Lemma 18 illustrate that

maximizing $\beta P_H + \frac{(1-\alpha)\mu_H N_H}{r}$, with $\beta = \beta_0$ satisfying (5.15), is equivalent to maximizing the overall platform value (surplus) $S(\alpha)$. Thus, when (5.23) holds, then $v^* = 0$ also maximizes $\beta P_H + \frac{(1-\alpha)\mu_H N_H}{r}$, with $\beta = \beta_0$ satisfying (5.15). Owing to $\mu_L = 0$ and the pricing equation (5.12), it follows under a fiat-based platform with $v^* = 0$ that $P_L = 0$, whereas $P_L \geq 0$ under a token-based platform with $v^* = v$. That is, setting $v^* = 0$ minimizes the term βP_L, while maximizing $\beta P_H + \frac{(1-\alpha)\mu_H N_H}{r}$. As a result, when (5.23) holds, then $v^* = 0$ maximizes both the overall platform value (surplus) $S(\alpha)$ and the developers' incentives $IC(\alpha)$. Hence a fiat-based platform is optimal when (5.23) holds.

In contrast, when (5.23) does not hold, a token-based platform leads to higher surplus $S(\alpha)$, in that v^* maximizes overall platform value. Hence, when (5.23) does not hold, then $v^* = v$ also maximizes $\beta P_H + \frac{(1-\alpha)\mu_H N_H}{r}$, with $\beta = \beta_0$ satisfying (5.15). It is therefore optimal to implement a token-based platform with $v^* = v$ if the incentive condition (5.13) does not affect the developers' optimization problem (i.e., does not constrain the optimal choice of v^*), which is the case if either I, κ or $1/\Lambda$ is sufficiently low (for details: see proof of Claim 1 in Proposition 19).

We demonstrate under what conditions (5.23) is satisfied. Note that—owing to $\xi > 0$—(5.23) holds in the limit $v \to \infty$ and thus, by continuity, for v sufficiently large. Likewise, because $vr > \mu_H$, it follows that the RHS of (5.23) tends to zero as $\xi \to 1$, so that (5.23) holds for ξ sufficiently large.

Last, we analyze μ_H and define $f(\mu_H) : \mu_H \mapsto \frac{\mu_H}{vr} - \left(\frac{\phi}{vr-\mu_H+\phi}\right)^{\frac{1}{1-\xi}}$ and note that (5.23) holds whenever $f(\mu_H) \geq 0$. Observe that for $\mu_H = vr$, (5.23) holds in equality and $f(\mu_H) = 0$. Next, calculate $f'(vr) = \frac{1}{vr} - \frac{1}{(1-\xi)\phi}$, which is negative if and only if $vr > (1-\xi)\phi$. Thus, if $vr > (1-\xi)\phi$, it follows that $f(\mu_H) > 0$ and (5.23) holds in a left neighbourhood of vr, i.e., for $\mu_H < vr$ sufficiently large. Note that $\mu_H \geq \phi(1-\xi)$ implies $vr > \phi(1-\xi)$, due to $vr > \mu_H$.

5.12 Token price volatility

5.12.1 Proof of Corollary 22

Take $\bar{A} \in \{A_L, A_H\}$. First, assume that utility features pin down the token price in both states, G, B. This results into the equilibrium pricing system (5.25):

$$P_G = P_G(\alpha) = v\left(\frac{\bar{A}+\varepsilon_G}{vr - \alpha\mu(\bar{A}+\varepsilon_G) + \phi - v\rho(P_B/P_G - 1)}\right)^{\frac{1}{1-\xi}}$$
$$P_B = P_B(\alpha) = v\left(\frac{\bar{A}+\varepsilon_B}{vr - \alpha\mu(\bar{A}+\varepsilon_B) + \phi - v\rho(P_G/P_B - 1)}\right)^{\frac{1}{1-\xi}}.$$

Linearize the above system (w.r.t. ξ and $\mu(\cdot)$) to obtain

$$P_G = P_G(\alpha) = v\left(\frac{\bar{A}+\varepsilon_G}{vr - \alpha\mu(\bar{A}) + \phi - v\rho(P_B/P_G - 1)}\right) + o(\xi) + o(\max_A |\mu'(A)|),$$
$$P_B = P_B(\alpha) = v\left(\frac{\bar{A}+\varepsilon_B}{vr - \alpha\mu(\bar{A}) + \phi - v\rho(P_G/P_B - 1)}\right) + o(\xi) + o(\max_A |\mu'(A)|).$$

Discarding the terms $o(\xi) + o(\max_A |\mu'(A)|)$, this becomes a system of two linear equations, which can be solved for

$$P_G \simeq \frac{v\ (\bar{A}(v(2\rho + r) - \alpha\mu(\bar{A}) + \phi)\bar{A} + \varepsilon_G(v(r+\rho) - \alpha\mu(\bar{A})) + \varepsilon_B v\rho)}{(\phi - \alpha\,\mu(\bar{A}) + vr)\ (\phi - \alpha\,\mu(\bar{A}) + v(r + 2\rho))}$$

and

$$P_B \simeq \frac{v\ (\bar{A}(v(2\rho + r) - \alpha\mu(\bar{A}) + \phi)\bar{A} + \varepsilon_B(v(r+\rho) - \alpha\mu(\bar{A})) + \varepsilon_G v\rho)}{(\phi - \alpha\,\mu(\bar{A}) + vr)\ (\phi - \alpha\,\mu(\bar{A}) + v(r + 2\rho))},$$

where we omit the remainder terms of order $o(\xi) + o(\max_A |\mu'(A)|)$. One can then calculate that

$$\sigma \simeq P_G - P_B = \frac{v(\varepsilon_G - \varepsilon_B)}{\phi + v(r + 2\rho) - \alpha\mu(\bar{A})}$$
$$\bar{\sigma} \simeq \frac{P_G - P_B}{0.5(P_G + P_B)} = \frac{2\ (\varepsilon_G - \varepsilon_B)\ (vr - \alpha\mu(\bar{A}) + \phi)}{(2\bar{A} + \varepsilon_G + \varepsilon_B)\,(v(r + 2\rho) - \alpha\mu(\bar{A}) + \phi)},$$

showing that σ and $\bar{\sigma}$ are increasing and convex in α, for ξ and $\max_A |\mu'(A)|$ sufficiently small.

The second claim directly follows from the main text and in particular after some calculations, departing from expression (5.26).

5.12.2 The problem before time τ with uncertainty after time τ

Set $\bar{A} = A_H$. At time τ, we have by assumption that $\mathbb{P}(\varepsilon_\tau = \varepsilon_G) = \mathbb{P}(\varepsilon_\tau = \varepsilon_B) = 1/2$. The solution to the model with persistent productivity shocks is then similar to that of Section 5.3 with the expected token price $\bar{P}$ and expected adoption levels $\bar{N}$, given by

$$\bar{P} := \frac{P_G + P_B}{2} \quad \text{and} \quad \bar{N} := \frac{N_G + N_B}{2},$$

replacing the token price P_H and adoption level N_H.

5.13 Micro-foundation for transaction protocol and holding period

5.13.1 Microfoundation for convenience yield

The flow utility from transacting, as stipulated in (6.1), can be micro-founded by a random search and matching protocol. For a detailed micro-foundation, we refer to Cong et al. (2020b,a).

5.13.2 Microfoundation for token holding period

In the following, we give several potential ways to micro-found the holding period vdt. For the exposition, we consider our framework after the milestone, i.e., $t > \tau$, and assume that all users hold the tokens merely for transaction purposes. That is, the token is priced according to its utility features (which is the case if $vr > \alpha\mu_H$).

Transaction settlement delays. There are transaction settlement delays of length vdt. Consider a transaction that is initiated at time t. After the transaction is initiated, its execution is delayed by vdt units of time and so its execution is completed at time $t + vdt$. Notably, over $[t, t + vdt]$, tokens used in the transaction, initiated at time t, are locked and cannot be used otherwise, i.e., cannot be sold. After the transaction is executed at time $t + vdt$, tokens used in the transaction can be sold again. It follows that any transaction requires to hold tokens over a time period of length vdt.

In our model, the value of any transaction is one dollar, i.e., $1/P_t$ tokens. That is, N_t is the number of dollar transactions which are initiated over a short period of time $[t, t+dt)$. Thus, over any time period $[t, t+dt)$, N_t initiated transactions—each worth one dollar—require to hold $1/P_t$ tokens for vdt units of time.

Transactions are equally spread over time (i.e., over the time interval $[t, t+dt)$). This implies that at any time t, tokens are only held for transactions that are initiated over the interval $[t - vdt, t)$ (and so executed over $[t, t+vdt)$). The number of transactions initiated over $[t - vdt, t)$ equals

$$\frac{1}{dt}\left(\int_{t-vdt}^{t} N_s ds\right) = \frac{1}{dt}\left(\int_{t-vdt}^{t} (N_t + (N_s - N_t))\, ds\right) = \frac{1}{dt}\left(N_t vdt - o((dt)^2)\right) = vN_t,$$

where the second equality uses that $N_t - N_s \simeq dN_t$ is infinitesimal in that $ds(N_t - N_s) = o((dt)^2)$. The last inequality ignores higher order terms in that $o((dt)^2) = 0$.

As a result, at any time t, vN_t transactions, that have been initiated yet not executed, require to hold one dollar in tokens, i.e., $1/P_t$ dollars. Hence, the aggregate token demand equals vN_t/P_t, which by virtue of market clearing must equal the token market supply $1 - \beta_t$. Therefore, the token price equals:

$$P_t = \frac{vN_t}{1-\beta_t}.$$

Deposits. Alternatively, we could obtain the same results by assuming that users have to hold fraction v of the overall transaction value in tokens over $[t, t+dt)$. Here, $v > 1$ implies that users have to put a deposit while $v < 1$ allows users to transact with margins. Specifically, a transaction of value x requires to holder vx tokens over $[t, t+dt)$. This implies the token demand vN_t/P_t and so—by virtue of market clearing—the token price:

$$P_t = \frac{vN_t}{1-\beta_t}.$$

5.14 Cash diversion

Consider the following formulation of the moral hazard problem. Over $[t, t+dt)$, with probability Λdt, the milestone τ arrives. If developers invest upon reaching the milestone amount $I > 0$, productivity becomes $A_t = A_H$ for $t \geq \tau$. If they invest less than I or do not invest at all, productivity becomes $A_t = A_L$ for $t \geq \tau$. Thus, in order to develop the project and to reach high platform productivity, developers must raise amount I at time zero and save/store this amount to be able to invest at the milestone. We assume stored/saved dollars to not earn interest. As in the baseline model, it is optimal not to hold more cash than necessary, i.e., no more than I dollars.

Following DeMarzo and Sannikov (2006), before time τ, developers can secretly divert cash and receive per dollar diverted $\lambda \in [0,1]$ dollars. After time τ, platform cash flows are observable, and there is no moral hazard problem anymore.

Given α and β, we analyze the developers' incentives to divert cash at any time t before time τ. It is clear that if cash diversion is optimal, then it is optimal to divert all cash I, yielding λI dollars. However, after diversion, the project has low productivity after the milestone, and the price becomes $P_t = P_L = vN_L$. That is, at time τ, developers earn βP_L dollars from selling all retained tokens and $(1-\alpha)\mu_L N_L/r$ dollars from future cash flows. Otherwise, if developers do not divert cash, token price equals $P_H = vN_H$, and they obtain βP_H dollars from selling all retained tokens and $(1-\alpha)\mu_H N_H/r$ dollars from future cash flows.

As at any time $t < \tau$ the expected time to reaching the milestone equals $1/\Lambda$, it follows that cash diversion is not optimal if and only if

$$\lambda I + \frac{\Lambda(\beta P_L + (1-\alpha)\mu_L N_L/r)}{r+\Lambda} \leq \frac{\Lambda(\beta P_H + (1-\alpha)\mu_H N_H/r)}{r+\Lambda}.$$

Rewriting yields

$$\Lambda\left(\beta P_H + \frac{(1-\alpha)\mu_H N_H}{r}\right) - \lambda(r+\Lambda)I - \Lambda\left(\beta P_L + \frac{(1-\alpha)\mu_L P_L}{r}\right),$$

which is similar to the incentive constraint in the baseline model, i.e., (5.13). For $\lambda \equiv \frac{\kappa(r+\Lambda)}{I}$, this model variant in fact is isomorphic to the baseline model.

5.15 Adverse selection

This Appendix introduces adverse selection in our model by considering that there are two possible types of firms (i.e., platforms) operated by developers: a good platform, as described in the baseline version of the model, and a bad platform whose productivity after the milestone equals $A_t = A_L$ with $\mu_t = \mu_L$ with certainty. Both platforms require an initial investment I. The platform is good with exogenous probability $\pi \in [0,1]$. Developers are privately informed about platform quality. Token investors only know the probability π that a platform is good. Assumption 2 implies that good platforms are profitable and have positive net present value (NPV), but bad platforms with low productivity A_L are inefficient to finance and have negative net present value.

Let us start the analysis by looking for a separating equilibrium. In a separating equilibrium, the good firm grants cash flow rights α to token holders and retains tokens β, resulting in (state-contingent) adoption and token prices for $j = H, L$ after time τ, given by

$$N_j(\alpha) = N_j = \left(\frac{A_j}{\max\{0, vr - \alpha\mu_j\} + \phi}\right)^{\frac{1}{1-\xi}}$$

and

$$P_t = P_j = \begin{cases} v\left(\frac{A_j}{vr-\alpha\mu_j+\phi}\right)^{\frac{1}{1-\xi}} & \text{if } vr > \alpha\mu_j \\ \frac{\alpha\mu_j}{r}\left(\frac{A_j}{\phi}\right)^{\frac{1}{1-\xi}} & \text{if } vr \le \alpha\mu_j. \end{cases}$$

In contrast, the bad firm chooses cash flow rights α_L and retention β_L. As the bad firm (platform) is inefficient (i.e., has negative NPV) and so does not receive financing, it follows that in a separating equilibrium, the payoff of a bad firm is equal to zero. On the other hand, mimicking the good firm and setting $\alpha_L = \alpha$ and retaining $\beta_L = \beta$ tokens yields strictly positive payoff $\frac{\Lambda\beta P_L}{r+\Lambda} > 0$. As a result, whenever the bad platform has negative NPV, there does not exist a separating equilibrium. Therefore, we study in the following a pooling equilibrium.

5.15.1 Pooling equilibrium

In a pooling equilibrium, both good and bad firms grant cash-flow rights α to token holders and retain initially β tokens. After time τ, token price equals P_H, if the firm is of good type and developers exert sufficient effort, and equals P_L otherwise. As in the baseline model, we assume that exerting effort is efficient and focus therefore on pooling equilibria, in which a good firm has productivity A_H after time τ and developers exert effort.

At time zero, given a fraction π of good firms, the token price equals

$$\overline{p} := \frac{\Lambda}{r+\Lambda}(\pi P_H + (1-\pi)P_L).$$

Any firm sells at time zero the minimal amount of tokens needed to cover initial financing needs I, in that the retention level is given by

$$\beta = 1 - \frac{I}{\overline{p}}.$$

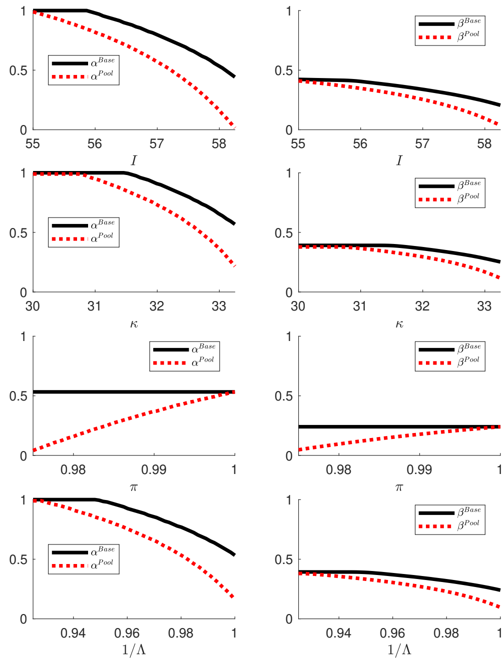

FIGURE 5.7: Comparison pooling equilibrium and baseline model.

Adverse selection worsens the financing conditions of a good firm. As a consequence, with adverse selection, developers (operating a good firm) must sell more tokens $1-\beta$ at time zero to raise funds I, thereby reducing the retention level β and developers' incentives.

For developers to have sufficient incentives to exert effort over $(0, \tau)$, the incentive condition (5.13) has to hold, i.e.,

$$IC(\alpha) := \underbrace{\Lambda\left(\beta P_H + \frac{(1-\alpha)\mu_H N_H}{r}\right) - \kappa}_{\text{Payoff under } a_t=1} - \underbrace{\Lambda\beta P_L}_{\text{Payoff under } a_t=0} \geq 0.$$

Recall that by assumption 2, $\mu_L = 0$. Finally, a good firm's problem boils down to solving

$$\max_{\alpha[0,1]} \Lambda\left(\beta P_H + \frac{(1-\alpha)\mu_H N_H}{r}\right) - \kappa \quad \text{s.t.} \quad IC(\alpha) \geq 0, \beta = 1 - \frac{I}{\bar{p}}.$$

In contrast, a bad firm chooses α and β to mimic a good firm, leading to (scaled) payoff $\Lambda\beta P_L > 0$. Because the bad firm's productivity is low with certainty, there is no moral hazard problem for bad type firms.

Fig. 5.7 illustrates the effects of introducing adverse selection on outcome variables. The top panels show that adverse selection reduces the level of security features attached to tokens of good platforms for two reasons. First, due to adverse selection (which implies that $\bar{p} \leq P_H$), developers have to sell more tokens at time zero to cover their initial financing needs, leading to lower initial token retention β. To maintain incentive compatibility, developers, in turn, must possess more equity incentives, which requires to grant less cash-flow rights to token holders. Second, by increasing α and spurring platform adoption, developers of a good platform increase the token price P_H but may reduce the average token price $P_H\pi + P_L(1-\pi)$. This is because granting cash-flow rights to token holders may reduce platform value in case the platform happens to be of low quality with $A_t = A_L$ for $t \geq \tau$. In fact, Proposition 18 highlights that granting cash-flow rights to token holders is only optimal if platform productivity is sufficiently high.

Fig. 5.7 demonstrates that even mild adverse selection can lead to a substantial reduction in token security features α and developers' token retention β. In effect, we pick the value $\pi = 0.98$ as the base case, because lower values of π make it inefficient to finance the platform. Fig. 5.7 also demonstrates that introducing adverse selection has no bearing on the predictions of the model regarding the effects of financing needs (I), the cost of effort (κ), or the expected time to platform development ($1/\Lambda$)) on the optimal level of retention (β) and token security features (α).

5.15.2 Alternative model assumptions and separating equilibrium

This section relaxes Assumption 2 by considering that the project has positive NPV and may produce cash-flows with $\mu_L > 0$ even if developers do not exert effort, in that

$$\max_{\alpha\in[0,1]} \frac{\Lambda}{r+\Lambda}\left(P_L + \frac{(1-\alpha)\mu_L N_L}{r}\right) > I. \tag{5.35}$$

As a result, a bad platform with low productivity A_L and cash-flow rate μ_L has a positive NPV and receives financing in frictionless markets. Under this assumption there can be a separating equilibrium. In the following, we assume that a bad platform has sufficiently low productivity and hence sufficiently low cash-flows in that $\mu_L \leq (1-\xi)\phi - vr\xi$, where $(1-\xi)\phi - vr\xi > 0$, so that under perfect information it is optimal to set cash-flow rights to $\alpha = \alpha_L \equiv 0$ for a bad type platform. Also recall that due to $vr > \mu_H \geq \mu_L$, it holds that $P_i = vN_i$ for $i = H, L$.

Consider first the problem of a bad type firm. By Proposition 18 (by replacing μ_H by μ_L in the statement of the Proposition), it is optimal to set token security features of a bad platform to $\alpha_L = 0$ when a bad project receives financing, because of $\mu_L \leq (1-\xi)\phi - vr\xi$. Developers' payoff is therefore given by

$$V_L := \frac{\Lambda}{r+\Lambda}\left(v + \frac{\mu_L}{r}\right)\left(\frac{A_L}{vr+\phi}\right)^{\frac{1}{1-\xi}} - I = \frac{\Lambda}{r+\Lambda}\left(\beta_L v + \frac{\mu_L}{r}\right)\left(\frac{A_L}{vr+\phi}\right)^{\frac{1}{1-\xi}},$$

which is positive, by Eq. (5.35), and where developers' initial retention level satisfies

$$\beta_L := 1 - \frac{(\Lambda+r)Iv}{\Lambda}\left(\frac{A_L}{vr+\phi}\right)^{\frac{-1}{1-\xi}}. \tag{5.36}$$

By (5.35) and the optimality of $\alpha = \alpha_L = 0$, it follows that $\beta_L \in [0,1]$.

When there are no frictions, developers of a good platform optimally set $\alpha = 1$. In the presence of moral hazard, they choose the highest level of $\alpha \in [0,1]$ that satisfies the incentive compatibility constraint $IC(\alpha) \geq 0$ (see Assumption 2 and Proposition 18). In addition, developers retain the maximum amount of tokens subject to their financing needs, in that $\beta = 1 - \frac{(\Lambda+r)Iv}{\Lambda N_H}$.

By Proposition 18, platforms with low (high) cash flows optimally feature tokens with low (high) security features. Because a bad platform produces low cash flows, a high level of token security features may differentiate a good platform project from a bad one. In other words, token security features can serve as a signal for good platform quality. That is, the benefit of mimicking the high type is that the low type can sell tokens at a higher price at time zero and thus retain more tokens (i.e., mimicking reduces the cost of investment). The cost of mimicking the high type is the increase in token security features, which reduces platform value (i.e., mimicking reduces the benefit of investment).

In the following, we therefore look for a separating equilibrium in which developers of a good platform choose α according to

$$\max \alpha \text{ s.t. } IC(\alpha) \geq 0 \text{ and } \beta = 1 - \frac{(\Lambda + r)Iv}{\Lambda N_H}$$

and developers of a bad platform choose $\alpha_L = 0$ and $\beta_L = 1 - \frac{(\Lambda+r)Iv}{\Lambda}\left(\frac{A_L}{vr+\phi}\right)^{\frac{-1}{1-\xi}}$.

If it exists, this separating equilibrium is also the least cost separating equilibrium because in equilibrium, developers do not face additional optimization constraints (relative to the baseline model) and, in fact, solve (5.8). We also emphasize that because developers are risk neutral, token retention is not costly for platform developers and hence does not signal platform quality. Chod and Lyandres (2019) consider a risk-averse entrepreneur launching a platform. In their framework, token retention is costly for the entrepreneur, and there exists a separating equilibrium in which token retention signals a high platform quality. By assuming risk neutrality, we eliminate the signaling effect of token retention and highlight that token security features signal high platform quality too. Our findings can be viewed as complementary to those in Chod and Lyandres (2019).

In a separating equilibrium, it must hold that the bad type does not want to mimic the good type. When the bad type mimics the good type, it can sell tokens at time zero at a higher price and thus can retain more tokens, i.e., $\beta > \beta_L$. The cost of mimicking the high type is that α increases, which reduces platform value. If the bad type does not mimic the good type, its payoff equals V_L. By contrast, its payoff upon mimicking the good type reads $V_L^{mimic} := \frac{\Lambda}{r+\Lambda}\left(\beta v + \frac{(1-\alpha)\mu_L}{r}\right) N_L$. In a separating equilibrium, the following must hold

$$V_L \geq \frac{\Lambda}{r + \Lambda}\left(\beta v + \frac{(1-\alpha)\mu_L}{r}\right) N_L = V_L^{mimic}. \tag{5.37}$$

Moreover, provided that exerting full effort is efficient, the good type firm does not have incentives to mimic the bad type of firm. Hence, (5.37) is sufficient for the existence of the separating equilibrium. The reason is that in the proposed separating equilibrium, the good type solves (5.8) and thus does not face additional optimization constraints (relative to the baseline model). That is, in the separating equilibrium, the good type already chooses the retention level β and token security features α that maximizes her payoff subject to the incentive compatibility constraint (5.13) and the financing constraint (5.3). Therefore, the good type cannot do better by choosing different levels of β and α.

Fig. 5.8 examines the effects of a firm's environment on the existence of a separating equilibrium. To make sure that the bad type platform has positive NPV, we depart from our baseline parameter values along two dimensions: we set $A_L = 0.9$ (instead of $A_L = 0.55$) and $I = 1$ (instead of $I = 58$). We also consider that cash flows are given by $\mu_H = \mu = 0.025$ and $\mu_L = 0$. Fig. 5.8 shows that when financing needs are small, the separating equilibrium described above exists. Indeed, when financing needs I are sufficiently low, β and β_L are low and so is the benefit of mimicking. In this case, (5.37) is satisfied, and the separating equilibrium exists. When κ and $1/\Lambda$ are sufficiently large, the development effort is no longer efficient, and the good type prefers to mimic the bad type so that there is no separating equilibrium. Fig. 5.8 also shows that strong network effects (ξ), high platform transaction frictions (v), or a high cash-flow rate (μ) facilitate the existence of the separating equilibrium, in line with the above discussion.

In this separating equilibrium and provided that $\bar{\alpha} = 1$ (see Assumption 2), adverse selection does not change the optimal level of token security features α and the optimal retention level β compared to the base case model. The reason is that signaling is de facto costless for the good type: the good type chooses the highest level of α satisfying incentive compatibility (5.13), which at the same time makes it (most) costly for the low type to mimic the high type. The following section relaxes Assumption 2 and considers parameter configurations with $\bar{\alpha} < 1$, so that adverse selection may boost the provision of token security features in a separating equilibrium.

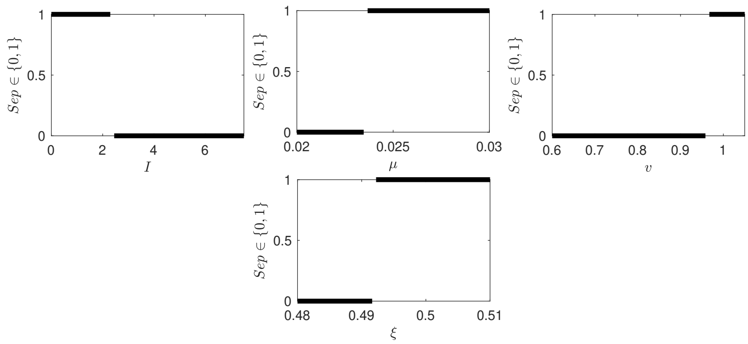

FIGURE 5.8: When does a separating equilibrium exist? $Sep = 1$ indicates that the prescribed separating equilibrium exists. Parameter values are set as in the base case environment except for $A_L = 0.9$ and $I = 1$ to ensure that the bad type platform has positive NPV.

As a result, adverse selection only affects the provision of token security features in the pooling equilibrium. Interestingly, adverse selection and moral hazard may interact and reinforce each other and hence jointly curb the provision of token security features. Moral hazard requires developers to possess sufficient equity incentives and leads to low token security features α. Low token security features α make it attractive for the low type to mimic the high type, thereby destabilizing the separating equilibrium and leading to a pooling equilibrium, which exacerbates moral hazard and reduces token security features even further (as shown in Appendix 5.15.1).

5.15.3 Costly signalling

In the previous section, signalling is de-facto costless for the good type firm, in that the good type firm finds it optimal to choose $\alpha \in [0,1]$ as high as possible, subject to incentive compatibility. This is a consequence of Assumption 3 which implies that absent friction $\alpha = \bar{\alpha} = 1$ is optimal. This section considers parameter configurations such that $\bar{\alpha} < 1$ is optimal for the good type firm and $\alpha_L = 0$ is optimal for the bad type firm, absent frictions. That is, input parameter values are such that

$$\bar{\alpha} = \frac{vr}{\mu_H} + \frac{1}{\xi} - \frac{\phi(1-\xi)}{\xi\mu_H} \in (0,1).$$

In addition, we consider that $vr > \mu_H$, so tokens are priced according to utility features.

We look for a separating equilibrium in which the good type firm chooses token security features α and token retention β, while the bad type firm chooses token security features $\alpha_L = 0$ and token retention β_L (with β_L characterized in (5.36)). In the separating equilibrium, the bad type firm must not have incentives to mimic the good type firm, so that (5.37) must hold. Next, consider the good type firm. In equilibrium, the good type firm's payoff equals

$$V_H := \frac{\Lambda}{r+\Lambda}\left(\beta v + \frac{(1-\alpha)\mu_H}{r}\right) N_H,$$

where $\beta = 1 - \frac{(\Lambda+r)Iv}{\Lambda N_H}$ and N_H is a function α, i.e., $N_H = N_H(\alpha)$. Alternatively, the good type can pick a a level of token security features $\alpha' \neq \alpha$, in which case the market perceives the good type firm as bad type firm. To cover initial financing needs I, the good type firm's retention level equals then β_L. As a result, upon deviating, the good type firm's payoff

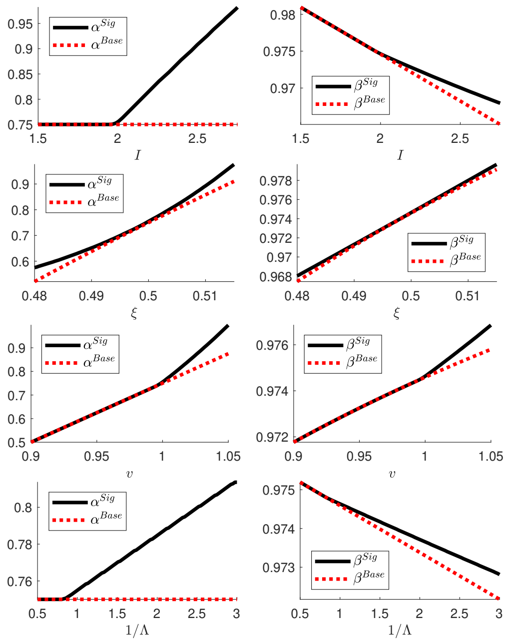

FIGURE 5.9: Comparison separating equilibrium and baseline model. We pick the parameters $I = 2 > 0 = \kappa$, $\mu_H = 0.2 > 0.1 = \mu_L$ and $A_L = 0.9$.

equals

$$V_H^{Dev} := \max_{\alpha' \in [0,1]} \left(\frac{\Lambda}{r+\Lambda} \left(\beta_L v + \frac{(1-\alpha')\mu_H}{r} \right) N_H \right),$$

subject to $IC(\alpha') \geq 0$, where N_H is a function of α', i.e., $N_H = N_H(\alpha')$. For the good type not to deviate, it must be that

$$V_H \geq V_H^{Dev}. \tag{5.38}$$

In a separating equilibrium, both (5.37) and (5.38) have to be satisfied.

Provided there exists at least one separating equilibrium, we focus on the least cost separating equilibrium, which can be found by solving

$$\max_{\alpha \in [0,1]} V_H \quad \text{s.t.} \quad (5.37) \quad \text{and} \quad (5.38).$$

This amounts to

$$\min_{\alpha \geq \bar{\alpha}} \alpha \quad \text{s.t.} \quad (5.37) \quad \text{and} \quad (5.38).$$

In the following, we assume a separating equilibrium exists and compare the model outcomes in the least cost separating equilibrium with those of the baseline model.[17] Fig. 5.9 illustrates that in the least cost separating equilibrium, the good type firm chooses higher α (i.e., $\alpha = \alpha^{Sig}$) than in the baseline model (i.e., $\alpha = \alpha^{Base}$). The reason is that by attaching high token security features, a good type firm signals good platform quality. Because a higher level of α boosts the token price, the good type firm's level of initial token retention β^{Sig} in the least cost separating equilibrium is higher than in the baseline model, i.e., than

[17] The question of the existence of a separating equilibrium was already discussed in the previous section.

β^{Base}. The intuition is that a good type firm signals both by token retention and attaching security features to tokens. Fig. 5.9 demonstrates that these effects are robust to parameter changes.

5.16 Transaction fees

Suppose now that developers can dynamically charge a fee $f > 0$ to users for transacting on the platform. This fee increases users' direct cost of transacting to $f + \phi$ and platform cash flows to

$$dD_t = (\mu(A_t) + f)N_t dt,$$

so the level of platform adoption becomes

$$N_H = \left(\frac{A_H}{\max\{0, vr - \alpha(\mu_H + f)\} + \phi + f}\right)^{\frac{1}{1-\xi}}, \tag{5.39}$$

when the platform charges transaction fees. In the following, we consider that developers cannot commit to future transaction fees. Appendix 5.16.3 analyzes the case of full commitment. For simplicity, we abstract from moral hazard w.r.t. effort by taking $\kappa, 1/\Lambda$ or I sufficiently small.

Without commitment, the optimal dynamic fee f maximizes at each point in time $t \geq \tau$ the dividends accruing to developers:

$$(1-\alpha)(\mu_H + f)N_H$$

and therefore maximizes platform cash flows $(\mu_H + f)N_H$.[18] This leads to the following result.

Proposition 23 *The optimal dynamic fee for platform developers satisfies*

$$f^* = \min\left\{\frac{(1-\xi)(vr+\phi) - (1-\alpha\xi)\mu_H}{\xi(1-\alpha)}, \frac{vr}{\alpha} - \mu_H\right\}.$$

If $(1-\xi)(vr+\phi) > \mu_H$, the fee increases in α for $\alpha \leq \alpha_1$ and decreases in α for $\alpha \geq \alpha_1$, where $\alpha_1 \in (0,1)$ is the unique solution to

$$\frac{(1-\xi)(vr+\phi) - (1-\alpha\xi)\mu_H}{\xi(1-\alpha)} = \frac{vr}{\alpha} - \mu_H.$$

The resulting adoption level satisfies

$$N_H^f = \begin{cases} \left(\frac{A_H\xi}{vr+\phi-\mu_H}\right)^{\frac{1}{1-\xi}} & \text{if } vr > \alpha(f+\mu_H) \\ \left(\frac{A_H}{vr/\alpha+\phi-\mu_H}\right)^{\frac{1}{1-\xi}} & \text{otherwise.} \end{cases}$$

Proposition 11 shows that the optimal dynamic fee depends on whether the token utility or security features pin down the token price (i.e., whether $vr > \alpha(f + \mu_H)$ or $vr \leq \alpha(f + \mu_H)$, respectively). If $(1-\xi)(vr+\phi) > \mu_H$, the optimal fee follows a hump-shaped pattern in α. The optimal level of security features with endogenous transaction fees is then characterized in the following corollary.

Corollary 7 *Tokens with $\alpha = 0$ are optimal if and only if*

$$(1-\xi)(\phi - \mu_H) \geq vr(\xi^{\frac{\xi}{\xi-1}} - 1). \tag{5.40}$$

Tokens with $\alpha = 1$ are optimal if and only if condition (5.40) *is not satisfied. Platform adoption is higher for $\alpha = 1$ than for $\alpha = 0$.*

[18] We assume that even if $\alpha = 1, \beta = 0$ developers set fees in order to maximize $(\mu_H + f)N_H$.

If the token is priced according to its utility features, then the optimal transaction fee satisfies $f^* < \frac{vr}{\alpha} - \mu_H$. In this case, users effectively incur the transaction fee $f^*(1-\alpha)$. The reason is that a fraction α of the transaction fees flows back to users in the form of dividends. Higher α in turn implies higher dividends, which allows developers to charge higher fees without endangering adoption. In this context, the issuance of a utility token (i.e., $\alpha = 0$) can be viewed as a commitment device not to charge high fees in the future.

If the token is priced according to its security features, then $f^* = \frac{vr}{\alpha} - \mu_H$ and developers charge lower transaction fees as token cash flow rights increase, in an attempt to limit purely return-driven investments and maximize adoption. Also note that, even if $\alpha = 1$, the proceeds from transaction fees do not fully flow back to users but partially accrue to (return-driven) token investors, who do not transact. In sum, both high and low token security features serve as commitment device for low future transaction fees and, thus, are particularly useful for platform building in the presence of commitment problems to future fees. As a result, either $\alpha = 1$ or $\alpha = 0$ is optimal.

Interestingly, the optimal transaction fee f can be negative. In this case, the startup firm subsidizes the user base in order to accelerate platform adoption.

Corollary 8 *Subsidies $f < 0$ are optimal if $\mu_H > \mathcal{S} := \frac{(1-\xi)(vr+\phi)}{1-\alpha\xi}$. If developers can commit to a fee structure $\{f\}$ at time zero, subsidies are optimal if $\mu_H > \mathcal{S} - \frac{vr}{1-\xi(1-\alpha)}$.*

As shown in Corollary 8, subsidies to the user base are more likely if the platform is financed with utility tokens (i.e., for $\alpha = 0$) or if the network effects are strong. In addition, subsidies are only optimal if the platform generates enough revenues μ_H to finance these subsidies. We also show that subsidies are more likely if the blockchain technology facilitates commitment.

5.16.1 Proof of Proposition 11

Define $\varepsilon = 1/(1-\xi)$. Note that $N_H = \left(\frac{A_H}{\max\{0, vr-\alpha(\mu_H+f)\}+\phi+f}\right)^{\frac{1}{1-\xi}}$ and that the optimal fee $f = f^*$ is such that: $f^* = \arg\max_{f\geq 0}(\mu_H + f)N_H$.

1. Assume that $vr - \alpha(\mu_H + f) > 0$. Then, the FOC $\frac{\partial(\mu_H+f)N_H}{\partial f} = 0$ must hold in optimum. That is,

$$
\begin{aligned}
0 = N_H - (\mu_H + f)\frac{\partial N_H}{\partial f} &= N_H - \frac{\varepsilon N_H(1-\alpha)(\mu_H+f)}{vr - \alpha\mu_H + \phi + f(1-\alpha)} \\
&\propto vr - \alpha\mu_H + \phi + f(1-\alpha) - \varepsilon(1-\alpha)(\mu_H + f).
\end{aligned}
$$

   Thus: $(1-\alpha)f = \frac{(vr+\phi)(1-\xi)-\mu_H}{\xi} + \alpha\mu_H$ for optimal $f = f^*$.

   Plugging the optimal fee expression into (5.39) yields the desired expressions for platform adoption, i.e.,

$$
\max_{\alpha}\left(\beta P_H + \frac{(1-\alpha)\mu_H N_H}{r}\right) - \frac{\kappa}{\Lambda}
$$

   and platform value (surplus), i.e.,

$$
\left(v + \frac{(1-\xi)(vr+\phi-\mu_H)}{\xi r}\right)\left(\frac{A_H\xi}{\phi+vr-\mu_H}\right)^{\frac{1}{1-\xi}}.
$$

   Both expressions do not depend explicitly on α. It follows that the developers' payoff and overall platform value (surplus) do not depend explicitly on α either.

2. Next, we assume that $vr \leq \alpha(\mu_H + f)$, implying that $N_H = \left(\frac{A_H}{\phi+f}\right)^{\frac{1}{1-\xi}}$. Then, if $vr < \alpha(\mu_H + f)$, the FOC $\frac{\partial(\mu_H+f)N_H}{\partial f} = 0$ must hold, so that

$$0 = N_H - \varepsilon\frac{\mu_H + f}{\phi + f}N_H \propto 1 - \varepsilon\frac{\mu_H + f}{\phi + f} \implies f = \frac{\phi - \varepsilon\mu_H}{\varepsilon - 1} = \frac{(1-\xi)\phi - \mu_H}{\xi}.$$

Otherwise, if $vr = \alpha(\mu_H + f)$, then $f = \frac{vr}{\alpha} - \mu_H = \frac{vr - \alpha\mu_H}{\alpha}$. Altogether, $f = f^* = \max\left\{\frac{vr-\alpha\mu_H}{\alpha}, \frac{(1-\xi)\phi-\mu_H}{\xi}\right\}$. Assumption 2 implies that $\frac{(1-\xi)\phi-\mu_H}{\xi} < \frac{vr}{\alpha} - \mu_H \Longleftarrow (1-\xi)(\phi - \mu_H) < vr\xi$, hence $f = f^* = \frac{vr-\alpha\mu_H}{\alpha}$.

In sum, we have shown that:[19]

$$f^* = \min\left\{\frac{(1-\xi)(vr+\phi) - (1-\alpha\xi)\mu_H}{\xi(1-\alpha)}, \frac{vr}{\alpha} - \mu_H\right\}.$$

If $(1-\xi)(vr+\phi) > \mu_H$, the first expression in the “min" operator increases in α while the second expression decreases in α. The first expression in the “min" operator tends to ∞ as $\alpha \to 1$ while the second one is always positive (due to $vr \geq \mu_H$) and tends to ∞ as $\alpha \to 0$. Hence, there exists a unique cutoff $\alpha_1 \in (0,1)$ solving $\frac{(1-\xi)(vr+\phi)-(1-\alpha\xi)\mu_H}{\xi(1-\alpha)} = \frac{vr}{\alpha} - \mu_H$ (in α). Below α_1, the payoff does not explicitly depend on α, as shown before.

5.16.2 Proof of Corollary 7

First, consider that α is such that $f^* = \frac{vr}{\alpha} - \mu_H$, which implies the adoption level $N_H = \left(\frac{A_H}{vr/\alpha-\mu_H+\phi}\right)^{1/(1-\xi)}$ and the price $P_H = \frac{\alpha(\mu_H+f)N_H}{r} = vN_H$. This is the case when $\alpha = 1$. Thus, the overall surplus is

$$S(\alpha) = vP_H + (1-\alpha)\frac{\mu_H + f}{r} = N_H\left(v + (1-\alpha)\frac{v}{\alpha}\right)N_H.$$

Next, for $\alpha > 0$:

$$\begin{aligned}
S'(\alpha) &= -\frac{v}{\alpha}N_H - \frac{(1-\alpha)v}{\alpha^2}N_H + \left(v + (1-\alpha)\frac{v}{\alpha}\right)N_H'(\alpha) \\
&\propto -v\alpha - (1-\alpha)v + \varepsilon\left(v + (1-\alpha)\frac{v}{\alpha}\right)\frac{vr}{vr/\alpha - \mu_H + \phi} \\
&\propto -v\alpha + \frac{\varepsilon v^2 r}{vr/\alpha - \mu_H + \phi} \propto -1 + \frac{\varepsilon vr}{vr + \alpha(\phi - \mu_H)}.
\end{aligned}$$

Hence, $S'(\alpha) = 0$ is solved by: $\varepsilon vr = vr + \alpha(\phi - \mu_H)$ so that $\alpha = \min\left\{\frac{\xi vr}{(1-\xi)(\phi-\mu_H)}, 1\right\}$ is optimal under these circumstances. Assumption 2 then implies that $\alpha = 1$, leading to adoption $N_H = \left(\frac{A_H}{\phi+vr-\mu_H}\right)^{\frac{1}{1-\xi}}$ and payoff (surplus) $v\left(\frac{A_H}{\phi+vr-\mu_H}\right)^{\frac{1}{1-\xi}}$. Note that for $\alpha = 1$: $f^* = vr - \mu_H$.

Second, consider α is such that $f^* = \frac{(1-\xi)(vr+\phi)-\mu_H}{\xi(1-\alpha)} + \frac{\alpha\mu_H}{1-\alpha}$. This is the case when $\alpha = 0$. In this case, the payoff does not depend on α (see previous results) and surplus (payoff) is given by $\left(v + \frac{(1-\xi)(vr+\phi-\mu_H)}{\xi r}\right)\left(\frac{A_H\xi}{\phi+vr-\mu_H}\right)^{\frac{1}{1-\xi}}$.

In sum, $\alpha \in \{0,1\}$ is optimal. Note that $\alpha = 0$ is optimal for developers if it leads to higher overall platform value (i.e., surplus) than $\alpha = 1$ (as developers can extract all residual

[19] For $\alpha = 1$, the expression for f^* with some slight abuse of notation simply becomes $\frac{vr}{\alpha} - \mu_H$.

payoff). Thus, $\alpha = 0$ is optimal if

$$\left(v + \frac{(1-\xi)(vr+\phi-\mu_H)}{\xi r}\right)\left(\frac{A_H\xi}{\phi+vr-\mu_H}\right)^{\frac{1}{1-\xi}} \geq v\left(\frac{A_H}{\phi+vr-\mu_H}\right)^{\frac{1}{1-\xi}}$$
$$\iff \left(v + \frac{(1-\xi)(vr+\phi-\mu_H)}{\xi r}\right)\xi^{\frac{1}{1-\xi}} \geq v \iff ((1-\xi)(\phi-\mu_H)+vr)\,\xi^{\frac{\xi}{1-\xi}} \geq vr$$
$$\iff (1-\xi)(\phi-\mu_H) \geq vr(\xi^{\frac{\xi}{\xi-1}}-1).$$

5.16.3 Full commitment to transaction fees

Blockchain technology facilitates commitment to various metrics of platform and token design. For example, Cong et al. (2020a) demonstrate that the commitment to predetermined rules of token supply stimulates platform building. In this section, we analyze the effects of full commitment to future transaction fees.

In line with economic intuition, Corollary 9 shows that developers charge lower transaction fees and that adoption is higher under full commitment.

Corollary 9 *Assume full commitment and $\phi > \mu_H$. Users incur the transaction fee:*

$$f^* = \min\left\{\frac{(1-\xi)\phi - \xi vr - (1-\alpha\xi)\mu_H}{\xi(1-\alpha)}, \frac{vr}{\alpha} - \mu_H\right\}.$$

If $(1-\xi)\phi - \xi vr - \mu_H > 0$, the fee increases in α for $\alpha \leq \alpha_2$ and decreases in α for $\alpha \geq \alpha_2$, where $\alpha_2 \in (0,1)$ is the unique solution to

$$\frac{(1-\xi)\phi - \xi vr - (1-\alpha\xi)\mu_H}{\xi(1-\alpha)} = \frac{vr}{\alpha} - \mu_H.$$

This implies the adoption level:

$$N_H^f = \begin{cases} \left(\frac{A_H\xi}{\phi-\mu_H}\right)^{\frac{1}{1-\xi}}, & \text{if } vr > \alpha(f+\mu_H) \\ \left(\frac{A_H}{vr/\alpha+\phi-\mu_H}\right)^{\frac{1}{1-\xi}}, & \text{otherwise.} \end{cases}$$

Remarkably, we find that the issuance of a utility token makes developers optimize platform adoption instead of cash flows under full commitment to transaction fees. It therefore follows that the ability to commit makes ICOs relatively more valuable.

Corollary 10 *Assume full commitment to fees $\{f\}$ and $\phi > \mu_H$. Then, $\alpha = 0$ is optimal if and only if:*

$$(1-\xi)(\phi-\mu_H) \geq vr\left(\left(\frac{(\phi-\mu_H)}{(\phi-\mu_H+vr)\xi}\right)^{\frac{1}{1-\xi}}\xi^{\frac{\xi}{\xi-1}}\right). \tag{5.41}$$

Otherwise, $\alpha = 1$ is optimal.

Proof of Corollary 9

1. Assume that $vr > (\mu_H + f)\alpha$. Under full commitment, developers choose the fee f in order to maximize (given α) $S(\alpha)$. Note that

$$\frac{\partial S(\alpha)}{\partial f} \propto (vr + (\mu_H+f)(1-\alpha))\,N_H'(f) + (1-\alpha)N_H(f)$$
$$\propto 1 - \frac{1}{1-\xi}\frac{(vr+(\mu_H+f)(1-\alpha))}{vr+\phi-\mu_H\alpha+(1-\alpha)f}.$$

In optimum, the FOC $\frac{\partial S(\alpha)}{\partial f} = 0$ must hold. We can solve for optimal $f = f^*$ via $(1-\alpha)f = \frac{\phi(1-\xi)-\mu_H-\xi vr}{\xi} + \alpha\mu_H$ leading to the adoption level $N_H^* = \left(\frac{A_H\xi}{\phi-\mu_H}\right)^{\frac{1}{1-\xi}}$.

2. Next, consider $vr \leq (\mu_H + f)\alpha$. If $vr = (\mu_H + f)\alpha$, then $f = \frac{vr}{\alpha} - \mu_H$ and $N_H^* = \left(\frac{A_H\xi}{\frac{vr}{\alpha}+\phi-\mu_H}\right)^{\frac{1}{1-\xi}}$. If $vr < (\mu_H + f)\alpha$, then $N_H = \left(\frac{A_H}{\phi+f}\right)^{\frac{1}{1-\xi}}$ and $P_H = \frac{\alpha(\mu_H+f)N_H}{r}$. We can wlog. assume that $\alpha = 1$. The FOC of maximization is

$$\frac{\partial S(\alpha)}{\partial f} \propto (\mu_H + f)N_H'(f) + N_H(f) \propto 1 - \frac{1}{1-\xi}\frac{\mu_H + f}{vr + f} = 0.$$

We can solve for: $f = \frac{(1-\xi)\phi-\mu_H}{\xi}$.

Overall, if $vr \leq \alpha(\mu_H + f)$, then $f = f^* = \max\left\{\frac{(1-\xi)\phi-\mu_H}{\xi}, \frac{vr}{\alpha} - \mu_H\right\}$. Because of $\frac{(1-\xi)\phi-\mu_H}{\xi} < \frac{vr}{\alpha} - \mu_H \Longleftarrow (1-\xi)(\phi-\mu_H) < vr\xi$, we have that $f = f^* = vr/\alpha - \mu_H$.

In sum, we have shown that $f^* = \min\left\{\frac{(1-\xi)\phi-\xi vr-(1-\alpha\xi)\mu_H}{\xi(1-\alpha)}, \frac{vr}{\alpha} - \mu_H\right\}$, as desired. Consider the equation $\frac{(1-\xi)\phi-\xi vr-(1-\alpha\xi)\mu_H}{\xi(1-\alpha)} = \frac{vr}{\alpha} - \mu_H$. If $(1-\xi)\phi - \xi vr - \mu_H > 0$, the above equation possesses a unique solution on $\alpha_2 \in (0,1)$.

Proof of Corollary 10

As in the proof of corollary 7, it suffices to compare payoffs under the polar cases $\alpha = 0$ and $\alpha = 1$. Notably, $\alpha = 0$ is optimal if and only if:

$$\left(vr + \frac{(1-\xi)(\phi-\mu_H) - \xi vr}{\xi}\right)\left(\frac{A_H\xi}{\phi-\mu_H}\right)^{\frac{1}{1-\xi}} \geq vr\left(\frac{A_H}{vr+\phi-\mu_H}\right)^{\frac{1}{1-\xi}}$$

$$\Longleftrightarrow \frac{(1-\xi)(\phi-\mu_H)}{\xi vr} \geq \left(\frac{(\phi-\mu_H)}{(\phi-\mu_H+vr)\xi}\right)^{\frac{1}{1-\xi}}$$

$$\Longleftrightarrow (1-\xi)(\phi-\mu_H) \geq vr\left(\left(\frac{(\phi-\mu_H)}{(\phi-\mu_H+vr)\xi}\right)^{\frac{1}{1-\xi}}\xi^{\frac{\xi}{\xi-1}}\right).$$

5.16.4 Proof of Corollary 8

First, absent commitment, the fee levied reads $f^* = \min\left\{\frac{(1-\xi)(vr+\phi)-(1-\alpha\xi)\mu_H}{\xi(1-\alpha)}, \frac{vr}{\alpha} - \mu_H\right\}$, which is—due to $vr \geq \mu_H$ —negative if and only if $\frac{(1-\xi)(vr+\phi)-(1-\alpha\xi)\mu_H}{\xi(1-\alpha)} < 0$, i.e., if and only if $\mu_H > \mathcal{S} := \frac{(1-\xi)(vr+\phi)}{1-\alpha\xi}$.

Second, with full commitment to a fee structure at time zero, Proposition 10 implies the optimal fee is given by $f^* = \min\left\{\frac{(1-\xi)\phi-\xi vr-(1-\alpha\xi)\mu_H}{\xi(1-\alpha)}, \frac{vr}{\alpha} - \mu_H\right\}$, which is smaller than zero if and only if $\frac{(1-\xi)\phi-\xi vr-\mu_H}{\xi} + \alpha\mu_H < 0$, i.e., if and only if $\mu_H > \mathcal{S} - \frac{vr}{1-\xi\alpha}$.

5.17 Dynamic trading

The objective of this Appendix is to introduce richer trading dynamics in the model by considering the role of speculators. Specifically, we consider that there are risk neutral speculators with discount rate $\rho < r$, capturing the notion that speculators are financially less constrained or more diversified than users and developers. Developers cannot commit at time zero to their trading of tokens.[20] In addition to speculators, we introduce convex costs of effort $\frac{\kappa a^2}{2}$ and assume that with effort $a \in [0, \bar{a}]$, the project succeeds at time τ and $A_t = A_H$ with probability pa and the project fails at time τ with $A_t = A_L$ with probability $1 - pa$. The convex cost is needed to generate smooth trading patterns, as will become clear below. We impose that $\bar{a} \leq 1$ and $p \in (0, 1)$ and assume throughout that optimal effort is interior.

We look for a Markov Perfect equilibrium with state variable β, where $d\beta_t = \eta_t dt - \beta_t \mathbf{1}_{\{t=\tau\}}$. That is, developers optimally sell all retained tokens at time τ because, as in the baseline model, there is no moral hazard problem after time τ. After time τ, the token price (adoption level) is given by P_H (N_H), if $A = A_H$, and is given by P_L (N_L), if $A = A_L$. Before time τ, the token price is a function of β, $P(\beta)$, and the developers's value function is also a function of β, $V(\beta)$; in addition, the developers trade tokens at endogenous rate η_t for $t < \tau$.

Fix α, which is chosen at time zero, and consider the developers' problem in state β. Define the developers' payoff from reaching the milestone with $A = A_i$:

$$T_i(\beta) := \beta P_i + \frac{(1-\alpha)\mu_i N_i}{r}.$$

This payoff consists of the value of token sales at the milestone, βP_i, and the present value value of future dividends, $\frac{(1-\alpha)\mu_i N_i}{r}$. Developers' value function $V(\beta)$ before time τ reads then

$$(r + \Lambda)V(\beta) = \max_{\eta, a} \left\{ \Lambda \left(paT_H(\beta) + (1 - pa)T_L(\beta) \right) - \frac{\kappa a^2}{2} + \eta(V'(\beta) - P(\beta)) \right\},$$

where the last term in the brackets captures the effects of trading. Thus, if effort a is interior (i.e., $a < \bar{a}$), it is given by

$$a = \frac{\Lambda p}{\kappa}(T_H(\beta) - T_L(\beta)). \tag{5.42}$$

That is, incentives are captured by the difference $T_H(\beta) - T_L(\beta)$. It is easy to see that $T_H(\beta) - T_L(\beta)$ increases in β, so that token retention incentivizes effort.

Using arguments similar to those presented in DeMarzo and Urošević (2006), one can show that in equilibrium, developers are indifferent between buying and selling tokens. That is, $\frac{\partial V(\beta)}{\partial \eta} = 0$ whenever $\beta \in (0, 1)$, i.e.,

$$P(\beta) = V'(\beta). \tag{5.43}$$

The reason is that developers' token sales exacerbate moral hazard and thereby depress platform value and token prices. As developers cannot commit to keeping tokens, they sell tokens and decrease the token price up to the point that they become marginally indifferent between buying and selling tokens. As such, all gains from trade are in equilibrium dissipated by the subsequent rise in agency costs (this observation is also related to that in DeMarzo and He (2020) on the effects of changes in capital structure on shareholder wealth in a no-commitment equilibrium).

We can insert (5.42) and (5.43) back into developers' HJB equation and solve for their value function in closed form

$$V(\beta) = \frac{\Lambda T_L(\beta) + \frac{1}{2\kappa}(\Lambda p(T_H(\beta) - T_L(\beta)))^2}{r + \Lambda}.$$

[20] Similar results could be obtained by assuming instead that users discount at rate $\rho < r$ instead of r. However, to facilitate comparison with the previous sections, we introduce speculators. In line with this assumption, Fahlenbrach and Frattaroli (2019) document that tokens are held by both speculators and platform users.

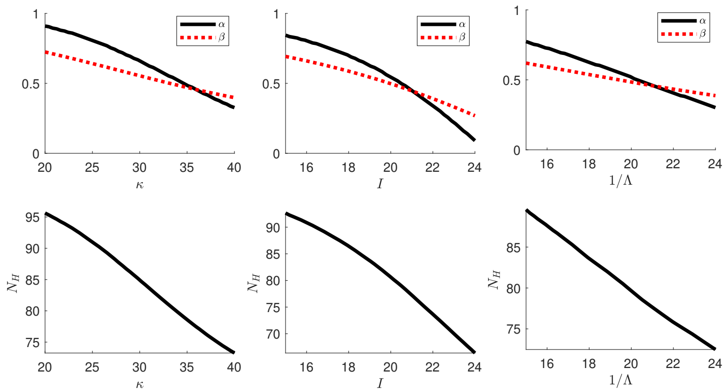

FIGURE 5.10: Model with optimal dynamic trading with $\kappa = 33.33$, $I = 20$, and $p = 0.5$.

Using (5.43) and differentiating the value function with respect to β, we obtain

$$P(\beta) = \frac{\Lambda T_L'(\beta) + \frac{(\Lambda p)^2}{\kappa}(T_H(\beta) - T_L(\beta))(T_H'(\beta) - T_L'(\beta))}{r + \Lambda}.$$

That is, the token price for $t < \tau$ is a function of β, in that $P_t = P(\beta_t)$. Before time τ, speculators are marginal token investors. Since they are risk neutral, they simply need to be compensated for their time preference ρ, in that $\rho P_t dt = \mathbb{E} dP_t$. This can be written as

$$\rho P(\beta) = \Lambda(paP_H + (1 - pa)P_L - P(\beta)) + P'(\beta)\eta.$$

Using this equation, we can solve for the trading rate η in closed form:

$$\eta = \frac{(\rho + \Lambda)P(\beta) - \Lambda(paP_H + (1 - pa)P_L)}{P'(\beta)}.$$

As $\rho < r$ and there are gains for developers from selling tokens, it follows that $\eta < 0$. That is, developers optimally sell their token at a rate before the milestone is reached.

The initial retention level β is set such that $(1 - \beta)P(\beta) = I$ and developers choose α to solve the problem

$$\max_{\alpha \in [0,1]} V(\beta).$$

We solve for the optimal value of α and initial retention $\beta = \beta_0$. Fig. 5.10 presents the model outcomes for different values of κ, I, and $1/\Lambda$. As in the baseline version of the model, an increase of κ, I, or $1/\Lambda$ reduces the provision of security features α as well as the initial retention level β.

5.18 Operating flow costs

Consider that instead of a cost I at time zero, developers incur monetary flow costs idt when developing the project over $[t, t + dt)$. There are no financing frictions and to cover these monetary flow costs, developers sell retained tokens. Developers can stop financing the platform, in which case it cannot be completed and future productivity equals zero. That is, for $t < \tau$, the milestone τ arrives with probability Λdt over $[t, t + dt)$ only if developers cover the development costs idt.

Starting with $\beta_0 = 1$ retained tokens at time zero, developers sell the retained tokens at rate $\eta < 0$ during platform development $[0, \tau)$ to cover development costs idt. In addition, developers sell all retained tokens at the milestone τ. Formally:

$$d\beta_t = \eta_t dt - \beta_t \mathbf{1}_{\{t=\tau\}}.$$

We look for a Markov Perfect Equilibrium with state variable β. Before the milestone is reached (i.e., for $t < \tau$), the developers' value function $V(\beta)$ and the token price $P(\beta)$ are functions of β. To solve the model with flow costs, we first fix a level of α and solve for $V(\beta)$ and $P(\beta)$. We then select the optimal level of α by maximizing developers' payoff at time zero.

As in the baseline version, there is a moral hazard problem, in that developers must exert effort to achieve high platform productivity and exerting full effort is optimal in equilibrium. As a result, the incentive condition (5.13) must be satisfied:

$$IC(\alpha) := \underbrace{\Lambda\left(\beta P_H + \frac{(1-\alpha)\mu_H N_H}{r}\right) - \kappa}_{\text{Payoff under } a_t=1} - \underbrace{\Lambda\left(\beta P_H + \frac{(1-\alpha)\mu_L N_L}{r}\right)}_{\text{Payoff under } a_t=0} \geq 0.$$

Crucially, selling tokens reduces the retention level β, thereby undermining developers' incentives to exert effort. Because users and developers both discount at rate r, there are no gains from trade, so that at any point in time $t < \tau$, developers sell the minimal amount of tokens that is needed to cover financing i, which maximizes incentives and thus is optimal. That is:

$$-\eta P(\beta) = i \iff \eta = \frac{i}{P(\beta)}.$$

For a given level of α, the minimum level of retention $\underline{\beta}$ (depending on α) required to maintain incentive compatibility satisfies $IC(\alpha) = 0$, so that

$$\underline{\beta} = \frac{\kappa}{\Lambda(P_H - P_L)} + \frac{(1-\alpha)(\mu_L N_L - \mu_H N_H)}{r(P_H - P_L)}.$$

In the following, we assume that i) $\underline{\beta} \geq 0$ and ii) the project is inefficient to finance when productivity is low. The latter assumption implies that platform development and financing is terminated once β reaches $\underline{\beta}$ (and the project is never started when $\underline{\beta} \geq 1$).

Conditional on full effort, the developers' value function solves the ODE

$$(r+\Lambda)V(\beta) = \Lambda\left(\beta P_H + \frac{(1-\alpha)\mu_H N_H}{r}\right) + V'(\beta)\eta,$$

subject to $V(\underline{\beta}) = 0$, while the token price solves the ODE

$$(r+\Lambda)P(\beta) = \Lambda P_H + P'(\beta)\eta,$$

subject to $P(\underline{\beta}) = 0$. The trading rate η is given by $\eta = i/P(\beta)$. The solution to this system of coupled ODEs is not available in closed-form.

Finally, the optimal level of token security features α is set to maximize developers' payoff at time zero $V(\beta_0) = V(1)$, so that developers solve

$$\max_{\alpha \in [0,1]} V(1).$$

Chapter 6

Managing Stablecoins[1]

6.1 Introduction

More than a decade ago, Bitcoin heralded a new era of digital payments. Cryptocurrencies challenge the bank-centric payment systems by offering fast and round-the-clock settlement, anonymity, low-cost international remittances, mobile payment for unbanked population, and smart contracting (Brunnermeier, James, and Landau, 2019; Duffie, 2019). However, the substantial volatility exhibited by first-generation cryptocurrencies limits their utility as a means of payment. Stablecoins aim to maintain a stable price against a reference currency, or a basket of currencies by pledging to hold a reserve of fiat currencies or other assets against which stablecoin holdings can be redeemed.[2]

This paper provides the first dynamic model of stablecoins that offers guidance to practitioners and lends itself to an evaluation of regulatory proposals. In our model, optimal stablecoin management features a rich set of strategies that are commonly observed in practice (Bullmann, Klemm, and Pinna, 2019), such as over-collateralization, dynamic reserve management and open market operations, transaction fees or subsidies, targeted price band, re-pegging, and issuance and repurchase of "secondary units" that function as equity shares of stablecoin platforms.

Our model addresses the fundamental questions on the credibility and sustainability of a fixed exchange rate. Bank deposits mediate payments in the traditional payment system with one-to-one convertibility to a fiat currency safeguarded by deposit insurance and regulatory supervision. For stablecoins, the issuer may find it desirable to fix the exchange rate through open market operations but does not face legal consequences for debasement. In the presence of shocks to reserves, such as fluctuations of collateral value and operational risk (e.g., cyber-attacks), commitment to redemption at par is only credible when reserves are sufficiently above the redemption value of outstanding stablecoins. Below this critical threshold, the issuer optimally debases the stablecoin.

In the debasement region, the redemption value comoves with reserves, effectively allowing the issuer to share risk with users and thereby to avoid liquidation. The issuer can also avoid liquidation by raising funds through the issuance of equity shares that are backed by future transaction fees charged to users (called "secondary units" among practitioners). However, as long as the financing costs exist, debasement happens before recapitalization. Moreover, we show that equity issuance must be accompanied by a jump in the stablecoin supply and a downward re-pegging; otherwise, arbitrage opportunities exist and the stablecoin issuer leaves money on the table.

The system exhibits a bimodal distribution of states. In states of high reserves, the stablecoin issuer honors the fixed exchange rate, so stablecoin demand is strong and transaction volume is high. Through open market operations and transaction fees, the issuer collects revenues that further grow its reserves. The level of reserves is capped by an endogenous upper bound, beyond which the stablecoin issuer pays out dividends, or, equivalently, repurchase equity shares. In states of low reserves, the issuer has to off-load risk to users. The

[1] This Chapter is based on Li and Mayer (2020).

[2] An alternative is to use algorithmic supply rules to stabilize price. Success has been limited in this area.

depressed stablecoin demand and transaction volume imply low revenues and slow accumulation of reserves. Therefore, a fixed exchange rate can last for a long time without any hint of instability, but once debasement happens, recovery is slow.

Stablecoins became the subject of heated debate after the technology giant Facebook and its partners announced their own stablecoin, Libra (now "Diem"), in June 2019.[3] Leveraging on their existing customer networks, global technology or financial firms are able to rapidly scale the reach of their stablecoins.[4] In our model, stronger network effects make a fixed exchange rate more sustainable and allow the stablecoin to recover more quickly from debasement. Stablecoin initiatives sponsored by companies with global customer networks attract attention from regulators for not only its potential of wide adoption but also concerns over monopoly power. In our model, stronger network effects allow the stablecoin issuer to earn more transaction fees and revenues from open market operations. However, because individual users do not internalize the positive network externalities, the stablecoin issuer has incentives to stimulate users' transactions by lowering fees and stabilizing the redemption value of the stablecoin. As a result, the split of welfare between the stablecoin issuer and users is rather insensitive to the degree degree of network effects.

The enormous transaction data brought by stablecoin systems offers a strong incentive for digital platforms to venture into payment services. Data enables new revenue sources, for example, advertisement targeting. In an extension, we model data as a productive asset that accumulates through users' transactions. A stablecoin platform faces a trade-off between reserve management and data acquisition. To preserve and grow reserves, the platform relies on transaction fees and, possibly, open market operations that off-load shocks to users (i.e., debasement). To acquire data, the platform needs to raise transaction volume by reducing fees (or even offering subsidies to users) and maintaining a stable redemption value of stablecoins. The optimal strategies then depend on the ratio of marginal value of reserves to marginal value of data (the "data q").

Finally, we evaluate two types of stablecoin regulations. The first is a standard capital requirement that stipulates the minimal degree of over-collateralization (equity). The capital requirement fails to eliminate debasement. As long as the threat of liquidation (or equity issuance costs) exists, whether it is due to reserve depletion or the violation of regulation, it is optimal for the stablecoin issuer and users to share risk through debasement. Therefore, the second type of regulation that forces a fixed exchange rate only hurts welfare by destroying the economic surplus from risk sharing.

Next, we provide more details on the model setup and mechanisms. The model is built to be technology-neutral so that it applies to stablecoins issued by central entities or on distributed ledgers.[5] In a continuous-time economy, a digital platform issues stablecoins ("tokens"). Users derive a flow utility from token holdings, which captures the transactional benefits, and network effect is modelled by embedding the aggregate holdings in individuals' utility (Cong, Li, and Wang, 2020c).[6] Users can redeem token holdings for numeraire goods ("dollars"). Redemption value can be continuously adjusted by the platform. The lack of commitment to a fixed exchange rate requires redemption value at any time to be optimal for the platform. Users are averse to the fluctuation of redemption value, so a volatile token value dampens token demand and transaction volume.

On the platform's balance sheet, the liability side has tokens and equity. On the asset side, it holds dollar reserves that earn a constant interest rate and load on Brownian shocks.

[3]The announcement triggered a globally-coordinated response under the umbrella of the G7. From then on, the G20, the Financial Stability Board (FSB), and central banks around the world have also embarked on efforts to address the potential risks while harnessing the potential of technological innovation.

[4]Another example is JPM Coin, a blockchain-based digital coin for fast payment settlement that is being developed by JP Morgan Chase and was announced in February 2019.

[5]At the current stage of stablecoin developments, policy makers take a technology-neutral approach that emphasizes economic insights over technological aspects of implementation ECB Crypto-Assets Task Force (2019).

[6]This money-in-utility approach follows the macroeconomics literature (Ljungqvist and Sargent, 2004). The modelling of network effect is in the traditional of social interaction (Glaeser, Sacerdote, and Scheinkman, 1996).

The shocks capture operational risk and unexpected fluctuation of reserve value. As standard in models of dynamic liquidity management (Bolton, Chen, and Wang, 2011; Décamps, Mariotti, Rochet, and Villeneuve, 2011; Hugonnier, Malamud, and Morellec, 2015b), shareholders' discount rate is above the interest rate of reserves, so the platform pays out reserves to shareholders when it has accumulated a sufficient amount as risk buffer. Beyond the interests on reserves, the platform earn fees charged to users and token-issuance proceeds. We allow open market operations in both directions, so the platform can also reduce token supply by selling dollars and buying back (and burn) tokens.

Through open market operations, the platform can implement any process of dollar price of token and preclude arbitrage between the secondary market and token redemption/issuance. Users are free to deposit dollars with the platform in exchange for tokens or redeem tokens for dollars under rational expectation of token price dynamics. Besides the token price process, the platform can also choose any processes of fees charged to users and payouts to shareholders.

For the platform's dynamic optimization, the amount of excess reserves (equity) is the state variable. The recursive formulation through Hamilton-Jacobi-Bellman (HJB) equation significantly simplifies the problem. First, the optimality condition on payouts to shareholders implies an endogenous upper bound on the state variable (the payout boundary). The natural lower bound is zero, the bankruptcy threshold, but is never reached because the platform can debase its token liabilities and will optimally do so to avoid costly and irreversible liquidation.

Then the choice of token price process boils down to choosing instantaneous drift and diffusion. Beyond the transactional benefits, users only care about the *net* appreciation or depreciation, i.e., drift minus fees that are proportional to token holdings, when choosing token demand. Therefore, the impact of drift and fees cannot be separately identified. The optimal set of fees, token-price drift and diffusion can thus be implemented through directly setting the aggregate token demand and choice of diffusion. Recovering the implied process of token redemption value can be formulated as a differential equation problem with the optimal token demand and token-price diffusion as inputs. The recovery delivers the optimal drift, which then implies optimal fees given the token demand.

In spite of positive equity (over-collateralization), the stablecoin issuer cannot always credibly promise redemption at par. To avoid costly liquidation, the platform opts for debasement whenever equity falls below a threshold. Debasement triggers a vicious cycle as the depressed token demand leads to a reduction in fee revenues, which causes a slow recovery of equity and persistent debasement. However, debasement is a necessary evil and a valuable option, as it allows the platform to share risk with users. When negative shocks decrease equity, debasement causes token liabilities to shrink. Above the debasement threshold, the platform credibly institutes token redemption at par. Then a strong token demand allows the platform to collect revenues to grow equity, which further strengthens the one-to-one convertibility to dollar. The virtuous cycle implies persistent expansion of platform equity until it reaches the payout boundary. The stationary distribution of platform equity is thus bimodal with two peaks near zero and the payout boundary, respectively.

Next, we allow the platform to raise equity. Under a fixed cost of equity issuance, the platform first resorts to debasement when equity falls below the threshold and only issues equity when equity falls to zero and when equity issuance leads to a higher shareholders' value than further debasement does. Once the fixed cost is paid, the platform raises equity all the way up to the payout boundary, where the marginal value of equity eventually falls to one. The jump in platform equity implies an immediate restoration of fixed exchange rate and, accordingly, a jump in the aggregate token demand. To preclude a predictable jump in token price (a secondary-market arbitrage opportunity), the platform must expand token supply at the payout boundary with the proceeds distributed to shareholders, an operation akin to issuing debts for share repurchase. Then the exchange rate is re-pegged. Therefore, the exchange rate starts at one, and every time after recapitalization (which follows debasement), the exchange rate is re-pegged downward to the pre-issuance level.

We further extend our model to incorporate data as a productive asset for the platform. Data improves the quality of the platform and thereby increases the users' flow utility from token holdings (the transactional benefits). Data accumulates through transactions. Under a constant money velocity, i.e., a constant ratio of transaction volume to users' token holdings, a feedback loop emerges — transactions generate more data, which improves the platform and leads to a stronger token demand and even more transactions. As a result, data accumulates exponentially over time. The platform has to balance between acquiring data and preserving reserves. The former requires lower fees and a more stable token while the latter calls for higher fees and risk-sharing with users through debasement. A key result is that the amount of reserves is no longer the key state variable driving the platform's decisions. The state variable is now the ratio of reserves to data stock. Data enters into the platform's decisions through a sufficient statistic, the data q, which is the marginal contribution of data to shareholders' value in analogy to Tobin's q of productive capital.

The model allows us to conduct several experiments that shed light on the heated debates surrounding stablecoins. An increase of data productivity captures the revolutionary progress in big data technology.[7] In response, the platform maintains more reserves before payout because data allows shareholders' investment to grow faster in expectation. However, more reserved do not lead to stabler tokens. The platform becomes more aggressive in stimulating transactions for data acquisition. A larger transaction volume amplifies operational risk, i.e., the shock loading of reserves. As a result, debasement is more likely to happen. Therefore, a paradox exists — stablecoins built primarily for the acquisition and utilization of transaction data can become increasingly unstable precisely when data becomes more valuable.

To understand the advantages of well-established digital networks in the stablecoin space, we compare platforms with different degrees of network effects. A stronger network effect is indeed associated with stabler tokens. The frequency of debasement declines because a stronger network effect allows the platform to accumulate fee revenues faster when equity runs out and, through a higher franchise (continuation) value, incentivizes the platform to build up a large equity position. As previously discussed, two counteracting forces limit the share of economic surplus seized by the platform. Under a stronger network effect, the platform can extract more rents from its users through fees or risk sharing, but it is also more eager to stimulate token demand by lowering fees and stabilizing tokens given that individual users do not internalize the positive network externalities.

As in our model, stablecoins are typically over-collateralized in practice (Bullmann, Klemm, and Pinna, 2019).[8] The issuer optimally maintains a risk buffer so that it does not need to immediately debase tokens following negative shocks. Even under voluntary over-collateralization, capital requirement can still add value. To avoid violating the regulation, the issuer accumulates more reserves that earn an interest rate below shareholders' discount rate, so shareholders' value declines. However, users' welfare increases because, even though debasement still happens, its frequency declines. The increase of users' welfare dominates the decrease of platform shareholders' value as the capital requirement rises up until an optimum level, which maximizes the total welfare. The optimal capital requirement increases in the strength of network effects, which amplifies the positive response of users to the regulation. In contrast, when data becomes more productive, the optimal capital requirement declines. By favoring reserve preservation over data acquisition, capital requirement stems the exponential growth of data that drives the long-run trajectory of welfare.

The current debate surrounding stablecoins largely adopts the "same business, same risk, same rules" principle (Financial Stability Board, 2020). Therefore, the conceptual framework aims to adopt banking regulations and analyzes stablecoin-enabled payment platforms from the perspective of systemically important financial institutions. Our model, albeit partial-equilibrium in nature, reveals unique features that distinguish stablecoins from traditional securities. Due to the debasement option, stablecoins share similarities with contingent convertible bonds (CoCos) that automatically share risk between equity owners and

[7] Alternative payment-service providers also benefit from regulatory initiatives that facilitate data sharing. A new European Union directive, PSD2, requires banks to provide non-bank service providers with data that would allow those providers to offer payment and other services to the banks' customers (Duffie, 2019).

[8] The concern over bank run and under-collateralization is partly justified by the inadequate auditing of reserves of certain stablecoin initiatives (Calle and Zalles, 2019).

debt (stablecoin) holders. Unlike CoCos, stablecoins do not pre-specify conversion rates and trigger events, which are both under discretion of the issuer.

Policy initiatives that aim to force a fixed exchange rate and regulate stablecoins as deposits ignore the ex-post efficiency of debasement option. Admittedly, our model omits several elements and thus can underestimate the value of a perfectly stable token. For example, debasement invites speculation that in turn amplifies price fluctuation and triggers a vicious cycle (Mayer, 2020b). To guard against speculation, stablecoin issuers often require users to post collateral themselves as the first line of defense before the issuers taps into their own reserves. Future research may explore the efficacy of double-deck scheme in stabilizing exchange rates in the presence speculative trading.

6.2 A dynamic model of stablecoins

Consider a continuous-time economy where a continuum of agents ("users") of unit measure conduct peer-to-peer transactions on a digital platform. The platform facilitates transactions by introducing a local currency ("token"). The generic consumption goods ("dollars") are the numeraire in this economy. The platform sets the exchange rate between tokens and dollars. Let P_t denote the token price in units of dollars (i.e., the exchange rate between tokens and dollars). At time t, users can redeem their token holdings for dollars or buy more tokens from the platform at the dollar price P_t. By no arbitrage, users also trade tokens among themselves at the dollar price P_t. The platform manages a stock of dollar reserves to meet potential token redemption and/or to manage token supply. Next, we first introduce users and then set up the platform's problem.

Users. We use $u_{i,t}$ to denote the numeraire (dollar) value of user i's holdings of tokens, so user i holds $k_{i,t} = u_{i,t}/P_t$ units of tokens. The aggregate value of token holdings is $N_t \equiv \int_{i \in [0,1]} u_{i,t} dt$.

A representative user i derives the following utility from token holdings

$$\frac{1}{\beta} N_t^{\alpha} u_{it}^{\beta} A^{(1-\alpha-\beta)} dt \,, \tag{6.1}$$

where $\alpha, \beta \in (0, 1)$ with $\alpha + \beta < 1$ and $A > 0$. We model the utility from holding means of payment following the classic models of monetary economics (e.g., Baumol, 1952; Tobin, 1956; Feenstra, 1986; Freeman and Kydland, 2000) and related empirical studies (e.g., Poterba and Rotemberg, 1986; Lucas and Nicolini, 2015; Nagel, 2016). In this literature, agents derive utility from the real value of holdings, i.e., $u_{i,t}$.[9] Following Rochet and Tirole (2003), we introduce network effect via N_t^{α}. As in Cong, Li, and Wang (2020c), it captures the fact that when tokens are more widely used as means of payment, each individual user's utility from using tokens is higher.[10] The quality of payment system is captured by parameter A which we will endogenize in Section 6.5.

User i pays a proportional fee on her token holdings, $u_{i,t} f_t dt$, where f_t is set by the platform. In practice, fees are often charged on transactions, in that f_t can be interpreted as the transaction fee per dollar transaction. Note that as long as the money (token) velocity is constant within a small time interval (dt), transaction volume is proportional to token holdings. There exists a technical upper bound on the volume of transactions that the platform can handle per unit of time. Without loss of generality, we model the bound as follows

$$N_t \leq \overline{N} \,. \tag{6.2}$$

[9] We refer readers to the textbook treatments (e.g., Galí, 2015; Ljungqvist and Sargent, 2004; Walsh, 2003). For the nominal value of means of payment (i.e., $k_{i,t}$) to affects agents' decisions, additional frictions, such as nominal illusion (e.g., Shafir, Diamond, and Tversky, 1997) or sticky prices (e.g., Christiano, Eichenbaum, and Evans, 2005), have to be introduced.

[10] For instance, when there are more people use tokens, it becomes easier to find a transaction counterparty that accepts tokens, so token holders expect more token usage of means of payment.

In sum, we define tokens' transaction utility from an ex ante perspective and (6.1) can be viewed as the utility from *expected* transaction needs in dt. We do not model the ex post circulation of tokens following the aforementioned literature on money-in-utility and cash-in-advance constraint.

The Platform. Let S_t denote the total units of tokens outstanding. The token market clearing condition is given by $S_t = \int_{i\in[0,1]} k_{i,t}dt$, or equivalently, in the numeraire (dollar) value:

$$N_t = S_t P_t. \tag{6.3}$$

The platform decides on the fees and controls the dollar price of tokens, P_t, by adjusting the token supply. This is akin to central banks using open market operations to intervene in the foreign exchange markets (e.g., Calvo and Reinhart, 2002). When the platform issues more tokens ($dS_t > 0$), it collects dollar revenues as users buy tokens with dollars. When the platform retires tokens ($dS_t < 0$), it loses dollars to users. Stablecoin platforms often claim a fixed exchange rate. However, we will show that optimal exchange rate depends on the platform's reserves.

Let M_t denote the platform's reserves (dollar holdings), which has a law of motion

$$dM_t = rM_t dt + (P_t + dP_t)dS_t + N_t f_t dt + N_t \sigma dZ_t - dDiv_t. \tag{6.4}$$

The first term is the interests earned on the reserves balance, and r is the constant interest rate. The second term is the revenues (losses) from issuing (buying back) tokens in dt from the secondary market. As the trade settles in dt, the amount of quantity adjustment, dS_t, is multiplied by the price in dt, $P_t + dP_t$. The third term is the fee revenues. The fourth term deserves more attention. Z_t is a standard Brownian motion, and its increment, dZ_t, captures the shocks to the net revenues, which can stem from operating expenses, risks involved in liquidity management, and activities beyond fees and token management.[11] This shock is the only source of uncertainty. Let Div_t denote the *cumulative* dividend process. The platform's reserves decrease when the platform pays its owners dividends, $dDiv_t$, which is non-negative under limited liabilities.

The platform maximizes the expected discounted value of dividend payouts to its owners:

$$V_0 \equiv \max_{\{f_t, dS_t, dDiv_t\}} \mathbb{E}\left[\int_0^\infty e^{-\rho t} dDiv_t\right] \quad \text{subject to} \quad (6.4) \text{ and } \quad dDiv_t \geq 0. \tag{6.5}$$

We assume that the platform's shareholders are impatient relative to other investors, $\rho > r$.[12]

Stability Preference. In equilibrium, the dollar price of token has a law of motion

$$\frac{dP_t}{P_t} = \mu_t^P dt + \sigma_t^P dZ_t\,, \tag{6.6}$$

which the atomic users take as given. In the next section, we will show how μ_t^P and σ_t^P depend on the platform's strategy. Let $R_{i,t}$ denote user i's (undiscounted) *cumulative* payoff from platform activities. The instantaneous payoff depends on user i's choice of $u_{i,t}$ and is given by

$$dR_{it} \equiv \frac{1}{\beta} N_t^\alpha u_{it}^\beta A^{(1-\alpha-\beta)} dt + u_{it}\left(\frac{dP_t}{P_t} - rdt - f_t dt\right), \tag{6.7}$$

[11] Examples include the revenues and costs associated with users' advertisement on the platform and loans to users, which are often enabled by the platform's possession of transaction data as will be discussed in Section 6.5. The risk of reserve portfolio may also arise from the platform holding different currencies and the fluctuation of these currencies' dollar exchange rates. Providers of global stablecoins (GSC) typically accept deposits in different currencies, and hold a portfolio of these currencies as backing (e.g., Libra Association, 2020).

[12] This impatience could be preference based or could arise indirectly because shareholders have other attractive investment opportunities. From a modeling perspective, impatience motivates the platform to pay out because, otherwise, the expected return on M_t is greater than r (due to revenues from the fees and token issuance), which then implies that the platform never pays out dividends.

where the first term is the flow utility (6.1) and the second term includes the return from token price change net off the forgone interests and fees.

A representative user i chooses $u_{i,t} \geq 0$ to maximize

$$\max_{u_{i,t}\geq 0} \mathbb{E}_t\left[dR_{it}\right] - \eta|\sigma_t^P|u_{i,t}dt = \max_{u_{i,t}} \frac{1}{\beta} N_t^{\alpha} u_{it}^{\beta} A^{(1-\alpha-\beta)}dt + u_{it}\left(\mu_t^P - \eta|\sigma_t^P| - r - f_t\right)dt\,. \quad (6.8)$$

First note that the classic Inada condition rules out $u_{i,t} < 0$. The user's aversion to token volatility is captured by the parameter η (> 0), and is defined on the absolute value of σ_t^P to capture the fact that users are averse to token price fluctuation no matter whether the price moves with ($\sigma_t^P > 0$) or against ($\sigma_t^P < 0$) the platform's dollar revenue shock dZ_t. The disutility from token volatility can be motivated by risk-averse preference or users' aversion to exchange-rate shocks that cause losses of net worth when assets and liabilities are denominated in different currencies (tokens and dollars) (Doepke and Schneider, 2017; Gopinath, Boz, Casas, Díez, Gourinchas, and Plagborg-Møller, 2020).

Liquidation. The platform liquidates if it can no longer guarantee the fulfillment of users' potential redemption en masse, i.e., $M_t < S_t P_t$.[13] In liquidation, token holders receive the dollar reserves pro rata and the platform owners obtain zero value. This criterion can be a legally binding commitment. It can also be motivated by the classic run mechanism (Diamond and Dybvig, 1983). As users redeem tokens on a first-come, first-served basis, a coordination failure can happen if $M_t < S_t P_t$ – an individual user i optimally chooses to redeem all of her token holdings when the rest of users do so. A run on stablecoins is a prominent concern of policy makers (Brainard, 2019; G7 Working Group on Stablecoins, 2019; ECB Crypto-Assets Task Force, 2019). A run does not necessarily happen when M_t falls right below $S_t P_t$. We may allow reserves to run down below $S_t P_t$ to a certain extent. As noted by Calle and Zalles (2019) and Duffie (2019), there has not been standard practice of reserve disclosure and auditing. Stablecoin users are likely to receive noisy signals about M_t. In a global-game framework, the run threshold can then be uniquely determined (Morris and Shin, 1998; Goldstein and Pauzner, 2005). As long as there exists a threshold of reserves below which liquidation happens, our analysis below carries through.

6.3 Equilibrium

In this section, we characterize the analytical properties of the dynamic equilibrium and, to sharpen the economic intuition, we also provide graphical illustrations based on the numerical solutions.

6.3.1 Managing the stablecoin platform

User Optimization. A representative user i solves a static problem in (6.8) and the solution is

$$u_{it} = \left(\frac{N_t^{\alpha} A^{(1-\alpha-\beta)}}{r + f_t - \mu_t^P + \eta|\sigma_t^P|}\right)^{\frac{1}{1-\beta}}. \quad (6.9)$$

Users' choices exhibit strategic complementarity as $u_{i,t}$ increases in the aggregate value N_t. In equilibrium $N_t = u_{i,t}$ under user homogeneity, which, through (6.9), implies

$$N_t = \frac{A}{\left(r + f_t - \mu_t^P + \eta|\sigma_t^P|\right)^{\frac{1}{1-\xi}}}, \quad (6.10)$$

where, to simplify the notations, we define $\xi \equiv \alpha + \beta$ (< 1). Aggregate token demand decreases in the fees charged by the platform, f_t, and depends on the token price dynamics,

[13] To avoid liquidation, the platform can potentially reset the redemption price P_t, but to make up the deficit, $S_t P_t - M_t$, the token price has to jump, causing users to expect an infinite rate of change and thus refrain from holding any tokens under their preference for token price stability.

which the platform controls. The capacity constraint (6.2) on N_t has to be satisfied when the platform sets its strategy.

Platform Optimization. To solve the platform's optimal strategy, we first note that, given the token price dynamics (i.e., μ_t^P and σ_t^P), the platform can directly set N_t through the fees f_t. Rearranging (6.10), we can back out the fees implied by the platform's choice of N_t:

$$f_t = \left(\frac{A}{N_t}\right)^{1-\xi} - r + \mu_t^P - \eta|\sigma_t^P|\,. \tag{6.11}$$

Using (6.11), we substitute out f_t in the law of motion of reserves (6.4) and obtain

$$dM_t - (P_t + dP_t)dS_t = rM_t dt + N_t^{\xi} A^{1-\xi} dt - rN_t dt + N_t\left(\mu_t^P - \eta|\sigma_t^P|\right)dt + N_t \sigma dZ_t - dDiv_t\,. \tag{6.12}$$

Next, we show the state variable for the platform's dynamic optimization is the *excess reserves*,

$$C_t \equiv M_t - P_t S_t\,. \tag{6.13}$$

To derive the law of motion of C_t, we first note that

$$\begin{aligned} dC_t &= dM_t - d(S_t P_t) = dM_t - (P_t + dP_t)dS_t - S_t dP_t \\ &= dM_t - (P_t + dP_t)dS_t - N_t\left(\mu_t^P dt + \sigma_t^P dZ_t\right)\,. \end{aligned} \tag{6.14}$$

The second equality uses $d(S_t P_t) = dS_t P_t + S_t dP_t + dS_t dP_t$ (by Itô's lemma) and the last equality uses (6.6) and $N_t = S_t P_t$. From a balance-sheet perspective, the reserves, M_t, are the platform's assets and the outstanding tokens, $P_t S_t$, are the liabilities. The excess reserves constitute the equity. Thus, equation (6.14) is essentially the differential form of balance-sheet identity. Using (6.12), we substitute out $dM_t - P_t dS_t$ the right side of (6.14) and obtain the following law of motion of C_t:

$$dC_t = \left(rC_t + N_t^{\xi} A^{1-\xi} - N_t \eta|\sigma_t^P|\right)dt + N_t(\sigma - \sigma_t^P)dZ_t - dDiv_t\,. \tag{6.15}$$

Note that $N_t \mu_t^P$ disappears. As shown in (6.12), the platform receives more fee revenues (see (6.11)) when users expect tokens to appreciate ($N_t \mu_t^P$), but such revenues do not increase the platform's equity (excess reserves) as they are cancelled out by the appreciation of token liabilities. Thus, the drift term, $rC_t + N_t^{\xi} A^{1-\xi} - N_t \eta|\sigma_t^P$, is the expected appreciation of the platform's equity position.

We characterize a Markov equilibrium with the platform's excess reserves, C_t, as the state variable. In the following, we solve the platform's control variables, $dDiv_t$, σ_t^P, and N_t, as functions of C_t, and thereby, show that (6.15) is an autonomous law of motion of the state variable.

The platform owners' value function at time t is given by

$$V_t = V(C_t) = \max_{\{N, \sigma^P, Div\}} \mathbb{E}\left[\int_{s=t}^{\infty} e^{-\rho(s-t)} dDiv_s\right]\,. \tag{6.16}$$

The platform pays dividends when the marginal value of excess reserves is equal to one, i.e., one dollar has the same value either held within the platform or paid out,

$$V'\left(\overline{C}\right) = 1\,. \tag{6.17}$$

The next proposition states that the value function is concave. The declining marginal value of excess reserves implies that $\overline{C}$ in (6.17) is an endogenous upper bound of the state variable C_t. At any $C_t \in \left(0, \overline{C}\right)$, the platform does not pay dividends to its owners because the marginal value of excess reserve, $V'\left(\overline{C}\right)$, is greater than one, i.e., the owners' value of dividend. And the optimality of payout at $\overline{C}$ also requires the following super-contact condition

(Dumas, 1991b),

$$V''\left(\overline{C}\right) = 0\,. \tag{6.18}$$

Proposition 24 (Value Function Concavity and Over-Collateralization) *There exists $\overline{C} > 0$ such that $C_t \in \left(0, \overline{C}\right]$. For $C_t \in \left(0, \overline{C}\right)$, the value function is strictly concave, and $V'\left(C_t\right) > 1$ so the platform maintains excess reserves. At $C_t = \overline{C}$, $V'\left(\overline{C}\right) = 1$ and the platform pays dividends.*

The platform's token liabilities are over collateralized. The intuition can be understood through the wedge between the liquidation value (zero) and the strictly positive value of platform as an ongoing concern. At $C_t = 0$, even a tiny negative shock triggers a downward jump of platform value to zero. By holding excess reserves, even by a small amount, the platform can prevent this. Next we confirm that a platform as ongoing concern always has a positive value. In the interior region $C \in \left(0, \overline{C}\right)$, $dDiv_t = 0$ and we obtain the following Hamilton-Jacobi-Bellman (HJB) equation:

$$\rho V(C) = \max_{\{N \in [0, \overline{N}], \sigma^P\}} \left\{ V'(C) \left(rC + N^{\xi} A^{1-\xi} - \eta N |\sigma_t^P| \right) + \frac{1}{2} V''(C) N^2 (\sigma - \sigma^P)^2 \right\}, \tag{6.19}$$

Setting first $\sigma^P = \sigma$ and then $N = 0$ is feasible in the HJB equation, which implies

$$V(C) \geq \frac{V'(C)}{\rho} \left(rC + \max_{\{N \in [0, \overline{N}]\}} \left\{ N^{\xi} A^{1-\xi} - \eta N \sigma \right\} \right) \geq \frac{V'(C) rC}{\rho} > 0\,. \tag{6.20}$$

Next, we analyze the platform's optimal choice of token exchange-rate process and transaction volume. Because μ_t^P disappears from (6.15), the platform's choice of redemption price, P_t (dollar exchange rate), boils down to the choice of $\sigma_t^P = \sigma^P\left(C_t\right)$. First, we consider how the choice of σ_t^P and N_t when C_t approaches zero. In the limit, σ_t^P must converge to σ to mute the shock exposure,

$$\lim_{C \to 0^+} \sigma^P(C) = \sigma\,; \tag{6.21}$$

otherwise, dC_t's loading on dZ_t in (6.15) is positive and a negative shock can trigger liquidation. Equation (6.21) implies that, when taking the right-limit on both sides of (6.19), we obtain

$$\lim_{C \to 0^+} \frac{V(C)}{V'(C)} = \frac{1}{\rho} \max_{\{N \in [0, \overline{N}]\}} \left\{ N^{\xi} A^{1-\xi} - \eta N \sigma \right\} = \frac{A}{\rho} \left(\frac{\xi}{\eta\sigma} \right)^{\frac{1}{1-\xi}} \left(\frac{1-\xi}{\xi} \right) \eta\sigma\,, \tag{6.22}$$

where the second equality follows from plugging in the optimal N_t given by

$$\underline{N} \equiv \lim_{C \to 0^+} N(C) = \arg\max_{N \in [0, \overline{N}]} \left\{ N^{\xi} A^{1-\xi} - \eta N \sigma \right\} = A \left(\frac{\xi}{\eta\sigma} \right)^{\frac{1}{1-\xi}}\,. \tag{6.23}$$

Therefore, (6.21) and (6.23) characterize respectively the limiting behavior of σ_t^P and N_t as C_t approaches zero. In the process, we find the value of $\lim_{C \to 0^+} V(C)$, a boundary condition for the HJB equation. As an interim summary, the next proposition summarizes the value function solution as solution to an ordinary differential equation (ODE) problem with an endogenous boundary. Figure 6.1 plots the numerical solution of value function (Panel A) and the decreasing marginal value of excess reserves with the red dotted line marking the payout boundary $\overline{C}$.

Proposition 25 (Solving Value Function) *The value function, $V(C)$, and the boundary $\overline{C}$ are solved by the ordinary differential equation (6.19) under the boundary conditions (6.17), (6.18), and (6.22).*

Next, we fully characterize the platform's optimal choices of σ_t^P and N_t as functions of the state variable, C_t (via the derivatives of $V(C)$). First, we define the platform's *effective risk aversion*:

$$\gamma\left(C\right) \equiv -\frac{V''(C)}{V'(C)}\,. \tag{6.24}$$

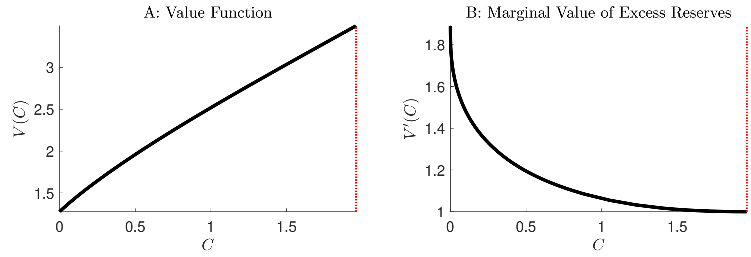

FIGURE 6.1: **Value Function.** This figure illustrates the level and first derivative of the platform's value function. The red dotted lines in both panels mark $\overline{C}$ (defined in Proposition 24). The parameters are $r = 0.05$, $\rho = 0.06$, $\sigma = 0.1$, $\overline{N} = 5$, $\eta = 0.15$, $\alpha = 0.45$, $\beta = 0.05$, and $A = 0.0025$.

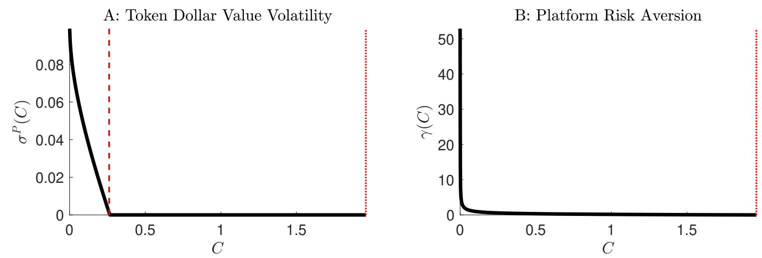

FIGURE 6.2: **Token Volatility and Platform Risk Aversion.** This figure plots token return volatility $\sigma^P(P)$ in Panel A and the platform's risk aversion $\gamma(C)$ in Panel B. In Panel A, the red dashed line marks $\widetilde{C}$ (in Proposition 26). The red dotted lines in both panels mark $\overline{C}$ (in Proposition 24). The parameterization follows Figure 6.1.

This definition is analogous to the classic measure of absolute risk aversion of consumers (Arrow, 1965; Pratt, 1964). From Proposition 24, $\gamma(C) \geq 0$ and, in $(0, \overline{C})$, $\gamma(C) > 0$. The *endogenous* risk aversion arises from the concavity of value function, which is in turn due to the gap between liquidation value and continuation value as previously discussed. The next proposition states the monotonicity of $\gamma(C)$ in C and summarizes the optimal $\sigma_t^P = \sigma^P(C_t)$ and $N = N(C_t)$.

Proposition 26 (Risk Aversion, Token Volatility, and Transaction Volume) *The platform's effective risk aversion, $\gamma(C)$, strictly decreases in the level of reserve holdings, C. There exists $\widetilde{C} \in \left(0, \overline{C}\right)$ such that, at $C \in \left(0, \widetilde{C}\right)$, $N(C) = \underline{N}$ and $\sigma^P(C)$ strictly decreases in C, given by,*

$$\sigma^P(C) = \sigma - \frac{\eta}{\gamma(C)\underline{N}} \in (0, \sigma), \tag{6.25}$$

and at $C \in \left[\widetilde{C}, \overline{C}\right]$, $\sigma^P(C) = 0$ and $N(C)$ increases in C, given by

$$N(C) = \min\left\{\left(\frac{\xi A^{1-\xi}}{\gamma(C)\sigma^2}\right)^{\frac{1}{2-\xi}}, \overline{N}\right\}. \tag{6.26}$$

When the platform's reserves are low, i.e., $C \in (0, \widetilde{C})$, it is the ratio of users' risk aversion to the platform's risk aversion that determines token volatility. Equation (6.25) shows that, in this region, when the platform accumulates more reserves and becomes less risk-averse, it

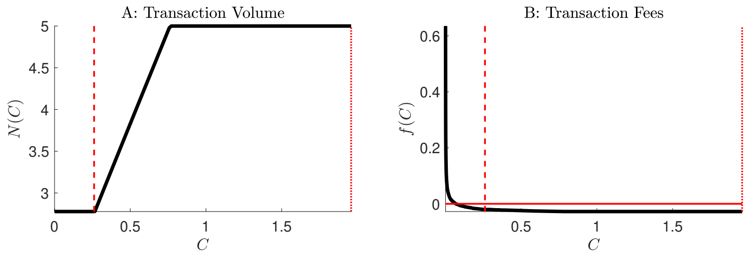

FIGURE 6.3: **Transaction Volume and Fees.** This figure plots transaction volume $N(C)$ in Panel A and fees per dollar of transaction $f(C)$ in Panel B. The red dotted lines mark $\overline{C}$ (in Proposition 24). In Panel A, the red dashed line marks $\widetilde{C}$ (in Proposition 26). In Panel B, the red solid line marks zero. The parameterization follows Figure 6.1.

absorbs risk from users by tuning down σ_t^P, and when the platform exhausts its reserves, it off-loads the risk in its dollar revenues to users.[14] The platform and its users engage actively in risk-sharing in $C \in (0, \widetilde{C})$. This is illustrated by the numerical solution in Panel A of figure 6.2 with $\widetilde{C}$ marked by the dashed line. In Panel B, we show that the platform's risk aversion declines in C. In this region of low reserves, the transaction volume, which is simply the dollar value of users' token holdings under our assumption of constant velocity, is pinned to the lowest level given by $\underline{N}$ in (6.23).

Once the platform's reserves surpass the critical threshold $\widetilde{C}$, its risk aversion becomes sufficiently low and it optimally absorbs all the risk in its dollar revenues, setting $\sigma^P(C)$ to zero which also implies that in this region $\mu^P(C) = 0$.[15] As a result, the transaction volume on the platform starts to rise above the "hibernation level", $\underline{N}$, as illustrated by Panel A of Figure 6.3. Therefore, reserves are absolutely essential for stimulating economic activities on a stablecoin platform.

Interestingly, even though the platform shelters its users from risk at any $C > \widetilde{C}$, its risk aversion still shows up in N_t given by (6.26). As shown in (6.11), the platform's choice of N_t is implemented through fees. Therefore, the intuition can be more easily explained when we substitute (6.26), the optimal N_t, and the optimal $\sigma_t^P = 0$ (as well as $\mu_t^P = 0$) into (6.11) to solve f_t: when $\left(\frac{\xi A^{1-\xi}}{\gamma(C)\sigma^2}\right)^{\frac{1}{2-\xi}} < \overline{N}$,

$$f_t = \left(\frac{A\gamma(C)\sigma^2}{\xi}\right)^{\frac{1-\xi}{2-\xi}} - r\,, \tag{6.27}$$

and when $\left(\frac{\xi A^{1-\xi}}{\gamma(C)\sigma^2}\right)^{\frac{1}{2-\xi}} \geq \overline{N}$, i.e., C is sufficiently high such that $\gamma(C)$ falls below $\frac{\xi A^{1-\xi}}{\sigma^2 \overline{N}^{2-\xi}}$,

$$f_t = \left(\frac{A}{\overline{N}}\right)^{1-\xi} - r\,. \tag{6.28}$$

As a platform accumulates reserves, its risk aversion declines, which, through (6.27), implies low fees charged on users and in turn a larger token demand N_t (in (6.26)).

[14] Equation (6.25) implies that the condition (6.22) is equivalent to $\gamma(C)$ (or $-V''(C)$) approaching infinity in the limit.

[15] This result arises because we express the equilibrium token price as a function of C, in that $P_t = P(C_t)$. Thus, token volatility and token returns can be expressed as functions of C too, in that $\sigma_t^P = \sigma^P(C_t)$ and $\mu_t^P = \mu^P(C_t)$. Since $\sigma^P(C) = 0$ for $C > \widetilde{C}$, token price $P(C)$ must be constant for $C > \widetilde{C}$, implying $\mu^P(C) = 0$ for $C > \widetilde{C}$.

The platform faces a risk-return trade-off. The fees serve as a compensation for risk exposure but discourages users from participation. So when the platform's risk aversion rises, it charges users more per dollar of transaction at the expense of a smaller volume. When the platform's risk aversion declines, the fees per dollar of transaction decline while the total transaction volume increases. Once reserves are sufficiently high such that $\gamma(C) \leq \frac{\xi A^{1-\xi}}{\sigma^2 \overline{N}^{2-\xi}}$, the fees no longer declines with the platform's risk aversion, as the platform has maxed out its transaction capacity, i.e., $N_t = \overline{N}$, and it becomes impossible to further stimulate user participation. Likewise, when the platform's reserves are below $\widetilde{C}$ and $\sigma^P(C) > 0$, $N_t = \underline{N}$, and the fees are given by

$$f_t = \left(\frac{A}{\underline{N}}\right)^{1-\xi} + \mu^P(C) - \eta\sigma^P(C) - r\,. \tag{6.29}$$

Even though the platform's risk aversion is high, it can no longer sacrifice transaction volume for higher fees because user participation already falls to the lowest level.

Panel B of Figure 6.3 plots the numerical solution of optimal fees that decrease in excess reserves. Depending on the parameters, fees can actually turn into user subsidies (i.e., fall below zero) when excess reserves are sufficiently high.[16] The next corollary summarizes the results on fees.

Corollary 11 (Optimal Transaction Fees) *Transaction fees, $f(C)$, decreases in excess reserves, C. At $C \in (0, \widetilde{C})$, where $\widetilde{C}$ is defined in Proposition 26, transaction fees are given by (6.29). At $C \in [\widetilde{C}, \widetilde{C}')$, where $\widetilde{C}'$ is defined by $\gamma(\widetilde{C}') = \frac{\xi A^{1-\xi}}{\sigma^2 \overline{N}^{2-\xi}}$, transaction fees are given by (6.27). At $C \in [\widetilde{C}', \overline{C})$, where $\overline{C}$ is defined in Proposition 24, transaction fees are given by (6.28).*

Another interesting implication of the optimal fees is that the platform charges (compensates) users the expected appreciation (depreciation) of tokens over risk-free rate, i.e., $\mu_t^P - r$ in f_t. To fully solve the fees, we need to know both $\gamma(C_t)$ and the function $\mu_t^P = \mu^P(C_t)$. In fact, the platform's choice of $\sigma_t^P = \sigma^P(C_t)$ already pins down the function of token price, $P_t = P(C_t)$, so the function $\mu^P(C_t)$ can be obtained from Itô's lemma. Next, we solve $P_t = P(C_t)$ from the function $\sigma^P(C_t)$. By Itô's lemma,

$$\sigma^P(C) = \frac{P'(C)}{P(C)} N(C) \left(\sigma - \sigma^P(C)\right), \tag{6.30}$$

where $N(C)\left(\sigma - \sigma^P(C)\right)$ is the diffusion of state variable C_t. Rearranging the equation, we solve

$$\frac{P'(C)}{P(C)} = \frac{1}{N(C)} \left(\frac{\sigma^P(C)}{\sigma - \sigma^P(C)}\right). \tag{6.31}$$

Using Proposition 25, we solve the value function $V(C)$ and obtain the function $\gamma(C)$. Then using Proposition 26, we obtain the functions $\sigma^P(C)$ and $N(C)$. Plugging $\sigma^P(C)$ and $N(C)$ into (6.31), we obtain a first-order ODE for the function of dollar price of token, $P(C)$.

To uniquely solve the function $P(C)$, we need to augment the ODE (6.31) with a boundary condition. In our model, both the platform and users are not concerned with the level of token price and only care about the expected token return, μ_t^P, and return volatility, σ_t^P. Therefore, we have the liberty to impose the following boundary condition:

$$P\left(\overline{C}\right) = 1\,. \tag{6.32}$$

i.e., the platform sets an exchange rate of one dollar for one token when C_t reaches $\overline{C}$. The next corollary states the solution of token price as solution to a first-order ODE problem.

[16] Specifically, under the particular parameterization, the condition is for fees to turn into subsidies near $\overline{C}$ is that $\frac{A^{1-\xi}}{\overline{N}^{1-\xi}} < r$ where we use (6.28) and the fact that $\mu^P(C) = 0$ for $C \in \left(\widetilde{C}, \overline{C}\right)$ (to be discussed later in this section).

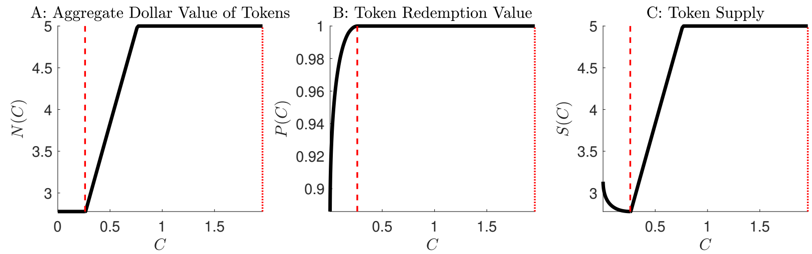

FIGURE 6.4: **Token Price and Quantity Dynamics.** This figure plots the aggregate dollar value of tokens $N(C)$ in Panel A, the token redemption value or dollar price $P(C)$ in Panel B, and token supply $S(C)$ in Panel C. The red dotted lines in all panels mark $\overline{C}$ (defined in Proposition 24). The red dashed lines in all panels mark $\widetilde{C}$ (defined in Proposition 26). The parameterization follows Figure 6.1.

Corollary 12 (Solving Token Price) *Given the solutions of $V(C)$ from Proposition 25 and $\sigma^P(C)$ and $N(C)$ from Proposition 26, the dollar price of token, $P(C)$, is solved by the ordinary differential equation (6.31) under the boundary condition (6.32).*

Proposition 26 states that, once C crosses above the critical threshold $\widetilde{C}$, $\sigma^P(C) = \mu^P(C) = 0$, which, by Itô's lemma (i.e., (6.30)), implies that $P'(C) = 0$. Therefore, if the platform's reserves are sufficiently high, it optimally fixes the dollar price (i.e., redemption value) of token at $P(C) = 1$. When C falls below $\widetilde{C}$, (6.30) implies that $P'(C) > 0$ (because $\sigma^P(C) \in (0, \sigma)$ in Proposition 26) so the token redemption value comoves with the platform's excess reserves.

The endogenous transition between redemption at par and redemption below par happens as the platform accumulates or depletes reserves through various activities laid out in (6.4) (and then (6.15)), including the platform's issuance of new tokens, users' token redemption, fee revenues, and shocks to the dollar reserve flow. The platform's choice of redemption price is optimally chosen and thus credible in the sense that the platform does not have incentive to deviate.

Proposition 27 (Credible Redemption Value Regimes) *At $C \in [\widetilde{C}, \overline{C}]$, where $\widetilde{C}$ is defined in Proposition 26, the platform credibly commits to redeem token at par, i.e., $P(C) = 1$. At $C \in (0, \widetilde{C})$, the redemption value of token comoves with the platform's excess reserves (i.e., $P'(C) > 0$).*

Proposition 27 states that redemption at par is credible if and only if the platform holds a sufficiently large amount of *excess* reserves ($C > \widetilde{C}$). When excess reserves fall below $\widetilde{C}$, the platform optimally debases its tokens. By allowing the redemption value to comove with its excess reserves, the platform off-loads the risk in its dollar revenues to users and thereby prevents liquidation.

Figure 6.4 plots the numerical solutions of aggregate token value, $N(C) = P(C)S(C)$ (Panel A), the redemption value of one token optimally set by the platform $P(C)$ (Panel B), and the total quantity of tokens $S(C)$ implied by $N(C)$ and $P(C)$ (Panel C). The dashed line marks $\widetilde{C}$. The platform implements the optimal token redemption value through the manipulation of token supply. When the platform has enough reserves to credibly sustain redemption at par (i.e., $C > \widetilde{C}$), token supply comoves with demand so that $P(C)$ is fixed. Below $\widetilde{C}$, a decrease of excess reserves triggers the platform to supply more tokens in exchange for dollars that replenishes reserves. The users respond to token debasement by reducing demand to $\underline{N}$, which in turn reinforces the debasement.

In practice, stablecoin platforms often claim commitment to redemption at par and substantiate this commitment by holding reserves that just cover their token liabilities. However, our analysis so far has drawn two conclusions that challenge such conventional wisdom. First, as long as there exists uncertainty in a platform's dollar revenues, over-collateralization

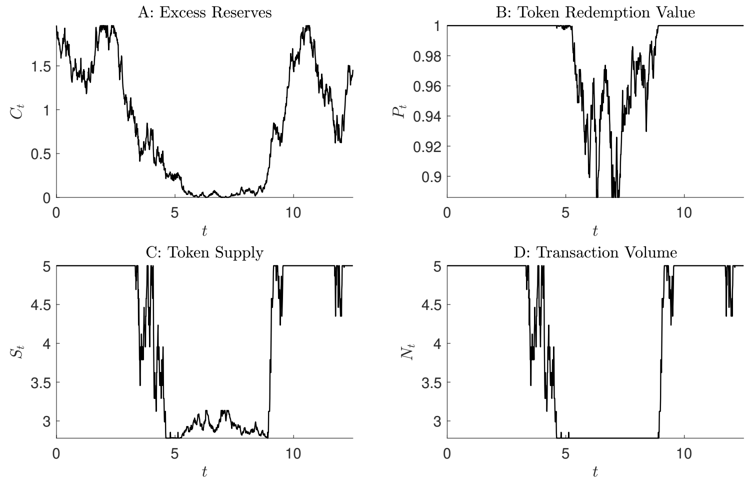

FIGURE 6.5: **Simulation.** Using the numerical solutions, we simulate a path of excess reserves (Panel A), token redemption value (Panel B), token supply (Panel C), and transaction volume on the platform (Panel D). The horizontal axis records the number of years. The parameterization follows Figure 6.1.

is not only necessary but also optimal from the platform's perspective. Second, redemption at par in every state of world is not credible or "incentive-compatible". As shown by Proposition 27, it is always in the platform's interest to debase its tokens once its excess reserves falls below the critical threshold. The credibility of commitment to redemption at par is thus contingent on the reserve level.

Simulation and Long-Run Dynamics. Using the parameters in Figure 6.1 and the numerical solutions, we simulate in figure 6.5 a path of excess reserves C_t (Panel A), token redemption value P_t (Panel B), token supply S_t (Panel C), and transaction volume N_t (Panel D). The horizontal axis records the number of years. In the first three years, in spite of the volatility in C_t, the platform manages to sustain redemption at par, and with the transaction volume (or token demand) at the full capacity at $\overline{N}$, a fixed dollar price of token implies a fixed token supply. Following a sequence of negative shocks between the third and fourth years, the platform raises fees. Users respond by reducing their token demand N_t, so the platform reduces token supply, maintaining redemption at par. The platform optimally trades off replenishing dollars reserves by raising fees and using dollar reserves in token buy-back. As more negative shocks hit between the fourth and ninth years, the platform gives up the peg and off-loads risk to users through the fluctuation of token redemption price. Users' token demand hits $\underline{N}$, and the platform starts actively expanding token supply in exchange for dollar revenues. Then following a sequence of positive shocks, the recovery started in the ninth year, and by the tenth year, the platform restores redemption at par.

We demonstrate the long-run dynamics of the model in Figure 6.6. Panel A plots the stationary probability density of excess reserves. It shows how much time over the long run the platform spends in different regions of C. The distribution is bimodal. The concentration of probability mass near $C = 0$ is due to the fact that, when the transaction volume (or token demand) gets stuck at the hibernation level $\underline{N}$, the platform can only grow out of this region very slowly by accumulating reserves through fee revenues and proceeds from expanding token supply. The platform also spends a lot of time near the payout boundary $\overline{C}$ as this is a stable region where, given a sufficiently high level of reserve buffer, shocks' impact is limited. In Panel B, we show that, even though redemption at par seems to be the norm, the system exhibits significant risk of token debasement ($P(C) < 1$).

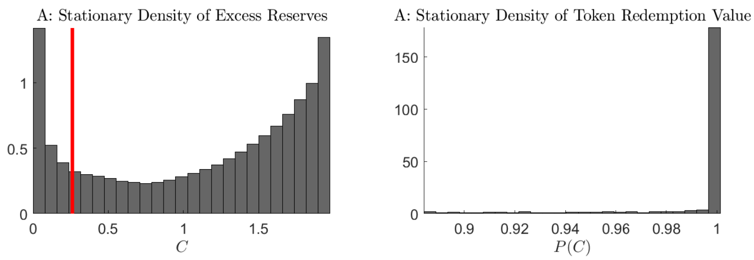

FIGURE 6.6: **Long-Run Dynamics and Stationary Density.** We plot stationary probability densities of excess reserves C (Panel A) and token value $P(C)$ (Panel B) in numerical solutions. The parameterization follows Figure 6.1.

Discussion: Liquidation. In our model, the platform never liquidates. As C declines and approaches zero, the platform gradually off-loads risk to users (see (6.21)). As a result, the platform avoids liquidation and gradually accumulates reserves through the interests on reserve holdings, the fee revenues, and, in case users' token demand increase, through the issuance of new tokens. As C increases either through the various sources of revenues or positive shocks, the platform can escape the low-C region where user participation is the lowest at the hibernation level $\underline{N}$.

In reality, stablecoin platforms can enter into liquidation, which is a major concern of practitioners and policy makers. Our model can be easily extended to make liquidation a positive probability event. The platform's dollar revenues can be subject to both small shocks (i.e., the diffusive, Brownian shock dZ_t) and large negative shocks that arrive following a Poisson process. When the level of reserves is sufficiently low, the arrival of a large negative shock triggers liquidation. Before the arrival of the large shock, the behavior of the extended model is akin to that of the current model except that the expectation of liquidation increases the concavity of value function and thereby induces more reserve holdings and risk-sharing between the platform and users.

6.3.2 The role of network effects

After Facebook announced its stablecoin project Libra (recently renamed to Diem), the interest in stablecoins among practitioners and regulators skyrocketed. Different from other stablecoin issuers, Facebook has the unique advantage of strong network effects brought by its comprehensive infrastructure covering social network, social media, and e-commerce (Facebook Shop). For individual users, the benefit of adopting Diem is enormous if other users on Facebook adopts Diem, because a great variety of activities can potentially be enabled by a universal means of payment for users around the globe. In our model, strong network effects are captured by a large value of α (see (6.1)).

Next, we compare stablecoin platforms with different degrees of network effects. Panel A of Figure 6.7 plots the payout boundary $\overline{C}$ as a measure of voluntary over-collateralization over different values of α. On the one hand, stronger network effects make the platform more profitable, which stimulates more precautionary savings to protect the franchise value. On the other hand, stronger network effects imply a higher level of user activities near $C = 0$, i.e., $\underline{N}$ in (6.23), which in turn implies a faster recovery out of the low-C region (through fee and token-issuance revenues) and thereby a weaker incentive to hold excess reserves. The two counteracting forces lead to the hump-shaped relationship between $\overline{C}$ and α in Panel A of Figure 6.7.

In Panel B of Figure 6.7, we show the long-run probability of $C > \widetilde{C}$ (based on stationary probability distribution of C) increases in α. Stronger network effects imply that, over the long run, the system spends more time in states with $P(C) = 1$. Under stronger network effects, recovery out of the low-C region is faster due to a higher level of user activities and

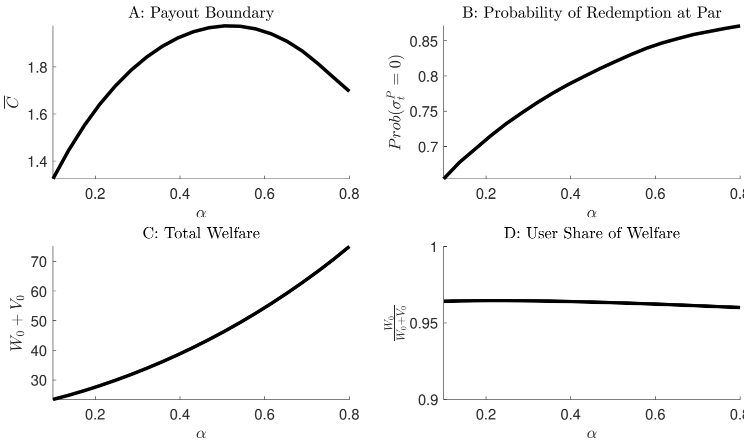

FIGURE 6.7: **Network Effects.** We plot the payout boundary (Panel A), the long-run probability of $C > \widetilde{C}$ based on stationary distribution (Panel B), the sum of platform value and users' welfare (Panel C), and users' share of total welfare (Panel D) over different values of α (degree of network externality). The parameterization follows Figure 6.1.

the resultant faster replenishment of reserves via fees and token-issuance proceeds. Our paper sheds light on why stablecoins issued by Facebook and other technology giants with strong network infrastructure are regarded as more promising than those issued by start-up payment service providers. Stronger network effects lead to stabler tokens and a lower likelihood of token debasement.

Finally, we examine the impact of network effects on welfare. In Panel C of Figure 6.7, we show that total welfare of the platform and its users increases in the degree of network effects. This explains why it is particular beneficial for technology giants with strong network infrastructure to introduce stablecoins as common means of payment among their customers. Network infrastructure is not limited to social network and e-commerce. Financial network is another example. JPMorgan Chase introduces JPM Coin to facilitate transactions among institutional clients.

Interestingly, as we gradually increase the degree of network effects in Panel D of Figure 6.7, the split of total welfare between the platform and its users is rather stable. Under stronger network effects, the monopolistic platform can extract more rents from its users through fees or off-loading risk in distress. However, precisely due to the network effects, individual users do not internalize the positive effect of their adoption on other users, so the platform has incentive to internalize the network externality by stimulating user activities through fee reductions (or subsidies) and token stability. These two counteracting forces imply that, as network effects become stronger, the platform's share of total surplus does not necessarily increase. This result alleviates the concern over technology giants abusing network effects in stablecoin projects.

6.3.3 Recapitalizing the stablecoin platform

At the end of this section, we take an excursion to analyze platform recapitalization. So far, the platform recovers from the low-C region through the accumulation of internal funds. We now we allow the platform to raise dollar funds by issuing equity shares to outside investors subject to a fixed financing cost, χ.[17] To characterize the optimal recapitalization policy, we first note that when issuing equity, the platform raises enough funds so that C jumps to $\overline{C}$

[17] Firms face significant financing costs due to asymmetric information and incentive issues. A large literature has sought to measure these costs, in particular, the costs arising from the negative stock price

where $V'(\overline{C}) = 1$. Once the fixed cost χ is paid, raising one more dollar does not incur further costs, so as long as the marginal value of reserves, $V'(C)$, is greater than one, the platform keeps raising funds.[18]

Moreover, the platform raises external funds only when C falls to zero. First, it is not optimal to issue equity at the payout boundary, $\overline{C}$, because newly raised funds will be paid out immediately and thus the issuance cost is incurred without any benefits. Therefore, let $\underline{C}$ denote the recapitalization boundary and we have $\underline{C} < \overline{C}$. Consider the *change* of existing shareholders' value after equity issuance: $[V(\overline{C}) - (\overline{C} - \underline{C}) - \chi] - V(\underline{C})$. To arrive at the existing shareholders' value, we deduct the issuance cost, χ, and competitive investors' equity value (equal to the funds raised), $(\overline{C} - \underline{C})$, from the total platform value post-issuance, $V(\overline{C})$. To calculate the change, we subtract $V(\underline{C})$, the value without issuance. Taking the derivative with respect to $\underline{C}$, we obtain $-V'(\underline{C}) + 1 < 0$ for $\underline{C} < \overline{C}$ because $V'(\underline{C}) > V'(\overline{C}) = 1$ under value function concavity.[19] Therefore, the platform prefers $\underline{C}$ to be as low as possible and thus optimally sets it to zero.

Finally, as in the baseline model, the platform can avoid liquidation by off-loading risk to users, as shown in (6.21), and obtain the value given by (6.22). Therefore, as C approaches zero, the platform only opts for recapitalization at $C = 0$ if recapitalization generates a higher value. Accordingly, the lower boundary condition (6.22) for the value function is modified to

$$\lim_{C \to 0^+} V(C) = \max \left\{ \lim_{C \to 0^+} V'(C) \frac{A}{\rho} \left(\frac{\xi}{\eta\sigma} \right)^{\frac{1}{1-\xi}} \left(\frac{1-\xi}{\xi} \right) \eta\sigma\,,\; V(\overline{C}) - \overline{C} - \chi \right\}, \tag{6.33}$$

The first term in the max operator is the value obtained from off-loading risk to users, given by (6.22). The second term is the post-issuance value for existing shareholders. The results in Proposition 24 to 25, 26 and Corollary 11 still hold except that the boundary condition (6.22) is replaced by (6.33).

Proposition 28 (Optimal Recapitalization) *The platform raises external funds through equity issuance only if* $V(\overline{C}) - \overline{C} - \chi > \lim_{C \to 0^+} V'(C) \frac{A}{\rho} \left(\frac{\xi}{\eta\sigma} \right)^{\frac{1}{1-\xi}} \left(\frac{1-\xi}{\xi} \right) \eta\sigma$ *(see (6.33)), where* $\overline{C}$ *is given by Proposition 24, and only when excess reserves fall to zero. The amount of funds raised is* $\overline{C}$.

When recapitalization happens, C_t jumps from zero to $\overline{C}$. If the platform is sufficiently risk-averse and charges high transaction fees and/or shares risk with users by setting $\sigma^P(0)$, it follows $N(0) < N(\overline{N})$, so that the aggregate token demand N_t jumps up at recapitalization, reflecting that the platform lowers fees $f(C)$ when it holds more reserves and becomes less risk-averse. If the platform does not adjust the token supply, S_t, there will be an upward *predictable* jump in the dollar price of token at $C = 0$, which implies an arbitrage opportunity.

To preclude arbitrage, the platform must expand token supply when it adjusts fees right after recapitalization so that the dollar price of token stays at its pre-issuance level. Since C_t already reaches the payout boundary, $\overline{C}$, the dollar proceeds from supplying new tokens are immediately paid out to shareholders. From a balance-sheet perspective, total assets

reaction in response to the announcement of a new issue. Lee, Lochhead, Ritter, and Zhao (1996) document that for initial public offerings (IPOs) of equity, the direct costs (underwriting, management, legal, auditing and registration fees) average 11.0% of the proceeds, and for seasoned equity offerings (SEOs), the direct costs average 7.1%. IPOs also incur a substantial indirect cost due to short-run underpricing. An early study by Asquith and Mullins (1986) found that the average stock price reaction to the announcement of a common stock issue was −3% and the loss in equity value as a percentage of the size of the new equity issue was as high as −31% (see Eckbo, Masulis, and Norli, 2007, for a survey).

[18] If a proportional cost of equity issuance is introduced, fund raising stops where $V'(C)$ is equal to one plus the proportional cost. For simplicity, we only introduce the fixed cost. Our analysis below on the token supply policy at the recapitalization boundary will not change in the presence of proportional cost of issuance.

[19] To prove the concavity of value function stated in Proposition 24, we only need the HJB equation (6.19) and the upper boundary conditions (6.17) and (6.18), so recapitalization does not affect value function concavity.

(reserves) stay at $\overline{C}$ while, on the liability side, token liability increases and equity decreases (through payout). This liability-structure adjustment is akin to corporations issuing debts for share repurchase.

Corollary 13 (Post-Recapitalization Adjustment of Liability Structure) *To preclude arbitrage in the token market, the platform adjusts its liability structure immediately after capitalization, supplying new tokens and paying out the dollar proceeds to the platform's shareholders.*

Finally, we revisit the results on the token price level in Corollary 12 and Proposition 27. Let $P_j(C)$ denote the token price function after the j-th recapitalization. According to Corollary (13), when recapitalization happens for the j-th time and C_t jumps from zero to $\overline{C}$, the platform sets

$$P_j(\overline{C}) = P_{j-1}(0)\,, \tag{6.34}$$

which replaces (6.32) as the boundary condition for the price-level ODE (6.31). Token price level before the first recapitalization, $P_0(C)$ is still solved under the boundary condition (6.32), i.e., $P_0(\overline{C}) = 1$.

Corollary 14 (Recapitalization and Token Price Level) *Token price level after the j-th recapitalization is solved by the ordinary differential equation (6.31) under the boundary condition (6.34).*

In the baseline model without recapitalization, the debasement of token is temporary: Token price level falls below 1 when C falls below $\widetilde{C}$ due to negative shocks and it recovers back to 1 when the platform accumulates sufficient amount of dollar revenues so that C crosses above $\widetilde{C}$ (Proposition 27). When recapitalization happens, the debasement is permanent. After the j-th recapitalization, token price level starts anew at a lower peg, $P_j(\overline{C}) = P_{j-1}(0)$, and if negative shocks deplete the platform's reserves and triggers another recapitalization, token price level declines along the process and, right after recapitalization, stabilizes at an even lower peg, $P_{j+1}(\overline{C}) = P_j(0)$.

Discussion: Financial Frictions. In our model, financial frictions play a key role. If costless recapitalization is possible, i.e., $\chi = 0$, then the platform will never allow the marginal value of reserves to exceed one because, when $V'(C) > 1$, it is profitable to raise funds from competitive investors that cost 1 per dollar and generates a value of $V'(C) > 1$. The constant marginal value of reserves implies that the platform is no longer risk-averse, i.e., $\gamma(C) = 0$, and thus, will absorb all risk, setting σ_t^P to zero. Tokens will always be redeemed at a fixed dollar value.

6.4 Regulating stablecoins

In this section, we apply our model to analyze two types of stablecoin regulations. The first type, which is of our focus, stipulates a minimum level of excess reserves held by the platform ("capital requirement"). Violating such a regulation triggers liquidation. The rationale behind is to guarantee a sufficient risk buffer so that the platform is likely to sustain stable redemption value. The second type is a more direct intervention that targets the fluctuation of token value. It requires the platform to keep token volatility below a certain level ("volatility regulation"). Our conclusion is that capital requirement delivers desirable equilibrium outcome, and if carefully designed, can achieve Pareto improvement for the platform and its users. Volatility regulation, in contrast, destroys the economic surplus from risk-sharing between the platform and its users.

6.4.1 Capital requirement

The regulator imposes a capital requirement, $C \geq C_L$, on the stablecoin platform and forces the platform to liquidate if the regulation is violated. Therefore, C_L replaces zero as the lower bound of excess reserves. In Figure 6.8, we plot the payout boundary $\overline{C}$ (Panel A), which is a measure of voluntary over-collateralization, and the welfare measures for different values

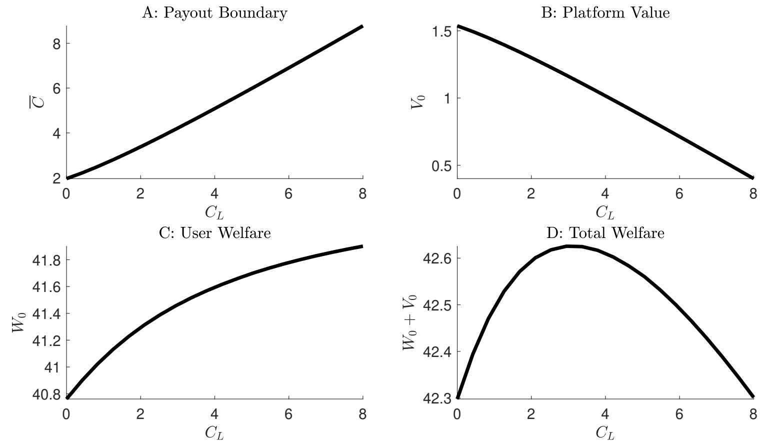

FIGURE 6.8: **Capital Requirement and Welfare.** We plot the numerical solutions of payout boundary $\overline{C}$ (Panel A), the platform shareholders' value at $t = 0$, V_0 (Panel B), users' welfare at $t = 0$, W_0 (Panel C), and total welfare $W_0 + V_0$ (Panel D). The parameterization follows Figure 6.1.

of C_L. Not so surprisingly, when the capital requirement tightens, the whole region of excess reserves is pushed to the right, resulting in a higher payout boundary $\overline{C}$ in Panel A. Because reserves earn an interest rate r that is below the shareholders' discount rate ρ, the platform shareholders' value, V_0, declines in C_L, as shown in Panel B. Panel C shows that users' welfare is improved by the capital requirement but there exists a significant degree of decreasing return as the regulator pushes up C_L.

What is interesting is that, in Panel D of Figure 6.8, the total welfare is non-monotonic in C_L. When the regulator increases C_L from zero, the increase of users' welfare overwhelms the decrease of platform value, but as the capital requirement is tightened, the loss of platform value eventually dominates. This suggests the existence of an optimal level of C_L that maximizes the total welfare.

As long as the users' welfare increases faster than the platform value decreases, the regulator can administer a transfer from users to the platform, making the regulation Pareto-improving. For example, the regulator can allow the platform to charge users a membership fees, i.e., a fixed cost of access, and imposes a cap on such fees. This type of access fees is commonly seen in the literature on regulation of utility networks (Laffont and Tirole, 1994; Armstrong, Doyle, and Vickers, 1996).

In Figure 6.9, we further demonstrate the stabilization effects of capital requirement. In Panel A, we plot the ratio of $\overline{C} - \widetilde{C}$ to $\overline{C} - C_L$ that measures the size of the stable subset of C where the platform maintains token redemption at par (i.e., $P(C) = 1$). As C_L increases, the stable region enlarges. In Panel B, we plot the probability of $C > \widetilde{C}$ (i.e., $\sigma^P(C) = 0$) based on the stationary distribution of C, which shows that over the long run the platform spends more time in the stable region when C_L increases. In Panel C, we plot the long-run average value of σ_t^P using the stationary probability distribution. A declining pattern emerges, indicating that capital requirement is indeed effective in reducing the token volatility. In Panel D, we plot the expected number of years it takes to reach $\widetilde{C}$ from C_L (denoted by $\tau(C_L)$). This recovery time decreases when the capital requirement is tightened. The intuition is that, as C_L increases, the platform near C_L still has abundant cash that self-accumulates over time by earning the interest rate r.

Lastly, Figure 6.10 demonstrates how network effects, as captured by α, drive the optimal capital requirement. Panel A plots the capital requirement C_L^* that maximizes total welfare. Panel B plots total welfare both with optimal capital requirements (solid line) and without capital requirements (dotted line). Panel C plots the welfare wedge between the optimally

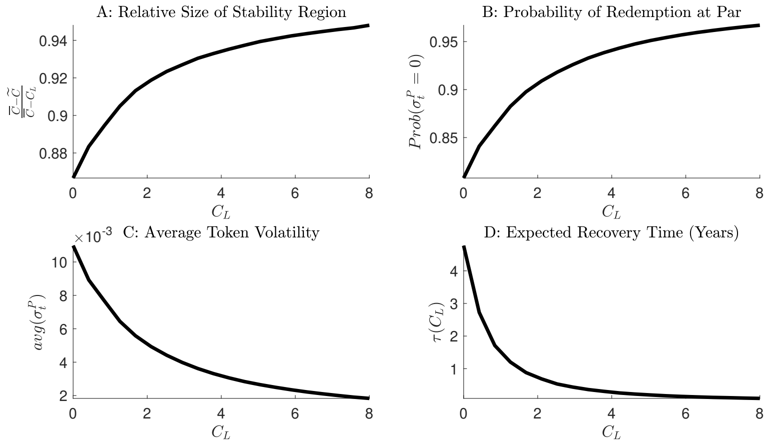

FIGURE 6.9: **Capital Requirement and Token Stability.** Using numeric solutions under different values of C_L, we plot the fraction of state space with redemption at par $\frac{\overline{C}-\widetilde{C}}{\overline{C}-C_L}$ (Panel A), stationary probability of zero token volatility (Panel B), the long-run average (based on stationary probability density) of token volatility (Panel C), and the expected time to reach $\widetilde{C}$ from C_L (Panel D). The parameterization follows Figure 6.1.

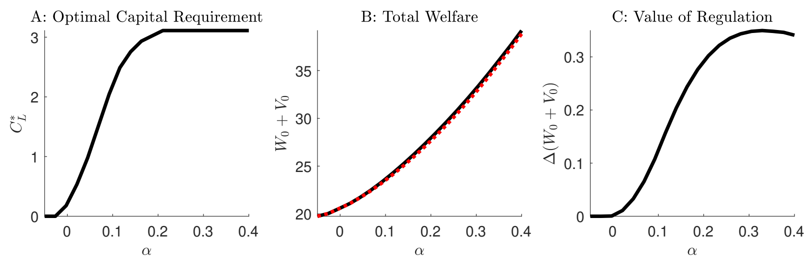

FIGURE 6.10: **Capital Requirement and Network Effects.** We calculate the optimal capital requirement C_L^* that maximizes total welfare (Panel A), total welfare both with capital requirement C_L^* (solid black line) and without capital requirement (dotted red line) (Panel B), and the welfare wedge between the optimally regulated equilibrium and the laissez-faire equilibrium (Panel C) over different values of α. Note that Panel C depicts the difference between the solid black line and the dotted red line from Panel B. The rest of parameterization follows Figure 6.1.

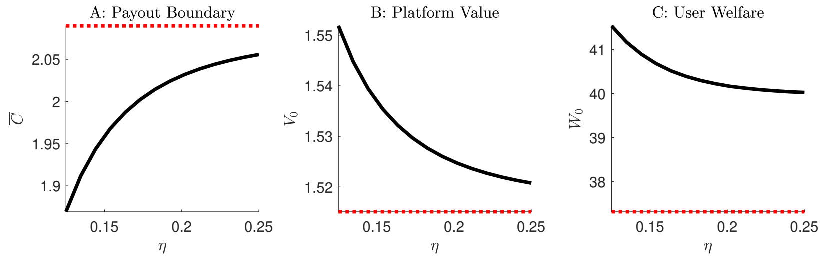

FIGURE 6.11: **Risk-Sharing, Volatility Regulation, and Welfare.** Using the numerical solutions, we calculate the payout boundary $\overline{C}$ (Panel A), the platform shareholders' value at $t = 0$, V_0 (Panel B), users' welfare at $t = 0$, W_0 (Panel C), and the long-run average fees based on stationary probability density (Panel D) over different values of users' risk aversion η for both the baseline model (solid line) and the model under zero-volatility regulation (red dotted line). The parameterization follows Figure 6.1.

regulated equilibrium and laissez-faire equilibrium. Importantly, it is optimal to raise capital requirement as network effects strengthen. In fact, platforms with no network effects ($\alpha = 0$) do not benefit at all from the regulation. The key to this result is the positive externality of individual users' token holdings, which explains the deviation of laissez-faire equilibrium from social optimum. The platform internalizes such externality in its decision to preserve reserves and stabilize tokens, but the internalization is not perfect. As shown in Figure 6.7, the platform cannot seize the full surplus as its share of total welfare is rather stable in α and always below 100%. Therefore, as α increases, the total welfare increases together with the component that is not internalized by the platform. This calls for a tighter capital requirement that moves the overall level of reserves closer to social optimum.

6.4.2 Volatility regulation

The concern over token volatility and debasement may also motivate more direct regulatory intervention. A volatility regulation imposes a cap on σ_t^P and requires the platform to implement a rule of token supply to achieve this goal. In Figure 6.11, we compares the baseline model (solid line) and the model under regulation that forces $\sigma_t^P = 0$ (dotted line) over a range of values of users' risk aversion η.[20] As shown by the HJB equation (6.19), under $\sigma_t^P = 0$, the platform's value function and control variables no longer depends on η, so the (dotted) lines are flat in all panels.

In Panel A of Figure 6.11, we show that under the zero-volatility requirement, the platform has to maintain a higher level of excess reserves to reduce the likelihood of liquidation because the option of off-loading risk to users is no longer available. Holding more reserves with an interest rate below the shareholders' discount rate leads to a lower platform value, as shown in Panel B of Figure 6.11.

An interesting finding is that imposing the volatility regulation even decreases users' welfare (Panel C of Figure 6.11) across all values of η. This seems to contradict the intuition that, by forcing the platform to maintain a perfectly stable token value, users will benefit, especially when they are more risk-averse. However, the argument ignores that, unable to off-load risk to users, the platform can compensate its risk exposure with higher fees. Volatility regulation is counterproductive because it limits the risk-sharing between the platform and its users. When the platform is close to liquidation, its effective risk aversion can be

[20] Because the platform can no longer off-load risk to users as C approaches zero, liquidation can happen and $C = 0$ becomes an absorbing state. The boundary condition $V(0) = 0$ implies $-V''(C)$ approaches infinity, which is the same boundary condition that we use to solve the baseline model (see Footnote 14).

higher than η, so there is economic surplus created from users' absorbing risk from the platform. Volatility regulation shuts down this insurance market.

6.5 Payment and big data

User-generated data is now a major asset of digital platforms. Social networks, such as Facebook and Twitter, utilize such data to target users for advertisement. Being able to utilize the enormous amount of transaction data has become a critical advantage of digital platforms relative to traditional payment service providers such as banks (Bank for International Settlements, 2019). Leading players, such as PayPal and Square, have become data centers and provide services beyond facilitating payments, for example, extending loans to consumers and businesses based on credit analysis that is enabled by transaction data. In this section, we follow Veldkamp (2005), Ordoñez (2013), Fajgelbaum, Schaal, and Taschereau-Dumouchel (2017), and Jones and Tonetti (2020) to model big data as a by-product of user activities.[21] Our analysis focuses how data as a productive asset affects the optimal strategies of stablecoin platforms and the efficacy of regulations.

6.5.1 Big data as a productive asset

In the baseline model, the quality parameter A is constant. We now interpret A as a measure of effective data units that enhance platform quality and assume the following law of motion

$$dA_t = \kappa A_t^{1-\xi} N_t^{\xi} dt \,. \tag{6.35}$$

Users' transactions generate a flow of raw data, $\kappa N_t^{\xi} dt$, where the parameter κ captures the technological efficiency of data processing and storage. To what extent the raw data contributes to the effective data units depends on the current amount of effective data via $A_t^{1-\xi}$. The complementarity between the old and new data captures the fact that the value of new data increases in the quality of statistical algorithms, which in turn depends on the amount of existing data that are needed to to select and train the algorithms.[22] The Cobb-Douglas form is chosen for analytical convenience.[23]

As the platform quality improves, we assume the transaction capacity increases accordingly, i.e., $\overline{N}_t = \overline{n} A_t$, where $\overline{n} > 0$ is constant. User optimization is static and follows the baseline model. As shown in (6.10), the dollar transaction volume (or token demand) $N_t \equiv n_t A_t$ where

$$n_t = \frac{1}{\left(r + f_t - \mu_t^P + \eta |\sigma_t^P|\right)^{\frac{1}{1-\xi}}} \wedge \overline{n} \,. \tag{6.36}$$

As in the baseline model, the platform sets n_t through the fees, f_t, and set the dynamics of token redemption price through its choice of σ_t^P. The model now has three natural state variables, reserves M_t, token supply S_t, and data stock A_t. Similar to the baseline model, $C_t = M_t - S_t P_t$ and A_t summarize payoff-relevant information, driving the platform value, $V_t = V(C_t, A_t)$, and the dollar value of token, $P_t = P(C_t, A_t)$. To simplify the notations, we will suppress the time subscripts.

We conjecture that the system is homogeneous in A, and in particular, the platform's value function and dollar value of token are given by $V(C, A) = v(c)A$ and $P(C, A) = P(c)$, respectively, where the excess reserves-to-data ratio is the key state variable for the platform's optimal strategies:

$$c \equiv \frac{C}{A} \,. \tag{6.37}$$

[21] Veldkamp and Chung (2019) provide an excellent survey of the literature of data and aggregate economy.

[22] Related, in Farboodi, Mihet, Philippon, and Veldkamp (2019), data have increasing return to scale.

[23] This data accumulation process is inspired by the specification of knowledge accumulation in Weitzman (1998).

We will confirm the conjecture as we solve the platform's optimization problem in the following. First, to derive the law of motion of c_t, we follow the derivation of the baseline model to obtain

$$dC_t = \left(rC_t + A_t n_t^{\xi} - \eta A_t n_t |\sigma_t^P|\right) dt + A_t n_t (\sigma - \sigma_t^P) dZ_t - dDiv_t \,. \tag{6.38}$$

Given (6.35) and (6.38), the law of motion of c_t is given by

$$dc_t = \left(rc_t + n_t^{\xi} - \eta n_t |\sigma_t^P| - \kappa n_t^{\xi} c_t\right) dt + n_t (\sigma - \sigma_t^P) dZ_t - \frac{dDiv_t}{A_t} \,. \tag{6.39}$$

Under the value function conjecture, $V(C, A) = v(c)A$, and the laws of motion of A (6.35) and c (6.39), the HJB equation for $v(c)$ in the interior region (where $dDiv_t = 0$) is given by

$$\rho v(c) = \max_{n \in [0,\overline{n}], \sigma^P} \left\{ \left[v(c) - v'(c)c\right] \kappa n^{\xi} + v'(c) \left(rc + n^{\xi} - \eta n |\sigma^P|\right) + \frac{1}{2} v''(c) n^2 (\sigma - \sigma^P)^2 \right\} , \tag{6.40}$$

When the marginal value of reserves, $V_A(C, A) = v'(c)$, falls to one, the platform pays out dividends. We define the payout boundary as $\overline{c}$ through $v'(\overline{c}) = 1$. The optimality of $\overline{c}$ also implies $v''(\overline{c}) = 0$. Note that as in the baseline model, when C (or c) approaches zero, the platform can avoid liquidation by setting $\sigma^P(c) = \sigma$ to off-load risk to its users and gradually replenish reserves.[24] For simplicity, we do not consider recapitalization (equity issuance). In sum, the platform's excess reserves, C_t, move in $(0, \overline{c}A]$. As data grows, the platform accumulates more excess reserves.

Proposition 29 (Data Growth and Over-Collateralization) *The amount of excess reserves, C_t, varies in $(0, \overline{c}A_t]$ where the upper bound increases with A_t, the effective data units.*

The intuition behind Proposition 29 is that data growth provides another channel through which the continuation value appreciates, as shown on the right side of (6.40), which makes the platform more patient in distributing excess reserves to shareholders. The first term on the right side contains the marginal value of data (which we call "data q")

$$q(c) = \frac{\partial V(C, A)}{\partial A} = v(c) - v'(c)c \,. \tag{6.41}$$

Retaining more reserves allows the platform to sustain a wider region of c with credible token redemption at par. A more stable token in turn stimulates transactions and thereby allows the platform to accumulate more data and earn the data q, $q(c)$. Data as a productive asset and by-product of user activities enhances the platform's incentive to reserve for tokens.

Next, we characterize the optimal transaction volume and token volatility. Following our analysis of the baseline model, we define the effective risk aversion based on $v(c)$:

$$\Gamma(c) = -\frac{v''(c)}{v'(c)} \,. \tag{6.42}$$

The following proposition summarizes the optimal choices of $n(c)$ and $\sigma^P(c)$.

Proposition 30 (Data q, Token Volatility, and Transaction Volume) *At c where the platform maintains the redemption of token at par, i.e., $\sigma^P(c) = 0$, the optimal transaction volume is*

$$N = n(c)A = \left[\frac{\xi}{\Gamma(c)\sigma^2}\left(1 + \frac{\kappa q(c)}{v'(c)}\right)\right]^{\frac{1}{2-\xi}} A \;\; \wedge \overline{n}A \,; \tag{6.43}$$

[24] The boundary condition for $v(c)$ is that as c approaches zero, $-v''(c)$ approaches infinity (see footnote 14).

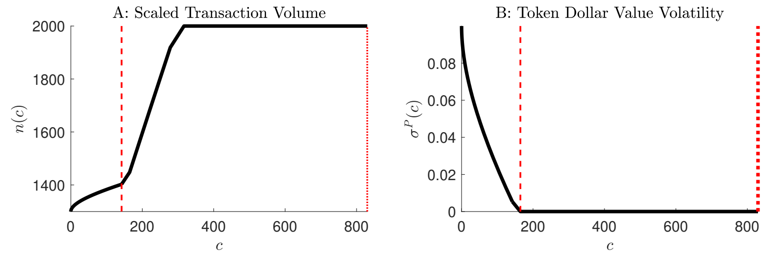

FIGURE 6.12: **Transaction Volume and Token Volatility.** This figure plots the A_t-scaled transaction volume $n(c)$ in Panel A and token return volatility $\sigma^P(c)$ in Panel B. In both panels, the red dotted lines mark the payout boundary $\overline{c}$, and the red dashed line marks $\widetilde{c}$, the threshold that separates the regions of volatile and constant token prices. The parameterization follows Figure 6.1 with the additional parameters $\overline{n} = 2000$ and $\kappa = 0.00025$. Note that $\overline{n} = 2000$ implies that for $A_0 = 0.0025$, $\overline{N}_t = \overline{n}A_t = 5$ as under the parameterization in the baseline (see Figure 6.1).

otherwise, the optimal token volatility is

$$\sigma^P(c) = \sigma - \frac{\eta}{\Gamma(c)n(c)} \in (0, \sigma)\,, \tag{6.44}$$

and the optimal transaction volume is

$$N = n(c)A = \left[\frac{\xi}{\eta\sigma}\left(1 + \frac{\kappa q(c)}{v'(c)}\right)\right]^{\frac{1}{1-\xi}} A \ \wedge \overline{n}A\,. \tag{6.45}$$

The optimal transaction volume is proportional to A, the effective data units. Therefore, as data grows following (6.35), the transaction volume grows too. With data as a productive asset, the platform faces a new trade-off. It can accumulate more reserves through higher fee revenues or, by reducing fees, boost the transaction volume to accumulate more data. Therefore, the ratio of marginal value of data (the data q) and marginal value of reserves, $q(c)/v'(c)$, emerges in both (6.43) and (6.45). When the data q is higher relative to the marginal value of reserves, the platform implements a higher transaction volume through lower fees. Note that given the token price dynamics, the monotonic relationship between transaction volume and fees is given by (6.36).

The optimal choice of token volatility resembles that in the baseline model. In the region where $\sigma^P(c) > 0$, it is the ratio of users' risk aversion to the platform's risk aversion that drives $\sigma^P(c)$. And in this region, the optimal transaction volume in (6.45), even scaled by A, is no longer the constant as in the baseline model but depends $q(c)/v'(c)$ instead, showing the trade-off between investing in data and accumulating reserves. Moreover, the optimal transaction volume depends on users' risk aversion η as η determines the cost of obtaining insurance from users (losing transaction volume after off-loading risk to users). When the platform absorbs all risk (i.e., $\sigma^P(c) = 0$), the optimal transaction volume varies with its own risk aversion $\Gamma(c)$ (6.43) because $\Gamma(c)$ drives the required risk compensation through higher fees that causes the transaction volume to decline.

Panel A of Figure 6.12 reports the optimal transaction volume. In contrast to Panel A of Figure 6.3 where the transaction volume is constant in the region where $\sigma^P(c) > 0$, the A-scaled transaction volume now increases in c. The intuition is that as reserves become more abundant relative to data, the platform is more willing to lower fees, so it acquires more data through more active transactions at the expense of lower dollar revenues added to the reserve buffer. Panel B of Figure 6.12 shows a similar token volatility dynamics as Panel A of Figure 6.2 from the baseline model.

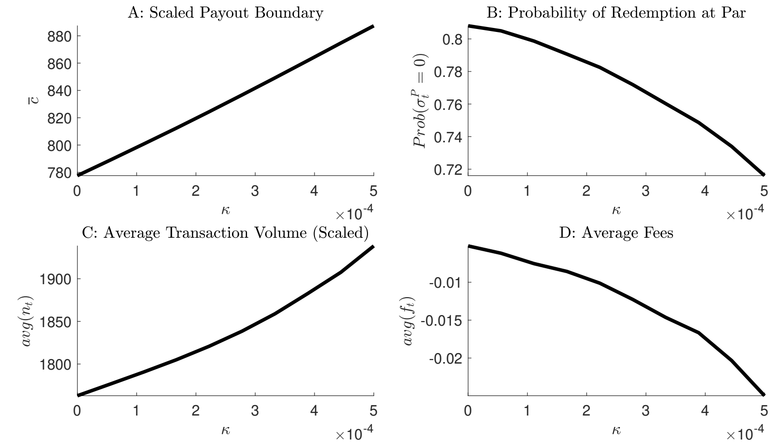

FIGURE 6.13: **Data Technology Progress and Platform Operation.** We plot the A-scaled payout boundary (Panel A), the probability of token redemption at par (Panel B), the average transaction volume (Panel C), and the average fees per dollar of transactions (Panel D) over κ (the efficiency of data technology). The moments in Panel B, C, and D are based on the stationary distribution of c. The parameterization follows Figure 6.1 with $\overline{n} = 2000$.

6.5.2 Data technology revolution and stablecoin platform strategies

The last few decades have witnessed enormous progress in data science. In our model, such technological advance can be captured by an increase of the parameter κ. In Figure 6.13, we examine the impact of big data technology on the operation of stablecoin platforms. In Panel A, we show that in response to an increase in κ, the platform optimally raises the (A-scaled) payout boundary, $\overline{u}$, which suggests a greater degree of over-collateralization. The intuition of such response can be understood jointly with the platform's decision on token volatility and fees.

To accumulate transaction data, the platform would increase the transaction volume. This can be achieved through lower fees. As shown in Panel C and D of Figure 6.13, the average fees (calculated from the stationary distribution of c) decline and the transaction volume increases in κ. The average fees per dollar of transaction even dips into the negative territory, becoming subsidies to users. This prediction is consistent with the practice that large digital platforms offer subsidies and fee services to retain and grow their customer base (Rochet and Tirole, 2006; Rysman, 2009).

However, a higher n implies a large exposure to operation risk as shown in (6.39). The platform responds by delaying payout, i.e., raising the boundary $\overline{c}$, to increase the reserve buffer, which explains why the payout boundary $\overline{c}$ increases in κ in Panel A of Figure 6.13. The platform can also respond by off-loading more risk to users. As shown in Panel B, the stationary distribution of c implies a smaller probability of $\sigma^P(c) = 0$ and a higher average $\sigma^P(c)$ when κ increases.

In sum, when transaction data can be better utilized, the platform becomes more aggressive in raising transaction volume through fee reduction (or subsidy). Accordingly, the platform maintains more reserves to buffer the resultant increase in operation risk. Part of the increased risk is shared with users through token price fluctuation. In Figure 6.14, we show that the improving efficiency of data technology increases the total welfare (Panel A) while the platform's share is rather stable and always below 100%. Therefore, even though the platform has monopolistic power as a unique marketplace where users transact with each other using tokens, the platform cannot possess the full economic surplus created by big data technology. Data originates from user activities, so to obtain data, the platform

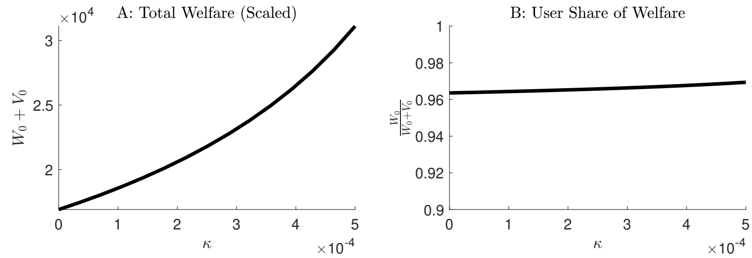

FIGURE 6.14: **Data Technology Progress and Welfare.** We plot the sum of platform value and users' welfare (Panel A), and users' share of total welfare (Panel B) against κ (the efficiency of data technology). The parameterization follows Figure 6.1. The parameterization follows follows Figure 6.1 with $\overline{n} = 2000$ and all quantities are scaled by data units A.

must share the economic surplus with users. These outcomes also suggest that regulations targeting and limiting the use of transaction data undermine the platform's incentives to accumulate liquidity reserves and are detrimental for both user and total welfare.

6.5.3 Data technology revolution and stablecoin regulation

Because the transaction volume is proportional to A, equation (6.35) implies an exponential growth of effective data units that scales up the platform value and users' welfare. The improving efficiency of data technology causes the exponential growth to be increasingly steeper. In such an environment, how should the optimal capital requirement adjust? In this subsection, we address this question.

As previously discussed, a larger transaction volume N amplifies the shock exposure of reserves, and to achieve a larger transaction volume, the platform has to lower fees, sacrificing the growth of dollar reserves. Therefore, there exists tension between precautionary management of reserves and data acquisition through users' transactions. Capital requirement favors preserving reserves over stimulating transaction volume for data acquisition. Therefore, as data becomes more productive, i.e., κ increases, capital requirement becomes more burdensome.

Panel A of Figure 6.15 confirms the intuition. The optimal requirement of excess reserves (scaled by A) declines in κ. We study the scaled capital requirement to preserve the homogeneity property of the system and keep the solution in one-dimensional space of $c = C/A$. Indeed, tightening capital requirement causes the platform to build up reserves at the expense of data acquisition. However, given the self-reinforcing growth of data in (6.35), such regulatory measure hurts the long-run exponential growth of both platform value and users' welfare. In Panel B of Figure 6.15, we compare the total welfare under *optimal* capital requirement with that from the laissez-faire equilibrium, and in Panel C we plot the wedge. The increase of welfare in κ is not surprising. What is interesting is that the benefit of capital requirement dwindles as κ increases in Panel C.

Moreover, as we show in Figure 6.13, data as a self-accumulating productive asset offers a new opportunity for shareholders' equity to growth over time. This effectively makes the platform more patient in paying out dividends. Therefore, the voluntary build-up of reserves is strengthened as κ increases. As a result, capital requirement is less needed for the internalization of user-network effects. In sum, as data becomes more productive, the role of capital requirement weakens.

6.6 Conclusion

The first-generation cryptocurrencies, such as Bitcoin and Ethereum, were built to serve as transaction medium, but the enormous volatility in price compromises such functionality.

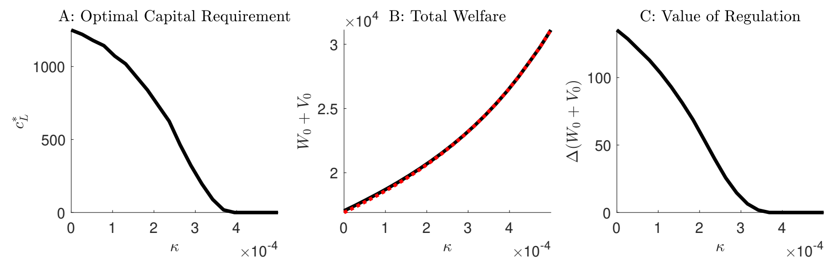

FIGURE 6.15: **Data Technology Progress and Capital Requirements.** We calculate the optimal scaled capital requirement $c_L^* \equiv C_L^*/A$ that maximizes total welfare (Panel A), scaled total welfare both with scaled capital requirement c_L^* (solid black line) and without capital requirement (dotted red line) (Panel B), and the welfare wedge between the optimally regulated equilibrium and laissez-faire equilibrium (Panel C) over different values of κ. Note that Panel C depicts the difference between the solid black line and the dotted red line from Panel B. The parameterization follows follows Figure 6.1 with $\overline{n} = 2000$ and all quantities are scaled by data units A.

For this reason, stablecoins — cryptocurrencies or tokens whose price is designed to be stable — have become increasingly popular. Typically, stablecoins are issued by private entities who promise to maintain price stability by holding collateral against which stablecoin holdings can be redeemed. However, as these private stablecoin issuers maximize their own payoffs rather than the total welfare, conflicts of interests between the issuers and users of stablecoins may arise, making room for welfare-enhancing regulations of stablecoin issuance. However, to this date, there has not been a systematic theoretical analysis of stablecoins, in spite of the enormous attention from regulators. In this paper, we fill this gap and develop a dynamic model of optimal stablecoin management and regulation.

In our model, users hold redeemable stablecoins to transact on a digital platform, and stablecoins can be exchanged for dollars, or vice versa, at an exchange rate set by the platform under its own discretion. The platform maintains liquidity reserves to meet token redemption and controls transaction volume through fees. When reserves are high, users can redeem tokens at par and pay low transaction fees charged by the platform, so that transaction volume is high. When reserves are low, the platform charges high fees and optimally debases token to share liquidity risks with users, in which case, user can only redeem tokens below par and transaction volume suffers. We then discuss optimal regulation of stablecoins. We find that regulation precluding token debasement is detrimental to welfare. In contrast, capital requirement, which stipulates a minimal level of reserve holdings, adds value by incentivizing the platform to internalize user-network effects.

The accumulation and utilization of transaction data launches self-reinforcing growth of transaction volume, platform value, and user welfare. By enhancing the platform's incentive to reserve for stablecoins, data as a productive asset weakens the role of capital requirement. Moreover, capital requirement imposed on a data-enabled stablecoin platform can have the unintended consequence of favoring reserve preservation over data acquisition, stemming the data-driven exponential growth.

6.7 Appendix

6.7.1 Value function concavity

We prove the concavity of value function in Proposition 24. Recall the HJB equation (6.19), that is,

$$\rho V(C) = \max_{\{N \in [0,\overline{N}], \sigma^P\}} \left\{ V'(C) \left(rC + N^{\xi} A^{1-\xi} - \eta N |\sigma^P| \right) + \frac{1}{2} V''(C) N^2 (\sigma - \sigma^P)^2 \right\}.$$

Using the envelope theorem, we differentiate both sides of the HJB equation (evaluated under the optimal controls) with respect to C to obtain

$$\rho V'(C) = rV'(V) + V''(V)\left(rC + N^{\xi}A^{1-\xi} - \eta N|\sigma_t^P|\right) + \frac{1}{2}V'''(C)N^2(\sigma - \sigma^P)^2.$$

We can solve for

$$V'''(C) = \frac{2}{N^2(\sigma - \sigma^P)^2}\left[(\rho - r)V'(C) - V''(V)\left(rC + N^{\xi}A^{1-\xi} - \eta N|\sigma_t^P|\right)\right]$$

Using the smooth pasting condition, $V'(\overline{C}) = 1$, and the super-contact condition, $V''(\overline{C})$, it follows that $V'''(\overline{C}) > 0$. As $V''(\overline{C}) = 0$, it follows that $V''(C) < 0$ in a left-neighbourhood of $\overline{C}$.

We show that $V''(C)$ for all $C \in [0, \overline{C})$. Suppose to the contrary that there exists $\hat{C} < \overline{C}$ with $V''(\hat{C}) \geq 0$ and set without loss of generality

$$\hat{C} = \sup\{C \geq 0 : V''(C) \geq 0\}. \tag{6.46}$$

As $V''(C)$ in a left-neighbourhood of $\overline{C}$ and the value function is twice continuously differentiable, it follows that $V''(\hat{C}) = 0$ and therefore $\sigma^P(\hat{C}) < \sigma$. In addition, $V'(\hat{C}) \geq 1$, so that $V'''(\hat{C}) > 0$. Thus, there exists $C' > \hat{C}$ with $V''(C') \geq 0$, a contradiction. Therefore, the value function is strictly concave on $[0, \overline{C})$.

6.7.2 Optimal control variables

In this section, we characterize the optimization in (6.19) and solve for the optimal control variables $N = N(C)$ and $\sigma^P = \sigma^P(C)$ in Proposition 26. To start with, we define

$$\underline{N} = \arg\max_{N \leq \overline{N}} N^{\xi}A^{1-\xi} - \eta N\sigma,$$

which yields

$$\underline{N} = \min\left\{\left(\frac{\xi A^{1-\xi}}{\eta\sigma}\right)^{\frac{1}{1-\xi}}, \overline{N}\right\}.$$

Now, first optimize the HJB equation over σ^P or equivalently over $N\sigma^P$.

If interior, the choice of σ^P satisfies the first order optimality condition. The first-order-condition in (6.19) with respect to σ^P is

$$-\eta V'(C) - V''(C)(N\sigma - N\sigma^P) = 0$$

Thus,

$$N\sigma^P = \max\left\{0, \frac{-\eta V'(C) + N\sigma V''(C)}{-V''(C)}\right\} = \max\left\{0, -\frac{\eta V'(C)}{-V''(C)} + N\sigma\right\}$$

We distinguish between two different cases

1. $\sigma^P > 0$, in which case we can insert σ^P into (6.19) to get

$$\rho V(C) = \max_{N \in [0, \overline{N}]}\left\{V'(C)\left[rC + N^{\xi}A^{1-\xi} - \eta N\sigma - \frac{\eta^2 V'(C)}{V''(C)}\right] + \frac{1}{V''(C)}\left[\frac{(\eta V'(C))^2}{2}\right]\right\}. \tag{6.47}$$

(6.48)

Thus, $N = \underline{N}$ is the optimal choice of N, so that

$$\sigma^P = \max\left\{0, -\frac{\eta V'(C)}{-V''(C)\underline{N}} + \sigma\right\}$$

2. $\sigma^P = 0$, so the HJB becomes

$$\rho V(C) = \max_{N \in [0, \overline{N}]} \left\{ V'(C)[rC + N^{\xi} A^{1-\xi}] + V''(C) \left[\frac{N^2 \sigma^2}{2} \right] \right\}, \tag{6.49}$$

and

$$N(C) = \min \left\{ \left(\frac{A^{1-\xi} \xi V'(C)}{-V''(C)\sigma^2} \right)^{\frac{1}{2-\xi}}, \overline{N} \right\} \tag{6.50}$$

is the optimal choice of N.

It follows that

$$N(C) = \underline{N} \iff -\frac{\eta V'(C)}{-V''(C)\underline{N}} + \sigma > 0,$$

as desired.

6.7.3 Effective risk aversion

We prove $\gamma'(C) < 0$, i.e., $\frac{d(-V''(C)/V'(C))}{dC} < 0$, in Proposition 26. Consider the following two cases:

1. $\sigma^P > 0$, $N = \underline{N}$. The HJB equation can be simplified to

$$\rho \frac{V(C)}{V'(C)} = rC + \underline{N}^{\xi} A^{1-\xi} - \eta \underline{N} \sigma - \frac{\eta^2}{2} \frac{V'(C)}{V''(C)}. \tag{6.51}$$

Differentiating the equation above with respect to C we obtain that in $(0, \overline{C})$

$$\rho \left(1 - \frac{V''(C)V(C)}{V'(C)^2} \right) = r - \frac{\eta^2}{2} \frac{d(V'(C)/V''(C))}{dC}. \tag{6.52}$$

which implies $\frac{d(V'(C)/V''(C))}{dC} < 0$ (because $V''(C) < 0$ and $\rho > r$), i.e., $\frac{d(-V''(C)/V'(C))}{dC} < 0$.

2. $\sigma^P = 0$, so the HJB becomes

$$\rho V(C) = \max_{N \in [0, \overline{N}]} \left\{ V'(C)[rC + N^{\xi} A^{1-\xi}] + V''(C) \left[\frac{N^2 \sigma^2}{2} \right] \right\}, \tag{6.53}$$

In this case, we further consider two cases:

a. $N(C) < \overline{N}$ and $N = \left(\frac{A^{1-\xi} \xi V'(C)}{-V''(C)\sigma^2} \right)^{\frac{1}{2-\xi}}$. In this case, the HJB can be simplified to

$$\rho \frac{V(C)}{V'(C)} = rC + \frac{1}{2} \left(\frac{\xi A^{1-\xi}}{\sigma^{\xi}} \right)^{\frac{2}{2-\xi}} \left(\frac{2-\xi}{\xi} \right) \left(\frac{V'(C)}{-V''(C)} \right)^{\frac{\xi}{2-\xi}}. \tag{6.54}$$

Differentiating the equation above with respect to C, we obtain

$$\rho \left(1 - \frac{V''(C)V(C)}{V'(C)^2} \right) = r - \frac{1}{2} \left(\frac{\xi A^{1-\xi}}{\sigma^{\xi}} \right)^{\frac{2}{2-\xi}} \left(\frac{V'(C)}{-V''(C)} \right)^{\frac{2\xi-2}{2-\xi}} \frac{d(-V'(C)/V''(C))}{dC}, \tag{6.55}$$

implying $\frac{d(V'(C)/V''(C))}{dC} < 0$ (because $V''(C) < 0$ and $\rho > r$), i.e., $\frac{d(-V''(C)/V'(C))}{dC} < 0$.

b. $N(C) = \overline{N}$. In this case, the HJB can be simplified to

$$\rho \frac{V(C)}{V'(C)} = rC + \overline{N}^{\xi} A^{1-\xi} + \frac{\overline{N}^2 \sigma^2}{2} \frac{V''(C)}{V'(C)}. \tag{6.56}$$

Differentiating the equation above with respect to C, we obtain

$$\rho\left(1-\frac{V''(C)V(C)}{V'(C)^2}\right)=r-\frac{\overline{N}^2\sigma^2}{2}\frac{d(-V''(C)/V'(C))}{dC}, \tag{6.57}$$

which implies $\frac{d(-V''(C)/V'(C))}{dC}<0$ (because $V''(C)<0$ and $\rho>r$).

6.7.4 Calculating user welfare

To start with, recall that any users' utility flow is

$$d\hat{R}_{it}\equiv N_t^{\alpha}A^{1-\xi}\frac{u_{it}^{\beta}}{\beta}dt+u_{it}\Big(\frac{dP_t}{P_t}-rdt-f_tdt-\eta|\sigma_t^P|\Big)$$

As such,

$$\mathbb{E}d\hat{R}_{it}=N_t^{\alpha}A^{1-\xi}\frac{u_{it}^{\beta}}{\beta}dt+u_{it}\Big(\mu_t^P-rdt-f_tdt-\eta|\sigma_t^P|\Big).$$

Inserting $u_{it}=N_t$ and (6.11) and using $\xi=\alpha+\beta$ yields

$$\begin{aligned}\mathbb{E}d\hat{R}_{it}&=\frac{N_t^{\xi}A^{1-\xi}}{\beta}dt+N_t\Big(\mu_t^P-rdt-(N_t^{\xi-1}A^{1-\xi}+\mu_t^P-r-\eta|\sigma_t^P|)dt-\eta|\sigma_t^P|\Big)\\&=\frac{N_t^{\xi}A^{1-\xi}}{\beta}dt-N_t^{\xi}A^{1-\xi}dt=\frac{(1-\beta)A^{1-\xi}}{\beta}N_t^{\xi}dt.\end{aligned}$$

As a next step, define the user welfare from time t onward, i.e.,

$$W_t:=\mathbb{E}\left[\int_t^{\infty}e^{-r(s-t)}\left(dR_{is}-\eta u_{is}|\sigma_s^P|ds\right)\right]. \tag{6.58}$$

As C is the payoff-relevant state variable, we can express user welfare as function of C, in that $W_t=W(C_t)$. The dynamic programming principle implies that user welfare solves on $[0,\overline{C}]$ the ODE

$$rW(C_t)dt=\mathbb{E}d\hat{R}_{it}+\mathbb{E}dW(C_t).$$

We can rewrite the ODE as

$$rW(C)=\frac{(1-\beta)A^{1-\xi}}{\beta}N(C)^{\xi}+W'(C)\mu_C(C)+\frac{W''(C)\sigma_C(C)^2}{2}, \tag{6.59}$$

whereby

$$\begin{aligned}\mu_C(C)&=rC+N(C)^{\xi}A^{1-\xi}-\eta N(C)|\sigma^P(C)|\\\sigma_C(C)&=N(C)(\sigma-\sigma^P(C))\end{aligned}$$

are drift and volatility of net liquidity C respectively.

The ODE (6.59) is solved subject to the boundary conditions

$$W'(\overline{C})=0$$

and

$$\lim_{C\to 0}W(C)=\frac{1}{r}\lim_{C\to 0}\left(\frac{(1-\beta)A^{1-\xi}}{\beta}N(C)^{\xi}+W'(C)\mu_C(C)\right).$$

6.7.5 Calculating the expected arrival time

Note that there exists $\tilde{C} \in (0, \overline{C})$ such that $\sigma^P(C) = 0$. Given $C_t = C$ at time t, we calculate

$$\tau(C_t) = \mathbb{E}[\tau^* - t | C_t = C] \quad \text{with} \quad \tau^* = \inf\{s \geq t : C_s \geq \tilde{C}\},$$

which is the expected time until net liquidity reaches $\tilde{C}$ and token price volatility vanishes.

We can rewrite $\tau(C_t)$ as

$$\tau(C_t) = \mathbb{E}_t \left[\int_t^{t^*} 1 dt \right]. \tag{6.60}$$

By definition, it holds that when $C_t = C \geq \tilde{C}$, then $t^* = t$ and

$$\tau(C_t) = \tau(C) = 0.$$

By the integral expression (6.60) and the dynamic programming principle, it follows that For $C \leq \tau(C)$, the function $\tau(C)$ solves the ODE

$$0 = 1 + \tau'(C)\mu_C(C) + \frac{\sigma_C(C)^2 \tau''(C)}{2}, \tag{6.61}$$

where

$$\mu_C(C) = rC + N(C)^{\xi} A^{1-\xi} - \eta N(C) |\sigma^P(C)|$$
$$\sigma_C(C) = N(C)(\sigma - \sigma^P(C))$$

are drift and volatility of net liquidity C respectively. The ODE (6.61) is solved subject to the boundary condition

$$\tau(\tilde{C}) = 0 \tag{6.62}$$

at $C = \tilde{C}$. At $C = C_L$, the lower boundary of the state space, the boundary condition

$$\lim_{C \to C_L} [1 + \tau'(C)\mu_C(C)] = 0.$$

Chapter 7

Conclusion

In this thesis, we have analyzed different situations of corporate decision making under agency conflicts and information asymmetries between a firm's insiders and outsiders. The thesis contains five different essays that analyze different situations and different types of firms but all have a common theme, that is agency conflicts.

Chapter 2 "Financing Breakghroughs under Failure Risk" studies agency conflicts inherent to the financing of innovative projects that are subject to substantial failure risk and information asymmetries. In Chapter 2, we demonstrate that to financing development of innovative projects, the optimal financing contract should involve different stage, whereby the provision of financing becomes more performance-sensitive over time. In addition, to incentivize truthful disclosure of bad outcomes such as failure, the optimal contract stipulates time-decreasing rewards for failure and excessively high rewards for success. The paper also derives implications for venture capital financing, R&D financing, and executive compensation.

Chapter 3 "Agency Conflicts and Short- vs. Long-Termism in Corporate Policies" (joint with Sebastian Gryglewicz and Erwan Morellec) analyzes the optimal time horizon of corporate policies and investment in the presence of agency conflicts between a firm's manager and shareholders. We demonstrate that due to these agency conflicts, corporate policies exhibit opitmal short-termism and long-termism depending on circumstances and firm characteristics. Our findings explain investment patterns observed in firm-level data and generate empirical implications regarding the optimal horizon of corporate policies.

Chapter 4 "Delegated Monitoring and Contracting" (joint with Sebastian Gryglewicz) studies double agency problems that arise in delegated investment and contracting. Specifically, in our model, investors finance a firm run by an agent while delegating monitoring and contracting with the agent to an intermediary. We show that agency conflicts arise on intermediary and firm level, with endogenous spillover effects determining the severity of these agency conflicts. Notably, the intermediary's and agent's incentives are linked via so-called trickle-up and trickle-down incentives. The model characterizes under what circumstances intermediation and increased investor participation in firm governance is valuable. In addition, we apply our model to private equity financing and derive several empirical implications.

Chapter 5 "Optimal Financing with Tokens" (joint with Sebastian Gryglewicz and Erwan Morellec) presents a unifying model of token-based financing that nests different types of tokens and token offerings. The model demonstrates under what circumstances token financing is preferred over equity financing, characterizes the optimal design of tokens, and explains the large variation in the observed designs of tokens. We also demonstrate under what circumstances, different types of token offerings, such as initial coin offerings (ICOs) or security token offerings (STOs), are the optimal way to finance a young firm operating a digital platform. As the paper characterizes optimal token utility and security features, it contributes to the ongoing policy discussion on whether tokens are securities and should be regulated as such.

Chapter 6 "Managing Stablecoins" (joint with Ye Li) develops a model of the issuance and design of stablecoins. We characterize the optimal structure of the liquidity reserves that back the stablecoin and study optimal regulation of stablecoin issuance. Notably, we show that even with over-collateralization, the optimal design of stablecoins features the possibility of price debasement and price volatility that arises when the platform's liquidity reserves are sufficiently low. The paper gives guidance for both practicioners who intend to launch stablecoins and for policy-makers who intend to regulate stablecoins. The paper also discusses the implications of user network effects and transaction data as productive capital on i) the development of token-based platforms featuring stablecoins, ii) the price-stability of cryptocurrencies, and iii) optimal regulation of stablecoins.

In future research, I would like to continue my research on stablecoins by studying the economics and consequences of Central Bank Digital Currency (CBDC) issuance in more detail. Currently, several central banks around the world are actively considering and planning the launch of CBDCs, with the People's Bank of China having drafted pilots on CBDC issuance and being one of the leaders on CBDC issuance. On a broader level, a CBDC can be understood as a deposit account at the central bank that is accessible to retail investors. As a result, by issuing a CBDC, a central bank competes with commercial banks in deposit taking. Changes in commercial banks' deposit taking likely propagate towards commercial banks' lending activities. Because commercial banks are generally more efficient lenders than a central bank and because efficient bank lending is key for the economy, it is important to study the effects of CBDC issuance on bank lending and the economy. In addition, by affecting bank lending, the issuance of CBDC has first-order implications on monetary policy transmission and, therefore, on the effects of quantitative easing (QE). I would like to answer these and related questions in future research.

Finally, I would like to conduct research on more general types of Decentralized Finance (DeFi) applications and platforms. With the rise of blockchain technology and cryptocurrencies, DeFi applications and platforms have become increasingly popular. Interestingly, many DeFi projects rely on tokenization, blockchain technology, and oftentimes make use of the Ethereum blockchain. Broadly, DeFi can be understood as a form of finance without financial intermediaries, such as banks or brokerages. Notable examples of DeFi projects are P2P lending platforms like Bondora, crowdfunding platforms like Kickstarter, or Maker DAO's stablecoin-based lending platform. In future research, I would like to study i) the impact of DeFi on traditional financial firms and their business models, ii) the sustanabiliy and fragility of DeFi platforms, iii) the decentralized governance of DeFi platforms, iv) the design of DeFi applications, and v) the regulation of DeFi.

Bibliography

Abadi, J. and M. Brunnermeier (2018). Blockchain economics. *Working Paper*.

Abel, A. and J. Eberly (2011). How q and cash flow affect investment without frictions: an analytic explanation. *Review of Economic Studies 78*, 1179–1200.

Acharya, V. V., P. DeMarzo, and I. Kremer (2011). Endogenous information flows and the clustering of announcements. *American Economic Review 101*(7), 2955–79.

Aghion, P. and J. Stein (2008). Growth versus margins: destabilizing consequences of giving the market what it wants. *Journal of Finance 63*, 1025–1057.

Aghion, P. and J. Tirole (1994). The management of innovation. *The Quarterly Journal of Economics 109*(4), 1185–1209.

Ai, H. and R. Li (2015). Investment and ceo compensation under limited commitment. *Journal of Financial Economics 116*(3), 452–472.

Albuquerque, R. and H. Hopenhayn (2004). Optimal lending contracts and firm dynamics. *Review of Economic Studies 71*, 285–315.

Armstrong, M., C. Doyle, and J. Vickers (1996). The access pricing problem: A synthesis. *The Journal of Industrial Economics 44*(2), 131–150.

Arrow, K. J. (1965). *Aspects of the Theory of Risk Bearing*. Helsinki: Yrjo Jahnssonin Saatio.

Asker, J., J. Farre-Mensa, and A. Ljunqvist (2015). Corporate investment and stock market listing: a puzzle. *Review of Financial Studies 28*, 342–390.

Asquith, P. and D. W. Mullins (1986). Signalling with dividends, stock repurchases, and equity issues. *Financial Management 15*(3), 27–44.

Athey, S., I. Parashkevov, V. Sarukkai, and J. Xia (2016). Bitcoin pricing, adoption, and usage: Theory and evidence. *Working Paper*.

Baliga, S. and T. Sjöström (1998). Decentralization and collusion. *Journal of Economic Theory 83*(2), 196–232.

Bank for International Settlements (2019). Chapter 3. Big tech in finance: opportunities and risks. Technical report, Bank for International Settlements.

Barton, D., J. Manyika, and S. Williamson (2017). The data: where long-termism pays off. *Harvard Business Review 95*(3), 67–68.

Baumol, W. J. (1952). The transactions demand for cash: An inventory theoretic approach. *The Quarterly Journal of Economics 66*(4), 545–556.

Bebchuk, L. A., A. Cohen, and S. Hirst (2017). The agency problems of institutional investors. *Journal of Economic Perspectives 31*(3), 89–102.

Bergemann, D. and U. Hege (1998). Venture capital financing, moral hazard, and learning. *Journal of Banking & Finance 22*(6-8), 703–735.

Bergemann, D. and U. Hege (2005). The financing of innovation: Learning and stopping. *RAND Journal of Economics*, 719–752.

Bernstein, S. (2015). Does going public affect innovation? *Journal of Finance 70*(1), 1365–1403.

Bertomeu, J., I. Marinovic, S. J. Terry, and F. Varas (2019). The dynamics of concealment. *Working Paper*.

Bettis, J. C., J. Bizjak, J. L. Coles, and S. Kalpathy (2018). Performance-vesting provisions in executive compensation. *Journal of Accounting and Economics 66*(1), 194–221.

Beyer, A., D. A. Cohen, T. Z. Lys, and B. R. Walther (2010). The financial reporting environment: Review of the recent literature. *Journal of Accounting and Economics 50*(2-3), 296–343.

Bhattacharya, S. and P. Pfleiderer (1985). Delegated portfolio management. *Journal of Economic Theory 36*(1), 1–25.

Bhattacharyya, S. and F. Lafontaine (1995). Double-sided moral hazard and the nature of share contracts. *The RAND Journal of Economics*, 761–781.

Biais, B., C. Bisiere, M. Bouvard, and C. Casamatta (2019). The blockchain folk theorem. *The Review of Financial Studies 32*(5), 1662–1715.

Biais, B., T. Mariotti, G. Plantin, and J.-C. Rochet (2007). Dynamic security design: Convergence to continuous time and asset pricing implications. *The Review of Economic Studies 74*(2), 345–390.

Biais, B., T. Mariotti, J.-C. Rochet, and S. Villeneuve (2010). Large risks, limited liability, and dynamic moral hazard. *Econometrica 78*(1), 73–118.

Board, S. and M. Meyer-ter Vehn (2013). Reputation for quality. *Econometrica 81*(6), 2381–2462.

Bolton, P., H. Chen, and N. Wang (2011). A unified theory of tobin's q, corporate investment, financing, and risk management. *The journal of Finance 66*(5), 1545–1578.

Bolton, P. and D. S. Scharfstein (1990). A theory of predation based on agency problems in financial contracting. *American Economic Review 80*(1), 93–106.

Bolton, P., J. Scheinkman, and W. Xiong (2006). Executive compensation and short-termist behavior in speculative markets. *Review of Economic Studies 73*(1), 577–610.

Bolton, P., N. Wang, and J. Yang (2017). Liquidity and risk management: Coordinating investment and compensation policies. forthcoming, *Journal of Finance*.

Bond, P. (2004). Bank and nonbank financial intermediation. *Journal of Finance 59*(6), 2489–2529.

Brainard, L. (2019, October). Digital currencies, stablecoins, and the evolving payments landscape. Speech by Governor Lael Brainard at The Future of Money in the Digital Age, Sponsored by the Peterson Institute for International Economics and Princeton University's Bendheim Center for Finance, Washington, D.C.

Brunnermeier, M. and Y. Sannikov (2014). A macroeconomic model with a financial sector. *American Economic Review 104*, 379–421.

Brunnermeier, M. K., H. James, and J.-P. Landau (2019). The digitalization of money. Technical report, Princeton University.

Buffa, A. M., Q. Liu, and L. White (2020). Optimal delegated contracting. Working paper.

Bullmann, D., J. Klemm, and A. Pinna (2019). In search for stability in crypto-assets: are stablecoins the solution? Occasional Paper Series 230, European Central Bank.

Buterin, V. and V. Griffith (2017). Casper the friendly finality gadget. arXiv preprint arXiv:1710.09437.

Calle, G. and D. B. Zalles (2019). Will businesses ever use stablecoins? white paper, R3 Reports.

Calvo, G. A. and C. M. Reinhart (2002). Fear of floating. *The Quarterly Journal of Economics 117*(2), 379–408.

Cao, S., L. W. Cong, and B. Yang (2019). Auditing and blockchains: Pricing, misstatements, and regulation. *Working Paper*.

Carlson, M., A. Fisher, and R. Giammarino (2006). Corporate investment and asset price dynamics: implications for SEO event-studies and long-run performance. *Journal of Finance 61*, 1009–1034.

Casamatta, C. (2003). Financing and advising: optimal financial contracts with venture capitalists. *The journal of finance 58*(5), 2059–2085.

Catalini, C. and J. S. Gans (2018). Initial coin offerings and the value of crypto tokens. *Working Paper*.

Chang, X., S. Dasgupta, G. Wong, and J. Yao (2014). Cash-flow sensitivities and the allocation of internal cash flow. *Review of Financial Studies 27*(12), 3628–3657.

Chen, L., L. W. Cong, and Y. Xiao (2019). A brief introduction to blockchain economics. *Working Paper*.

Chod, J. and E. Lyandres (2019). A theory of icos: Diversification, agency, and information asymmetry. *Management Science (forthcoming)*.

Chod, J., N. Trichakis, and S. A. Yang (2019). Platform tokenization: Financing, governance, and moral hazard. *Working Paper*.

Christiano, L. J., M. Eichenbaum, and C. L. Evans (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy 113*(1), 1–45.

Cong, L. W. and Z. He (2019). Blockchain disruption and smart contracts. *The Review of Financial Studies 32*(5), 1754–1797.

Cong, L. W., Z. He, and J. Li (2020). Decentralized mining in centralized pools. *Review of Financial Studies forthcoming*.

Cong, L. W., Y. Li, and N. Wang (2020a). Tokenomics and platform finance. *Working Paper*.

Cong, L. W., Y. Li, and N. Wang (2020b). Tokenomics: Dynamic adoption and valuation. *Review of Financial Studies forthcoming*.

Cong, L. W., Y. Li, and N. Wang (2020c). Tokenomics: Dynamic adoption and valuation. *Review of Financial Studies (forthcoming)*.

Cox, J., J. Ingersoll, and S. Ross (1985). An intertemporal general equilibrium model of asset prices. *Econometrica 53*, 363–384.

Cuoco, D. and R. Kaniel (2011). Equilibrium prices in the presence of delegated portfolio management. *Journal of Financial Economics 101*(2), 264–296.

Daley, B. and B. Green (2012). Waiting for news in the market for lemons. *Econometrica 80*(4), 1433–1504.

Daley, B. and B. Green (2016). An information-based theory of time-varying liquidity. *The Journal of Finance 71*(2), 809–870.

Daley, B. and B. S. Green (2019). Bargaining and news. *American Economic Review (forthcoming)*.

Décamps, J.-P., S. Gryglewicz, E. Morellec, and S. Villeneuve (2017). Corporate policies with permanent and transitory shocks. *Review of Financial Studies 30*(1), 162–210.

Décamps, J.-P., T. Mariotti, J.-C. Rochet, and S. Villeneuve (2011). Free cash flow, issuance costs, and stock prices. *The Journal of Finance 66*(5), 1501–1544.

DeMarzo, P. and Z. He (2020). Leverage dynamics without commitment. *Journal of Finance forthcoming*.

DeMarzo, P. M. and M. J. Fishman (2007a). Agency and optimal investment dynamics. *The Review of Financial Studies 20*(1), 151–188.

DeMarzo, P. M. and M. J. Fishman (2007b). Optimal long-term financial contracting. *The Review of Financial Studies 20*(6), 2079–2128.

DeMarzo, P. M., M. J. Fishman, Z. He, and N. Wang (2012). Dynamic agency and the q theory of investment. *The Journal of Finance 67*(6), 2295–2340.

DeMarzo, P. M., D. Livdan, and A. Tchistyi (2014). Risking other people's money: gambling, limited liability, and optimal incentives. Unpublished working paper, Stanford GSB.

DeMarzo, P. M. and Y. Sannikov (2006). Optimal security design and dynamic capital structure in a continuous-time agency model. *The Journal of Finance 61*(6), 2681–2724.

DeMarzo, P. M. and Y. Sannikov (2016). Learning, termination, and payout policy in dynamic incentive contracts. *The Review of Economic Studies 84*(1), 182–236.

DeMarzo, P. M. and B. Urošević (2006). Ownership dynamics and asset pricing with a large shareholder. *Journal of Political Economy 114*(4), 774–815.

Diamond, D. W. (1984). Financial intermediation and delegated monitoring. *The Review of Economic Studies 51*(3), 393–414.

Diamond, D. W. and P. H. Dybvig (1983). Bank runs, deposit insurance, and liquidity. *Journal of Political Economy 91*(3), 401–419.

Diamond, D. W. and R. E. Verrecchia (1991). Disclosure, liquidity, and the cost of capital. *The Journal of Finance 46*(4), 1325–1359.

Dilmé, F. and D. F. Garrett (2019). Residual deterrence. *Journal of the European Economic Association 17*(5), 1654–1686.

Dimson, E., O. Karakaş, and X. Li (2015). Active ownership. *Review of Financial Studies 28*(12), 3225–3268.

Doepke, M. and M. Schneider (2017). Money as a unit of account. *Econometrica 85*(5), 1537–1574.

Donaldson, J. R., G. Piacentino, and A. Thakor (2020). Intermediation variety. *Journal of Finance (forthcoming)*.

Duffie, D. (2019). Digital currencies and fast payment systems: Disruption is coming. working paper, Stanford University Graduate School of Business.

Dumas, B. (1991a). Super contact and related optimality conditions. *Journal of Economic Dynamics and Control 15*(4), 675–685.

Dumas, B. (1991b). Super contact and related optimality conditions. *Journal of Economic Dynamics and Control 15*(4), 675 – 685.

Easley, D., M. O'Hara, and S. Basu (2019). From mining to markets: The evolution of bitcoin transaction fees. *Journal of Financial Economics*.

ECB Crypto-Assets Task Force (2019). Stablecoins: Implications for monetary policy, financial stability, market infrastructure and payments, and banking supervision in the euro area. Occasional Paper Series 247, European Central Bank.

Eckbo, B. E., R. W. Masulis, and O. Norli (2007). Chapter 6 – security offerings. In B. E. Eckbo (Ed.), *Handbook of Corporate Finance: Empirical Corporate Finance*, Volume 1 of *Handbook of Finance Series*, pp. 233–373. Elsevier/North-Holland.

Edmans, A. and X. Gabaix (2016). Executive compensation: A modern primer. *Journal of Economic literature 54*(4), 1232–87.

Edmans, A., X. Gabaix, and D. Jenter (2017). Executive compensation: a survey of theory and evidence. In B. E. Hermalin and M. S. Weisbach (Eds.), *The Handbook of the Economics of Corporate Governance*, pp. 383–539. Amsterdam: Elsevier.

Edmans, A., X. Gabaix, T. Sadzik, and Y. Sannikov (2012). Dynamic ceo compensation. *The Journal of Finance 67*(5), 1603–1647.

Eeckhout, J., N. Persico, and P. E. Todd (2010). A theory of optimal random crackdowns. *American Economic Review 100*(3), 1104–35.

Ewens, M., R. Nanda, and M. Rhodes-Kropf (2018). Cost of experimentation and the evolution of venture capital. *Journal of Financial Economics 128*(3), 422–442.

Fahlenbrach, R. and M. Frattaroli (2019). Ico investors. *Working Paper*.

Fajgelbaum, P. D., E. Schaal, and M. Taschereau-Dumouchel (2017, 05). Uncertainty Traps*. *The Quarterly Journal of Economics 132*(4), 1641–1692.

Fang, L., V. Ivashina, and J. Lerner (2015). The disintermediation of financial markets: Direct investing in private equity. *Journal of Financial Economics 116*(1), 160–178.

Farboodi, M., R. Mihet, T. Philippon, and L. Veldkamp (2019, January). Big data and firm dynamics. Working Paper 25515, National Bureau of Economic Research.

Faure-Grimaud, A., J.-J. Laffont, and D. Martimort (2003). Collusion, delegation and supervision with soft information. *The Review of Economic Studies 70*(2), 253–279.

Feenstra, R. C. (1986). Functional equivalence between liquidity costs and the utility of money. *Journal of Monetary Economics 17*(2), 271 – 291.

Financial Stability Board (2020). Regulation, supervision and oversight of "global stablecoin" arrangements. Technical report, Financial Stability Board.

Fishman, M. J. and K. M. Hagerty (1989). Disclosure decisions by firms and the competition for price efficiency. *The Journal of Finance 44*(3), 633–646.

Fishman, M. J. and K. M. Hagerty (1990). The optimal amount of discretion to allow in disclosure. *The Quarterly Journal of Economics 105*(2), 427–444.

Francis, B., H. Iftekhar, J. Kose, and S. Zenu (2013). Asymmetric benchmarking of pay in firms. *Journal of Corporate Finance 23*, 39–53.

Freeman, S. and F. E. Kydland (2000, December). Monetary aggregates and output. *American Economic Review 90*(5), 1125–1135.

G7 Working Group on Stablecoins (2019). Investigating the impact of global stablecoins. Technical report, Committee on Payments and Market Infrastructures, Bank for International Settlements.

Gale, D. and M. Hellwig (1985). Incentive-compatible debt contracts: The one-period problem. *Review of Economic Studies 52*(4), 647–663.

Galí, J. (2015). *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework and Its Applications Second edition*. Number 10495 in Economics Books. Princeton University Press.

Garvey, G. and T. Milbourn (2006). Asymmetric benchmarking in compensation: executives are rewarded for good luck but not penalized for bad. *Journal of Financial Economics 82*(1), 197–225.

Gertler, M. and N. Kiyotaki (2010). Financial intermediation and credit policy in business cycle analysis. In B. Friedman and M. Woodford (Eds.), *Handbook of Monetary Economics*, pp. 547–599. Amsterdam: Elsevier.

Giannetti, M. and X. Yu (2018). Adapting to radical change: the benefits of short-horizon investors. Unpublished working paper, Stockholm School of Economics.

Glaeser, E. L., B. Sacerdote, and J. A. Scheinkman (1996, 05). Crime and Social Interactions*. *The Quarterly Journal of Economics 111*(2), 507–548.

Goldstein, I., D. Gupta, and R. Sverchkov (2019). Initial coin offerings as a commitment to competition. *Working Paper*.

Goldstein, I. and A. Pauzner (2005). Demand–deposit contracts and the probability of bank runs. *The Journal of Finance 60*(3), 1293–1327.

Gompers, P. A., W. Gornall, S. N. Kaplan, and I. A. Strebulaev (2020). How do venture capitalists make decisions? *Journal of Financial Economics 135*(1), 169–190.

Gopinath, G., E. Boz, C. Casas, F. J. Díez, P.-O. Gourinchas, and M. Plagborg-Møller (2020, March). Dominant currency paradigm. *American Economic Review 110*(3), 677–719.

Gornall, W. and I. A. Strebulaev (2020). Squaring venture capital valuations with reality. *Journal of Financial Economics 135*(1), 120–143.

Green, B. and C. R. Taylor (2016). Breakthroughs, deadlines, and self-reported progress: contracting for multistage projects. *American Economic Review 106*(12), 3660–99.

Grenadier, S. R. and A. Malenko (2011). Real options signaling games with applications to corporate finance. *The Review of Financial Studies 24*(12), 3993–4036.

Gryglewicz, S. and B. Hartman-Glaser (2019). Investment timing and incentive costs. *Review of Financial Studies, forthcoming*.

Gryglewicz, S., L. Mancini, E. Morellec, E. J. Schroth, and P. Valta (2020). Understanding cash flow risk. Unpublished working paper, Swiss Finance Institute.

Gryglewicz, S. and S. Mayer (2019). Delegated monitoring and contracting. *Available at SSRN 3175528*.

Gryglewicz, S., S. Mayer, and E. Morellec (2020). Agency conflicts and short-versus long-termism in corporate policies. *Journal of Financial Economics 136*(3), 718–742.

Gryglewicz, S., S. Mayer, and E. Morellec (2021). Optimal financing with tokens. *Journal of Financial Economics (forthcoming)*.

Guerrieri, V. and P. Kondor (2012). Fund managers, career concerns, and asset price volatility. *American Economic Review 102*(5), 1986–2017.

Gutierrez, G. and T. Philippon (2017). Investment-less growth: an empirical investigation. Unpublished working paper, New York University.

Guttman, I., I. Kremer, and A. Skrzypacz (2014). Not only what but also when: A theory of dynamic voluntary disclosure. *American Economic Review 104*(8), 2400–2420.

Hackbarth, D., A. Rivera, and T.-Y. Wong (2018). Optimal short-termism. Unpublished working paper, Boston University.

Halac, a. and I. Kremer (2019). Experimenting with career concerns. *American Economic Journal Microeconomics (forthcoming)*.

Halac, M. and A. Prat (2016, October). Managerial attention and worker performance. *American Economic Review 106*(10), 3104–32.

Hartman-Glaser, B., T. Piskorski, and A. Tchistyi (2012). Optimal securitization with moral hazard. *Journal of Financial Economics 104*(1), 186–202.

Hautsch, N., C. Scheuch, and S. Voigt (2019). Limits to arbitrage in markets with stochastic settlement latency. *Working Paper*.

He, Z. (2009a). Optimal executive compensation when firm size follows geometric brownian motion. *The Review of Financial Studies 22*(2), 859–892.

He, Z. (2009b). Optimal executive compensation when firm size follows geometric Brownian motion. *Review of Financial Studies 22*(2), 859–892.

He, Z. (2011). A model of dynamic compensation and capital structure. *Journal of Financial Economics 100*(2), 351–366.

He, Z. (2012). Dynamic compensation contracts with private savings. *The Review of Financial Studies 25*(5), 1494–1549.

He, Z. and A. Krishnamurthy (2011). A model of capital and crises. *The Review of Economic Studies 79*(2), 735–777.

He, Z. and A. Krishnamurthy (2013). Intermediary asset pricing. *American Economic Review 103*(2), 732–70.

He, Z., B. Wei, J. Yu, and F. Gao (2017). Optimal long-term contracting with learning. *The Review of Financial Studies 30*(6), 2006–2065.

Hellwig, M. F. (2000). Financial intermediation with risk aversion. *Review of Economic Studies 67*(4), 719–742.

Hinzen, F. J., K. John, and F. Saleh (2019). Bitcoin's fatal flaw: The limited adoption problem. *Working Paper*.

Hoffmann, F., R. Inderst, and M. Opp (2019). Only time will tell: A theory of deferred compensation and its regulation. *Working Paper*.

Hoffmann, F. and S. Pfeil (2010). Reward for luck in a dynamic agency model. *The Review of Financial Studies 23*(9), 3329–3345.

Hoffmann, F. and S. Pfeil (2021). Dynamic multitasking and managerial investment incentives. *Journal of Financial Economics (forthcoming)*.

Holden, R. and A. Malani (2019). The ico paradox: Transactions costs, token velocity, and token value. *Working Paper*.

Holmstrom, B. (1989). Agency costs and innovation. *Journal of Economic Behavior & Organization 12*(3), 305–327.

Holmstrom, B. and P. Milgrom (1987). Aggregation and linearity in the provision of intertemporal incentives. *Econometrica: Journal of the Econometric Society*, 303–328.

Holmstrom, B. and J. Tirole (1997). Financial intermediation, loanable funds, and the real sector. *the Quarterly Journal of economics 112*(3), 663–691.

Hori, K. and H. Osano (2013). Managerial incentives and the role of advisors in the continuous-time agency model. *The Review of Financial Studies 26*(10), 2620–2647.

Howell, S. T., M. Niessner, and D. Yermack (2020). Initial coin offerings: Financing growth with cryptocurrency token sales. *The Review of Financial Studies 33*(9), 3925–3974.

Huberman, G., J. Leshno, and C. C. Moallemi (2019). An economic analysis of the bitcoin payment system. *Working Paper*.

Hugonnier, J., S. Malamud, and E. Morellec (2015a). Capital supply uncertainty, cash holdings, and investment. *Review of Financial Studies 28*(2), 391–445.

Hugonnier, J., S. Malamud, and E. Morellec (2015b). Capital supply uncertainty, cash holdings, and investment. *The Review of Financial Studies 28*(2), 391–445.

Inderst, R. and H. M. Mueller (2010). Ceo replacement under private information. *The Review of Financial Studies 23*(8), 2935–2969.

Innes, R. D. (1990). Limited liability and incentive contracting with ex-ante action choices. *Journal of Economic Theory 52*(1), 45–67.

Jones, C. I. and C. Tonetti (2020, September). Nonrivalry and the economics of data. *American Economic Review 110*(9), 2819–58.

Jovanovic, B. (1982). Truthful disclosure of information. *The Bell Journal of Economics*.

Kaniel, R. and P. Kondor (2012). The delegated Lucas tree. *Review of Financial Studies 26*(4), 929–984.

Kaplan, S. N. and P. Strömberg (2003). Financial contracting theory meets the real world: An empirical analysis of venture capital contracts. *The review of economic studies 70*(2), 281–315.

Kaplan, S. N. and P. Strömberg (2009). Leveraged buyouts and private equity. *Journal of Economic Perspectives 23*(1), 121–46.

Kaplan, S. N. and P. E. Strömberg (2004). Characteristics, contracts, and actions: Evidence from venture capitalist analyses. *The Journal of Finance 59*(5), 2177–2210.

Kerr, W. R. and R. Nanda (2015). Financing innovation. *Annual Review of Financial Economics 7*, 445–462.

Kerr, W. R., R. Nanda, and M. Rhodes-Kropf (2014). Entrepreneurship as experimentation. *Journal of Economic Perspectives 28*(3), 25–48.

Kogan, L. (2004). Asset prices and real investment. *Journal of Financial Economics 73*, 411–432.

Laffont, J.-J. and J. Tirole (1994). Access pricing and competition. *European Economic Review 38*(9), 1673 – 1710.

Lazear, E. P. (2006). Speeding, terrorism, and teaching to the test. *The Quarterly Journal of Economics 121*(3), 1029–1061.

Lee, D., H. Shin, and R. Stulz (2018). Why does capital no longer flow more to the industries with the best growth opportunities. Unpublished working paper, Ohio State University.

Lee, I., S. Lochhead, J. Ritter, and Q. Zhao (1996). The costs of raising capital. *Journal of Financial Research 19*(1), 59–74.

Lee, J. and C. A. Parlour (2018). Crowdfunding, initial coin offerings, and consumer surplus. *Working Paper*.

Leland, H. (1994). Corporate debt value, bond covenants, and optimal capital structure. *Journal of Finance 49*(4), 1213–1252.

Lerner, J. and R. Nanda (2020). Venture capital's role in financing innovation: What we know and how much we still need to learn. *Journal of Economic Perspectives 34*(3), 237–61.

Li, J. and W. Mann (2019). Initial coin offerings: Current research and future directions. *Palgrave-MacMillan Handbook of Alternative Finance*.

Li, J. and W. Mann (2020). Digital tokens and platform building. *Working Paper*.

Li, Y. and S. Mayer (2020). Managing stablecoins: Optimal strategies, regulation, and transaction data as productive capital. *Fisher College of Business Working Paper* (2020-03), 030.

Libra Association (2020). White Paper: v2.0. Technical Report April 16, Libra Association.

Liu, Y. and A. Tsyvinski (2019). Risks and returns of cryptocurrency. *Working Paper*.

Ljungqvist, L. and T. J. Sargent (2004). *Recursive Macroeconomic Theory, 2nd Edition*, Volume 1 of *MIT Press Books*. The MIT Press.

Lucas, R. E. and J. P. Nicolini (2015). On the stability of money demand. *Journal of Monetary Economics 73*, 48 – 65.

Lyandres, E. (2019). Product market competition with crypto tokens and smart contracts. *Working Paper*.

Lyandres, E., B. Palazzo, and D. Rabetti (2019). Do tokens behave like securities? an anatomy of initial coin offerings. *Working Paper*.

Macho-Stadler, I. and J. D. Pérez-Castrillo (1998). Centralized and decentralized contracts in a moral hazard environment. *The Journal of Industrial Economics 46*(4), 489–510.

Makarov, I. and A. Schoar (2019). Trading and arbitrage in cryptocurrency markets. *Journal of Financial Economics*.

Malenko, A. (2019a, 08). Optimal dynamic capital budgeting. *Review of Economic Studies 86*(4), 1747–1778.

Malenko, A. (2019b, 08). Optimal dynamic capital budgeting. *Review of Economic Studies 86*(4), 1747–1778.

Malinova, K. and A. Park (2017). Market design with blockchain technology. *Working Paper*.

Manso, G. (2011). Motivating innovation. *The Journal of Finance 66*(5), 1823–1860.

Marinovic, I. and M. Szydlowski (2019). Monitor reputation and transparency. *Working paper*.

Marinovic, I. and F. Varas (2016). No news is good news: Voluntary disclosure in the face of litigation. *The RAND Journal of Economics 47*(4), 822–856.

Marinovic, I. and F. Varas (2019a). Ceo horizon, optimal pay duration, and the escalation of short-termism. *The Journal of Finance 74*(4), 2011–2053.

Marinovic, I. and F. Varas (2019b). CEO horizon, optimal pay duration, and the escalation of short-termism. *Journal of Finance 74*(4), 2011–2053.

Mason, R. and J. Välimäki (2015). Getting it done: dynamic incentives to complete a project. *Journal of the European Economic Association 13*(1), 62–97.

Mayer, S. (2020a). Financing breakthroughs under failure risk. *Available at SSRN 3320502*.

Mayer, S. (2020b). Token-based platforms and speculators. *Available at SSRN 3471977*.

Melumad, N. D., D. Mookherjee, and S. Reichelstein (1995). Hierarchical decentralization of incentive contracts. *The RAND Journal of Economics*, 654–672.

Metrick, A. and A. Yasuda (2010). The economics of private equity funds. *Review of Financial Studies 23*(6), 2303–2341.

Miao, J. and A. Rivera (2016). Robust contracts in continuous time. *Econometrica 84*, 1405–1440.

Mookherjee, D. (2013). Incentives in hierarchies. In R. S. Gibbons and J. Roberts (Eds.), *The Handbook of Organizational Economics*, pp. 764–798. Princeton University Press.

Mookherjee, D. and M. Tsumagari (2004). The organization of supplier networks: effects of delegation and intermediation. *Econometrica 72*(4), 1179–1219.

Morellec, E. and N. Schürhoff (2011). Corporate investment and financing under asymmetric information. *Journal of financial Economics 99*(2), 262–288.

Morris, S. and H. S. Shin (1998). Unique equilibrium in a model of self-fulfilling currency attacks. *The American Economic Review 88*(3), 587–597.

Nagel, S. (2016, 07). The Liquidity Premium of Near-Money Assets. *The Quarterly Journal of Economics 131*(4), 1927–1971.

Nanda, R. and M. Rhodes-Kropf (2013). Investment cycles and startup innovation. *Journal of Financial Economics 110*(2), 403–418.

Nanda, R. and M. Rhodes-Kropf (2017). Financing risk and innovation. *Management Science 63*(4), 901–918.

Nanda, R., K. Younge, and L. Fleming (2014). Innovation and entrepreneurship in renewable energy. In *The changing frontier: Rethinking science and innovation policy*, pp. 199–232. University of Chicago Press.

Narayanan, M. (1985). Managerial incentives for short-term results. *Journal of Finance 40*(6), 1469–1484.

Ordoñez, G. (2013). The asymmetric effects of financial frictions. *Journal of Political Economy 121*(5), 844–895.

Orlov, D. (2018). Optimal design of internal disclosure. *Working Paper*.

Ou-Yang, H. (2003). Optimal contracts in a continuous-time delegated portfolio management problem. *The Review of Financial Studies 16*(1), 173–208.

Pagnotta, E. (2020). Bitcoin as decentralized money: Prices, mining rewards, and network security. *Working Paper*.

Piskorski, T. and M. M. Westerfield (2016). Optimal dynamic contracts with moral hazard and costly monitoring. *Journal of Economic Theory 166*, 242–281.

Poterba, J. M. and J. J. Rotemberg (1986, 1986). Money in the utility function: An empirical implementation. Working Paper 1796, National Bureau of Economic Research.

Prat, J. and B. Walter (2019). An equilibrium model of the market for bitcoin mining. *Working Paper*.

Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica 32*(1/2), 122–136.

Rahman, D. (2012). But who will monitor the monitor? *American Economic Review 102*(6), 2767–97.

Rampini, A. and S. Viswanathan (2013). Collateral and capital structure. *Journal of Financial Economics 109*, 466–492.

Rochet, J.-C. and J. Tirole (2003). Platform competition in two-sided markets. *Journal of the European Economic Association 1*(4), 990–1029.

Rochet, J.-C. and J. Tirole (2006). Two-sided markets: A progress report. *The RAND Journal of Economics 37*(3), 645–667.

Rysman, M. (2009). The economics of two-sided markets. *The Journal of Economic Perspectives 23*(3), 125–143.

Saleh, F. (2019). Blockchain without waste: Proof-of-stake. *Review of Financial Studies (forthcoming)*.

Sannikov, Y. (2008). A continuous-time version of the principal-agent problem. *The Review of Economic Studies 75*(3), 957–984.

Sannikov, Y. (2014). Moral hazard and long-run incentives. *Working paper, Stanford University*.

Schmidt, K. M. (2003). Convertible securities and venture capital finance. *The Journal of Finance 58*(3), 1139–1166.

Shafir, E., P. Diamond, and A. Tversky (1997, 05). Money Illusion. *The Quarterly Journal of Economics 112*(2), 341–374.

Shreve, S. (2004). *Stochastic calculus for finance II: continuous-time models*. Berlin: Springer Verlag.

Sockin, M. and W. Xiong (2020). A model of cryptocurrencies. *Working Paper*.

Stein, J. (1989). Efficient capital markets, inefficient firms: a model of myopic corporate behavior. *Quarterly Journal of Economics 104*(4), 655–669.

Strausz, R. (1997). Delegation of monitoring in a principal-agent relationship. *The Review of Economic Studies 64*(3), 337–357.

Strebulaev, I. (2007). Do tests of capital structure theory mean what they say? *Journal of Finance 62*(4), 1747–1787.

Stulz, R. M. (2019). Fintech, bigtech, and the future of banks. *Journal of Applied Corporate Finance 31*(4), 86–97.

Szydlowski, M. (2015). Incentives, project choice, and dynamic multitasking. *Working Paper*.

Szydlowski, M. (2019). Optimal financing and disclosure. *Management Science (forthcoming)*.

Thakor, R. (2018). A theory of efficient short-termism. Unpublished working paper, University of Minnesota.

Tobin, J. (1956). The interest-elasticity of transactions demand for cash. *The Review of Economics and Statistics 38*(3), 241–247.

Townsend, R. M. (1979). Optimal contracts and competitive markets with costly state verification. *Journal of Economic Theory 21*(2), 265–293.

Varas, F. (2017). Managerial short-termism, turnover policy, and the dynamics of incentives. *The Review of Financial Studies*.

Varas, F., I. Marinovic, and A. Skrzypacz (2020a). Random inspections and periodic reviews: Optimal dynamic monitoring. *Review of Economic Studies (forthcoming)*.

Varas, F., I. Marinovic, and A. Skrzypacz (2020b). Random inspections and periodic reviews: Optimal dynamic monitoring. *The Review of Economic Studies 87*(6), 2893–2937.

Veldkamp, L. and C. Chung (2019). Data and the aggregate economy. Working paper, Columbia University.

Veldkamp, L. L. (2005). Slow boom, sudden crash. *Journal of Economic Theory 124*(2), 230 – 257. Learning and Bounded Rationality.

Von Thadden, E.-L. (1995). Long-term contracts, short-term investment and monitoring. *Review of Economic Studies 62*(6), 557–575.

Walsh, C. E. (2003). *Monetary Theory and Policy, 2nd Edition* (2 ed.). The MIT Press.

Weitzman, M. L. (1998, 05). Recombinant Growth. *The Quarterly Journal of Economics 113*(2), 331–360.

Zhu, J. (2018). Myopic agency. *Review of Economic Studies 85*, 1352–1388.

Zhu, J. Y. (2012). Optimal contracts with shirking. *Review of Economic Studies 80*(2), 812–839.