On the Effects of CEO Compensation

ISBN: 978 90 3610 531 6

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

This book is no. **726** 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.

On the Effects of CEO Compensation

Over de effecten van CEO compensatie

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the
Rector Magnificus

Prof. dr. R.C.M.E. Engels

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

Thursday 1 November 2018 at 13:30 hrs

by

Yuhao Zhu
born in Anqing, China

Erasmus University Rotterdam

Doctoral committee

Doctoral dissertation supervisor:

Prof. dr. I. Dittmann

Other members:

Prof. dr. P. G. J. Roosenboom
Prof. dr. O. G. Spalt
Dr. S. Gryglewicz

Co-supervisor:

Prof. dr. S. van Bekkum

Abstract

The separation of ownership and control within companies cause agency problems. Executive compensation is a tool to align the interests between shareholders and top executives. This thesis studies the potential effects of executive compensation packages on the firm value and other factors. I have three main findings in this thesis. In the first paper, I show that when CEOs are probability weighting, the optimal contracts are convex. This explains the existence of option components in CEO's compensation packages. In the second paper, I find that the wage of the employee is increasing in the CEO pay. This relationship is found both across firms and across time. I ascribe this relationship to the behindness aversion of workers. The result suggests CEO compensation incurs extra costs to the firms. In the third paper, I show that firms with low wage gaps between CEO and workers are overpriced on the stock market. The effect should be even stronger in the presence of inequality-averse investors. This finding suggests that investors do trade on the pay inequality, and show that the mis-pricing comes from the overvaluation of low wage gap stocks.

De scheiding van eigendom en controle binnen bedrijven veroorzaakt bureauproblemen. CEO compensatie is een hulpmiddel om de belangen af te stemmen tussen aandeelhouders en topbestuurders. Dit proefschrift bestudeert de mogelijke effecten van een vergoedingspakket voor bestuurders op de bedrijfswaarde en andere factoren. Ik heb drie belangrijke bevindingen in dit proefschrift. Ten eerste laat ik zien dat wanneer CEO's kansen wegen, de optimale contracten convex zijn. Dit verklaart het bestaan van optiecomponenten in de beloningspakketten van de CEO. Ten tweede vind ik dat het loon van de werknemer stijgt in de CEO-beloning. Deze relatie is bewezen, zowel binnen bedrijven als in de loop van de tijd. Ik schrijf deze relatie toe aan de behindness aversie van arbeiders. Het resultaat suggereert dat de vergoeding van de CEO voor de bedrijven extra kosten met zich mee kan brengen. Ten derde laat ik zien dat bedrijven met lage loonverschillen tussen CEO en werknemers te duur zijn op de aandelenmarkt. Het effect zou zelfs sterker moeten zijn in de aanwezigheid van ongelijkheidsaverse investeerders. Deze bevinding suggereert dat beleggers wel handel drijven in de loonongelijkheid en laten zien dat de verkeerde prijsstelling voortkomt uit de overwaardering van aandelen met een laag loonverschil.

Acknowledgement

I would like to express my gratitude to my supervisor, Prof. dr. Ingolf Dittmann, who guided me through my Ph.D. study with great wisdom, patience, and kindness.

I would also like to thank those who provided me with many valuable feedbacks and comments for my thesis.

Special gratitude goes to my family, my friends and my beloved one. Thanks for all your encouragement that enables me to shoulder my way through the difficulties.

Contents

Chapter I

Introduction

The separation of ownership and control within companies causes agency problems. The top executives may increase their own benefits at the costs of shareholders. Executive compensation is a tool to align the interests of shareholders and top executives. Studying executive compensation is important to firms. On the one hand, an executive compensation package can give incentives to executives, which directly affects the value of the firm. Not only the total amount but also the structure of the compensation package plays a role in providing incentives. A well-designed executive compensation contract can provide high incentives at low costs. On the other hand, executive compensation may also affect the decisions of other people, e.g., rank-and-file workers or investors. The choices of these people can affect the firm value. Thus, the executive compensation package indirectly affects the firm value.

The question then arises: what are the potential effects of the executive compensation packages on the firm value and other factors? The answer to this question can reveal the direct and indirect effects of CEO compensation packages on the firm value. It provides guidance to shareholders for determining the CEO compensation. Previously literature provides abundant studies on the optimal executive compensation contract and its effects on firm value. However, some phenomena are not well explained by existing theories. For example, why stock option is a component of the CEO compensation package? Why high CEO compensation causes strong opposition from rank-and-file employees? Is the wage gap between CEO and workers correctly priced by investors in the stock market? In my thesis, I dive into these questions. By incorporating behavioral theories into the principal-agent problem, I give answers to the questions that are not answered by the models introduced in the previous literature. The predictions are also supported by empirical tests.

In Chapter II, I discuss the shape of the optimal CEO compensation contract. A typical CEO compensation package contains various components, e.g., fixed salary, bonus, shares, and stock options. Whether stock options should be a part of the optimal CEO compensation package remains debatable. If the CEO is risk-averse, her compensation packages should contain little fixed salary and no options (Dittmann and Maug, 2007). Thus, alternative theories should be used to rationalize the convexity of the CEO compensation contract.

I analyze a principal-agent model where the CEO is probability weighting. I approximate the probability weighting function with parameters that shift the normal distribution. Using this approximation of the probability weighting function, I show that the optimal general contracts exhibit convexity when firm performance is high. The model explains the considerable number of options in the observed CEO compensation packages. To see whether my model fits the data well, I calibrate the model with a wide range of parameters using the observed U.S. CEO compensation contracts. I show that the probability weighting model performs better in explaining the shape of the observed contracts than the traditional CRRA model.

The findings suggest that shareholders exploit probability weighting to provide cheap incentives that encourage CEOs to exert more effort. It provides an alternative theory for convexity in contracts. My paper can act as a complement to the literature that tries to rationalize the positive options pay in observed CEO contracts, e.g., the loss-aversion theory (Dittmann, Maug, and Spalt, 2010) and risk-seeking incentive theory (Dittmann, Yu, and Zhang, 2017).

In Chapter III, we discuss the effect of CEO compensation on the employees' pay, and the costs associated with this. Some strongest opposition against high CEO compensation comes from normal employees. This phenomenon cannot be rationalized with traditional theories because workers should only care about their own wage. A potential explanation is that workers are behindness averse, i.e., they suffer dis-utility from the wage gap between the CEO pay and their own pay. This behavioral pattern increases the labor costs of the firm and influence the design of the CEO compensation package.

We establish a principal-agent model where the principal designs a contract with two agents: the CEO and the employee who is behindness averse. We find that the wage of the employee is increasing in the CEO pay. This relationship is found, both across firms and across time, by statistical testing on a matched CEO-employee panel data set for German firms. To alleviate the endogeneity problems, we use the difference-in-difference setting. We find that the workers receive a significant increase in wage when CEO compensation is made public for the first time. The implication of the results is that the CEO compensation can affect the firm value indirectly. Envy of workers associated with high CEO compensation brings extra costs to the firms.

The findings in this chapter contribute to the empirical studies which examine the relation between CEO compensation and employee wages or productivity. This study is the first to show that there exists a positive relation between CEO and rank-and-file employee pay and ascribes this relation to the behindness aversion of employees.

In Chapter IV, we move further into the stock market and see how CEO-workers wage gap affects the preferences and choices of investors. Despite populist anger towards high CEO compensation, the literature shows that the larger wage gap reflects higher CEO skills (Faleye, Reis, and Venkateswaran, 2013). However, this information is not correctly priced in the stock market. Mueller, Ouimet, and Simintzi (2017) show that stocks with low pay inequality yield negative risk-adjusted returns.

We set up an asset pricing model with noise traders and short-sales constraints, in which the optimal wage gap between the CEO and rank-and-file workers increases with managerial skills. In equilibrium, we show that firms with low wage gaps should be overpriced, and the effect should be even stronger in the presence of inequality-averse investors. To provide empirical evidence, we use the data set of German firms. We find that a long-short portfolio of stocks with high and low wage gaps yields positive and robust risk-adjusted returns.

The findings in this chapter confirm the previous literature that investors do trade on the pay inequality and show that the mispricing comes from the overvaluation of low wage gap stocks. Our findings also contribute to the recent literature that studies the impact of values on investor behavior. Previous research shows that investors consider nonpecuniary factors in their trading strategies. The findings in this chapter provide evidence that investors, much like the general public, dislike pay inequality within firms.

Other researchers also contribute to the completion of this thesis. Chapter III is a joint work with Ingolf Dittmann and Christoph Schneider. Chapter IV is a joint work with Ingolf Dittmann and Maurizio Montone. I would like to express my sincere thanks to them.

Chapter II

Probability weighting CEOs and optimal contracts

1. Introduction

In this paper, I analyze a principal-agent model where the CEO is probability weighting. I approximate the probability weighting and rank-dependent expected utility model (Quiggin, 1982; Tversky and Kahneman, 1992) with parameters that shift the shape of normal distributions. I show that, using this approximation of the probability weighting function, the optimal general contracts exhibit convexity when firm performance is high. This theoretically explains considerable option components in CEO's compensation packages. To see whether my model predicts the observed contracts well, I calibrate the model with a wide range of parameters using the observed U.S. CEO contracts. I show that the probability weighting model can explain the shape of the observed contracts better than the normal constant relative risk aversion (CRRA) model without probability weighting.

A typical CEO compensation package includes multiple components, e.g., fixed salary, bonus, shares, stock options, and other long-term incentives. The realization of incentive pay is contingent on the future performance of the firms. Incentive pay not only incentivizes the CEOs to exert effort to increase the stock prices of their firms, but also lead to a convex shape of the contracts. Whether a convex contract structure is optimal remains debatable. Dittmann and Maug (2007) solve and calibrate a standard principal-agent model with CRRA agents using observed U.S. CEO contracts. Their solution is a general contract with concave shape. Therefore, Dittmann and Maug (2007) find that neither positive option holding nor positive fixed salary is a part of the optimal scheme. A reason for this concavity is driven by the decreasing marginal utility. It becomes inefficient to keep the CEO pay sensitive to performance at high levels of firm value.

To explain the difference between observed contracts and theoretical optimal contracts, previous literature adopts various behavioral models to rationalize the convexity of the CEO contracts. Dittmann, Maug, and Spalt (2010) incorporate the loss-aversion into the model. They find that the loss-aversion model can better explain the positive option holding than the traditional CRRA model. This is because the optimal contracts with the loss-aversion model are locally convex around the reference point. Dittmann, Yu, and Zhang (2017) improve the CRRA model with the risk-seeking incentives. They find that convex contracts can provide incentives to CEOs to implement projects that are of higher risk. However, the analytical optimal general contracts in these models are only locally convex. That means that when the firm performance becomes large enough, the theoretical optimal general contract becomes concave. Thus, these models have explanatory power for the shapes of the observed CEO compensation contracts only when firm performance is not too high. Other behavioral models are also used to exploit the shape of the contracts for different types of agents. Otto (2014) finds that when CEOs are optimistic, they tend to receive less options and bonus.

The paper by Spalt (2013), who shows that probability weighting model can explain the employee stock option plan, can shed light on the convexity of CEO contracts. Although the study is on rank-and-file employees, the idea can be borrowed to the study of CEO compensation contracts. The probability weighting means that people tend to overweight the probability of the extreme outcomes. It results in a different preference than traditional CRRA. When CEO exhibits probability weighting in her preference, this trait can have two effects on her pay structure. The first effect is the income effect. Options protect the CEO from bad outcomesand benefit the CEO from the good outcomes. Therefore, probability weighting CEOs find options more valuable than its true expected value. The firms can then substitute fixed salary with options of less expected value. This reduces the expected costs for the firm. The second effect is the incentive effect. When a CEO attaches a higher probability to very bad or very good outcomes, her marginal effort can then increase the probability of extreme outcomes more than normal. Since she is protected in bad outcomes with options, she tends to exert more efforts in her work. This effort increases the expected value of the firm.

In this paper, I suggest a new approach to explaining the convexity of the observed CEO compensation contracts. I introduce the probability weighting into the standard Hölmstrom (1979) principal-agent model, and use the sigma-mu transformation (changing the shape of a normal distribution) to approximate the probability weighting function. The paper makes three main contributions.

A potential challenge of incorporating the probability weighting feature into the model of CEO compensation contract is that reaching the closed-form solution for the optimal contract is almost impossible. The probability weighting feature transforms the cumulative distribution function (CDF) of the original distribution to a new one. This means that the new distribution cannot be described by any typical distribution function. As the first contribution, I find that when stock returns follow a normal distribution, probability weighting transforms the distribution into a similar normal distribution with a different set of parameters, namely, σ and μ. This means that the probability weighting feature can be approximated by transforming the parameters of the normal distribution. For each probability weighting parameter δ, I can always find a set of parameters η_s and η_m that transforms the original normal distribution into a new normal distribution with the similar shape. This sigma-mu transformation helps me to reach the closed-form solution for the optimal general contract.

I establish a principal-agent model with a risk-neutral principal and a risk-averse and probability weighting agent. After solving this model, I show that the optimal general contract is convex even when firm performance is very high. The shape is different from the optimal contracts listed in previous literature, which are concave when firm performance is high. The optimal contract of my model gives theoretical evidence for the convexity of CEO compensation contracts.

After providing theoretical evidence for convex CEO compensation contracts, I continue with calibrating the probability weighting model with observed contracts of U.S. CEOs. I numerically solve for the optimal piecewise linear contract using the observed contracts for a wide range of parameters, i.e., combinations of η_s and η_m. The third contribution of my paper is the empirical evidence that the probability weighting model performs better in predicting positive option holding and positive fixed salary than the CRRA model.

My finding suggests that shareholders exploit probability weighting to provide cheap incentives that encourage CEOs to exert more effort. It provides an alternative theory for convexity in contracts. My paper can act as a complement to the literature that tries to rationalize the existence of option components in the observed CEO contracts, e.g., the loss-aversion theory (Dittmann, Maug, and Spalt, 2010) and risk-seeking incentive theory (Dittmann, Yu, and Zhang, 2017).

My results can also be linked to a wide range of behavioral finance literature on CEO compensation. For example, the probability weighting can be linked with CEO overconfidence. The effect of probability weighting on the normal distribution is equivalent to transforming the mean and volatility to higher values. This is related to the theories on overconfidence. Previous literature defines that a manager is "optimistic" if she thinks that the future average performance is higher than the true mean. Malmendier and Tate (2005) find that overconfident CEOs tend to postpone the exercising of their options. Optimism can be linked to a positive η_m in my paper. On the other hand, a manager is mis-calibrated or overprecise if she underestimates the volatility of the future return (Ben-David, Graham, and Harvey, 2013). This effect is similar to an $\eta_s < 1$ in my paper. My results suggest that probability weighting effect may exceed the mis-calibration effect in shaping the CEO compensation contracts.

The paper proceeds as follows. Section 2 introduces the model, the sigma-mu transformation function, and the optimal general contract. Sections 3 introduces the calibration strategy. Section 4 summarizes the data set that is used for the calibration. Section 5 provides the empirical results of the calibration. Section 6 concludes.

2. Model

The baseline model is the principal-agent model introduced by Hölmstrom (1979) with the hidden efforts. The principal (firm) is risk-neutral while the agent (CEO) is risk-averse. There are two stages in the model: the starting stage at time 0, and the paying stage at time T. At time 0, the contract is signed between the firm and the CEO. According to the contract, the CEO receives certain payment from the firm at time T. The payment W_T at time T is contingent on the performance of the firm at time T. The CEO can choose the effort level e that will influence the firm's intrinsic value.

2.1 Production function

P_0 is the intrinsic value of the firm at time 0. It can be seen as the expected discounted future cash flow of the firm evaluated at the stock market. The intrinsic value is affected by the effort level e of the CEO. When CEO exerts more effort, the income from future operation increases and the firm value P_0 increases. The stock return of the firm at time T follows a normal distribution, which means that the future value of the firm follows the log-normal distribution. The firm value at time T is equal to the firm value at time 0 multiplied by a stochastic factor that follows the log-normal distribution. Denote r_f as the risk-free rate, d as the dividend rate, σ^2 as the yearly volatility during between time 0 and time T. The production function of the firm at a fixed effort level e is given by:

$$\begin{aligned}\widetilde{P_T} =& P_0(e)\exp\left\{(r_f-d)T-\frac{\sigma^2}{2}T+\widetilde{u}\sigma\sqrt{T}\right\}\\ =& \exp\{\mu+\widetilde{u}\sigma\},\end{aligned} \tag{1}$$

where $\widetilde{u} \sim \mathcal{N}(0,1)$, and $\mu = \ln P_0(e) + (r_f - d)T - \frac{\sigma^2}{2}T$. The firm value at time 0 should be the principal's unconditional expected long-term firm value $\widetilde{P_T}$ discounted by the risk-free rate r and the dividend rate d. That is, $P_0(e) = \mathbb{E}\left[\exp\{-(r_f-d)T\}\widetilde{P_T}\right]$. Throughout this paper, I put a "tilde" on all random variables, e.g., $\widetilde{P_T}$, to distinguish them from normal variables[1].

2.2 Utility function

I assume throughout the paper that the principal (shareholder) is risk-neutral, while the agent (CEO) is risk-averse. I use the CRRA utility function[2] for the CEO, that is:

$$V(W) = \begin{cases} \frac{W^{1-\gamma}-1}{1-\gamma}, & \text{if } \gamma \neq 1 \\ \ln(W), & \text{if } \gamma = 1, \end{cases} \tag{2}$$

where W is the wealth of the CEO at the time of evaluation. The aggregated utility function is the utility derived from the personal wealth deducted by the loss of utility due to the cost of efforts. That is:

$$U(W) = V(W) - C(e) = \begin{cases} \frac{W^{1-\gamma}-1}{1-\gamma} - C(e), & \text{if } \gamma \neq 1 \\ \ln(W) - C(e), & \text{if } \gamma = 1, \end{cases}$$

where $C(e)$ is the cost function of the effort. The functional form is not explicitly known, but is increasing and convex, i.e., $\frac{dC(e)}{de} > 0$ and $\frac{d^2C(e)}{de^2} > 0$.

2.3 Principal-agent problem

I extend the Hölmstrom (1979) principal-agent model by introducing the probability weighting feature of the CEO. Namely, she assigns the original probability with decision weights. This feature means that firm and the CEO have different views towards future outcomes. I assume that the shareholders perceive the correct probability distribution of firm performance, but the CEO has a biased, subjective probability distribution due to probability weighting. Therefore, The CEO attaches a probability, different from that of shareholders, to each possible future outcome.

I use $\mathbb{E}$ to stand for the expectation of the principal, and use $\mathbb{E}_A$ to stand for the subjective

[1]The differentiation between random variable and a normal variable is valuable when I want to define cumulative probability distribution function. For example, $\Pr\left\{\widetilde{P_T} \leqslant P_T\right\} = F(P_T)$. It causes no misunderstanding.

[2]The original CRRA utility function for the case when $\gamma = 1$ is $V(W) = \frac{W^{1-\gamma}-1}{1-\gamma}$. I use $V(W) = \frac{W^{1-\gamma}}{1-\gamma}$ in calibration for numerical simplicity. It does not affect the calibration result since it is monotonicity transformation.

expectation of the agent. The principal-agent problem is formulated as follows:

$$\max_{W_T(\cdot),e} \mathbb{E}\left[\widetilde{P_T} - W_T\left(P_T\right) \mid e\right] \tag{3}$$

$$s.t. \quad \mathbb{E}_A\left[V\left(W_T\left(\widetilde{P_T}\right)\right) \mid e\right] - C\left(e\right) \geqslant \underline{U} \tag{4}$$

$$e \in \arg\max_{\tilde{e}} \left\{\mathbb{E}_A\left[V\left(W_T\left(\widetilde{P_T}\right)\right) \mid \tilde{e}\right] - C\left(\tilde{e}\right) - \underline{U}\right\} \tag{5}$$

where W_T is the personal wealth at time T, which is determined by the stock performance $\widetilde{P_T}$. The wealth of the agent at time T is the sum of the personal initial wealth $\omega_0 \exp\{rT\}$ and the compensation π_T. $\underline{U}$ is the reservation utility of the outside option for the CEO. Expression (3) is the objective function of the principal. Expression (5) is the CEOs' incentive compatibility constraint, which can be re-written in the form of first-order condition $\frac{\partial}{\partial e}\mathbb{E}_A\left[V(W_T(\widetilde{P_T})) \mid e\right] = \frac{d}{de}C(e)$.

To solve this principal-agent problem, a two-step method can be adopted. In the first step, we fixed the effort level e to every possible value, and solve the principal-agent problem for the optimal wealth structure $W_T\left(\cdot\right)$. In the second step, we search for the effort level that generate the highest profit. Because I am more interested in the pay structure W_T, I particularly focus on the first step. For a given effort level e, the optimal wealth structure $W_T\left(\cdot\right)$ can be calculated by solving the equivalent cost-minimizing program:

$$\min_{W_T(\cdot)} \mathbb{E}\left[W_T\left(\widetilde{P_T}\right) \mid e\right] \tag{6}$$

$$s.t. \quad \mathbb{E}_A\left[V(W_T(\widetilde{P_T})) \mid e\right] - C(e) \geqslant \underline{U} \tag{7}$$

$$\frac{\partial}{\partial e}\mathbb{E}_A\left[V(W_T(\widetilde{P_T})) \mid e\right] = \frac{d}{de}C(e) \tag{8}$$

2.4 Generalized optimal contract

In this subsection, I solve the principal-agent problem (6) to (8). Denote the "true" cumulative distribution function as $F\left(\cdot\right)$. This is the belief of the firm. The CEO transforms the objective distribution function into the subjective distribution function by a probability transformation function $\Psi\left(\cdot\right)$. To simplify the calulation, I denote the subjective cumulative distribution function after transformation as $G\left(\cdot\right)$. I use $G\left(P_T \mid e\right)$ to replace $\Psi\left(F\left(P_T \mid e\right)\right)$ in the equations. The principal-agent problem (6) to (8) can be re-written in the integral form as follows:

$$\begin{aligned}
&\min_{W_T(\cdot)} \int_0^\infty W_T\left(P_T\right) dF\left(P_T \mid e\right) \\
&s.t. \quad \int_0^\infty V\left[W_T\left(P_T\right)\right] dG\left(P_T \mid e\right) - C\left(e\right) \geqslant \underline{U} \\
&\frac{d}{de}\left[\int_0^\infty V\left[W_T\left(P_T\right)\right] dG\left(P_T \mid e\right) - C\left(e\right)\right] = 0
\end{aligned} \tag{9}$$

where $F(\cdot)$ is the objective probability distribution of the future stock price, i.e., $F\left(P_T \mid e\right) = \Pr\left\{\widetilde{P_T} \leqslant P_T \mid e\right\}$. $\Psi\left(\cdot\right)$ is the general probability transformation function. Because the stock price is log-normally distributed, the integral domain is from 0 to infinity.

Equation (9) can be written as:

$$\frac{d}{de}\left[\int_0^{\infty} V\left[W_T\left(P_T\right)\right] g\left(P_T \mid e\right) dP_T - C\left(e\right)\right] = 0$$
$$\implies \int_0^{\infty} V\left[W_T\left(P_T\right)\right] g_e\left(P_T \mid e\right) dP_T - C'\left(e\right) = 0,$$

where $g_e\left(P_T \mid e\right) = \frac{d}{de} g\left(P_T \mid e\right)$.

Construct the Lagrangian:

$$\mathcal{L} = \int_0^{\infty} W_T\left(P_T\right) f\left(P_T \mid e\right) dP_T - \lambda_o \left[\int_0^{\infty} V\left(W_T\left(P_T\right)\right) g\left(P_T \mid e\right) dP_T - C\left(e\right) - \underline{U}\right]$$
$$- \lambda_e \left[\int_0^{\infty} V\left(W_T\left(P_T\right)\right) g_e\left(P_T \mid e\right) dP_T - C'\left(e\right)\right]$$

Take the first order derivative of the Lagrangian with respect to the wealth W_T at every point of P_T, we have:

$$0 = \frac{d\mathcal{L}}{dW_T(\cdot)} = f\left(P_T \mid e\right) - \lambda_o V'\left(W_T\left(P_T\right)\right) g\left(P_T \mid e\right) - \lambda_e V'\left(W_T\left(P_T\right)\right) g_e\left(P_T \mid e\right)$$

Rearrange it:

$$\left\{V'\left(W_T\left(P_T\right)\right)\right\}^{-1} = \lambda_o \frac{g\left(P_T \mid e\right)}{f\left(P_T \mid e\right)} + \lambda_e \frac{g_e\left(P_T \mid e\right)}{f\left(P_T \mid e\right)} \tag{10}$$
$$= \frac{g\left(P_T \mid e\right)}{f\left(P_T \mid e\right)} \left[\lambda_o + \lambda_e \frac{g_e\left(P_T \mid e\right)}{g\left(P_T \mid e\right)}\right] \tag{11}$$

Assume the utility function of CEO is CRRA, i.e., $V'\left(W_T\left(P_T\right)\right) = \left(W_T\left(P_T\right)\right)^{-1}$. So the wealth of the CEO at time T has the following structure,

$$W_T\left(P_T\right) = \left\{\frac{g\left(P_T\right)}{f\left(P_T\right)} \left[\lambda_o + \lambda_e \frac{g_e\left(P_T \mid e\right)}{g\left(P_T \mid e\right)}\right]\right\}^{1/\gamma}. \tag{12}$$

2.5 Probability-weighting and sigma-mu transformation

To further solve the generalized optimal contract in Equation (12), I need to assume the explicit functional form of $\Psi\left(\cdot\right)$. I use the weighting function proposed by Tversky and Kahneman (1992) and the rank-dependent expected utility proposed by Quiggin (1982). In this paper, I do not use the cumulative prospect theory. I use CRRA utility function instead of the loss-aversion utility function. The reason is that loss-aversion alone have explanatory power in explaining the convexity of the observed contracts in the central region Dittmann, Maug, and Spalt (2010). I exclude loss aversion from my model to see whether probability weighting can also explain the convexity of the observed contracts.

If the CEO is probability weighting, he transforms the cumulative probability with a weighting function. In other words, he attaches extreme outcomes with higher decision weight. The parametrized form of weighting function proposed by Tversky and Kahneman (1992) is:

$$\Psi_{pw}(p) = \frac{p^{\delta}}{\left(p^{\delta} + (1-p)^{\delta}\right)^{1/\delta}}, \tag{13}$$

where p is the original cumulative probability and δ is the parameter. I use this function in my model as the probability weighting function that transforms the original cumulative probabilities into decision weights when calculating expected utility. The green dotted curve in Figure 1 shows the probability weighting function when δ is arbitrarily set to 0.6. The big curvature at the two ends means an exaggeration of the probability of extreme outcomes. The relatively flat curve at the middle part shows that the agent underestimate the probability of the middle outcomes. In my model, the CEO transforms the original cumulative distribution function $J(x)$ to a new cumulative distribution function $K(x)$ with the probability weighting function $\Psi_{pw}(\cdot)$. The subjective cumulative distribution function $K(x)$ is

$$K(x) = \Psi_{pw}(J(x)) = \frac{J(x)^{\delta}}{\left(J(x)^{\delta} + (1-J(x))^{\delta}\right)^{1/\delta}}.$$

The probability weighting is parametrized with a single parameter δ in Equation (13). However, using parametrization with δ may cause inconvenience if I want to have an closed-form solution for the optimal contract described in Equation (12). To be more specific, if the original distribution function $J(\cdot)$ describes a normal distribution, after transformed by probability weighting function $\Psi_{pw}(J(x))$, the new distribution function $K(\cdot)$ cannot be described by any popular or known distribution function. This not only causes difficulty in theoretical analysis, but brings troubles in empirical calibration. To solve this problem, I consider other probability transformation functions with similar properties as the probability weighting function $\Psi_{pw}(\cdot)$.

One solution is to assume that the new distribution after being transformed is still a normal distribution. That is to say that a function $\Psi_{sm}(\cdot)$ transforms the original normal distribution to a new normal distribution. And the function $\Psi_{sm}(\cdot)$ is similar in shape with the probability weighting function $\Psi_{pw}(\cdot)$. I call the function $\Psi_{sm}(\cdot)$ as "sigma-mu transformation". Namely, only the mean and variance of the objective distribution is changed.

Using the sigma-mu transformation as an alternative to the probability weighting function has three advantages. First, the sigma-mu transformation function $\Psi_{sm}(\cdot)$ can transform normal distribution to another normal distribution, which exhibits similar properties with the probability weighting function. The red dashed curve in Figure 1 shows the transformation of normal distributions by changing mean and variance. For certain sigma and mu, the sigma-mu transformation function has the similar shape of a typical probability weighting function: It is concave and convex at the two ends. It exhibits insensitivity in the central region. As the second advantage, after the CEO applies probability weighting using sigma-mu transformation, the subjective stock price still follows the normal distribution. Thus, the probability density function of the new distribution can be explicitly written, which results in a closed-form solution for the analytical optimal general contract. Third, using sigma-mu transformation simplifies the empirical calibration of the model because I do not need to numerically calculate the PDF of the new distribution after adopting the probability weighting function $\Psi_{pw}(\cdot)$.

Now, I show how to sigma-mu transformation to approximate the probability weighting function. $J(x)$ is the cumulative distribution function of the random variable $\widetilde{X} \sim \mathcal{N}(\mu, \sigma^2)$, which describes the objective distribution of the stock return. The true mean is μ and the true variance is σ^2. So $J(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$ and $J(p) = \sigma\Phi^{-1}(p) + \mu$, where $\Phi(\cdot)$ is the CDF of the standard normal distribution. $K(y)$ is the cumulative distribution function of the random variable of $\widetilde{Y} \sim \mathcal{N}(\mu_A, \sigma_A^2)$, which describes the agent's subjective distribution of the future return. So $J(y) = \Phi\left(\frac{y-\mu_A}{\sigma_A}\right)$. Because the true CDF, $J(x)$, is transformed to a the subjective CDF, $K(x)$, using the sigma-mu transformation function $\Psi_{sm}(\cdot)$, we have,

$$
\begin{aligned}
& \Psi_{sm}(J(x)) = K(x) \\
\Longrightarrow\ & \Psi_{sm}(J(x)) = K\left(J^{-1}(J(x))\right) \\
\Longrightarrow\ & \Psi_{sm}(J(x)) = \Phi\left(\frac{\sigma\Phi^{-1}(J(x)) + \mu - \mu_A}{\sigma_A}\right) \\
\Longrightarrow\ & \Psi_{sm}(J(x)) = \Phi\left(\frac{\sigma}{\sigma_A}\Phi^{-1}(J(x)) + \frac{\mu-\mu_A}{\sigma_A}\right) \\
\Longrightarrow\ & \Psi_{sm}(p) = \Phi\left(\frac{\sigma}{\sigma_A}\Phi^{-1}(p) + \frac{\mu-\mu_A}{\sigma_A}\right) \\
\Longrightarrow\ & \Psi_{sm}(p) = \Phi\left(\frac{\Phi^{-1}(p) - \frac{\mu_A-\mu}{\sigma}}{\frac{\sigma_A}{\sigma}}\right)
\end{aligned}
$$

To ensure that the transformation function $\Psi_{sm}(\cdot)$ has a fixed functional form, $\frac{\sigma_A}{\sigma}$ and $\frac{\mu-\mu_A}{\sigma_A}$ need to be constant. Denote

$$
\begin{cases}
\eta_s & = \frac{\sigma_A}{\sigma} \\
\eta_m & = \frac{\mu_A-\mu}{\sigma}
\end{cases}.
$$

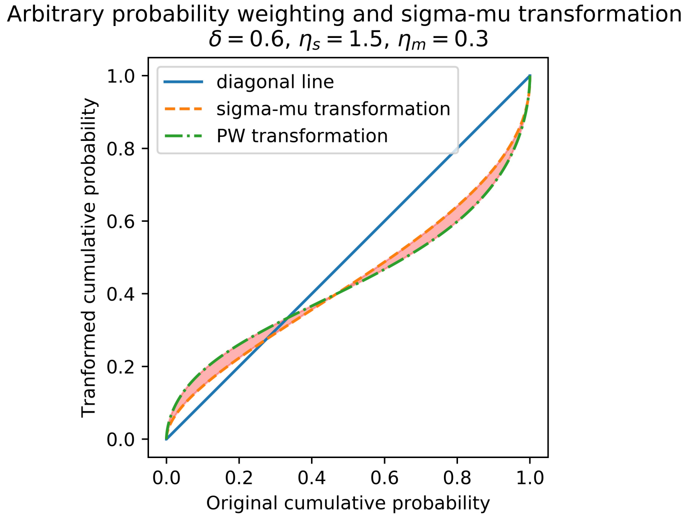
The sigma-mu transformation transfer the subjective normal distribution to the agent's objective normal distribution by changing the parameters σ and μ:

$$
\begin{cases}
\sigma_A & = \eta_s\sigma \\
\mu_A & = \mu + \eta_m\sigma
\end{cases} \tag{14}
$$

In another words, if a random variable follows normal distribution $\widetilde{X} \sim \mathcal{N}\left(\mu,\ \sigma^2\right)$, then the new random variable after the sigma-mu transformation also follows normal distribution. The new random variable is: $\widetilde{Y} = \eta_s\widetilde{X} + \mu - \eta_s\mu + \eta_m\sigma$, and $\widetilde{Y} \sim \mathcal{N}\left(\mu + \eta_m\sigma,\ \eta_s^2\sigma^2\right)$. I make a comparison between the probability weighting function $\Psi_{pw}(\cdot)$ and the sigma-mu transformation function $\Psi_{sm}(\cdot)$. The yellow dotted curve in Figure 1 is the sigma-mu transformation function when η_s is arbitrarily set to 1.5 and η_m is set to 0.3. It has similar properties with the probability weighting function with $\delta = 0.6$ shown by the green dotted curve. They are both curved at the two ends and

Figure 1: The approximation of arbitrary probability weighting function

This figure shows the shape of the probability-weighting function in green dashed curve and the sigma-mu transformation function in yellow dashed curve. The x-axis is the original cumulative probability of the normal distribution, and the y-axis is the transformed cumulative probability. $\delta = 0.6$, $\eta_s = 1.5$, and $\eta_m = 0.3$.



relatively flat in the middle.

To approximate the probability weighting function using sigma-mu transformation, I minimize the squared distance between the two curves, which is depicted in the Figure 1 as the red shadow. In other words, when the probability weighting is parameterized by a given parameter δ, I search for two parameters η_s and η_m that gives the similar shape.

I define the distance metrics as the squared distance between the probability weighting function and the sigma-mu transformation function as:

$$\mathbb{E}\left[(\psi_{pm}(P) - \psi_{sm}(P))^2\right] = \int_0^1 (\psi_{pm}(P) - \psi_{sm}(P))^2 \, dP \tag{15}$$

I numerically search for the optimal pair of (η_s,η_m) that minimizes distance metrics defined in Equation (15). For $\delta = 0.6$, the optimal solution is that $\eta_s = 1.7877$ and $\eta_m = 0.3576$. Figure 2 shows the shape of the probability-weighting function with $\delta = 0.6$ as the green dashed curve, and the shape of the sigma-mu transformation function with optimized parameters as the yellow dashed curve. We can see that the curves are very close to each other.

Table 1 shows the optimized pairs of (η_s,η_m) and the squared difference, corresponding to different δ. δ takes values from 0.3 to 0.9. When δ is close to 1, the curve of the probability weighting tends to be a straight line, so η_s is close to 1 and η_m is close to 0. When δ is smaller, the curvation of the probability weighting is larger. There are larger distortion of the probability at the two ends. Both η_s and η_m are decreasing in δ.

Figure 2: The optimal approximation of the probability weighting function

This figure shows the shape of the probability-weighting function in green dashed curve with $\delta = 0.6$, and the sigma-mu transformation function in yellow dashed curve with $\eta_s = 1.7877$ and $\eta_m = 0.3576$. The x-axis is the original cumulative probability of the normal distribution, and the y-axis is the transformed cumulative probability.

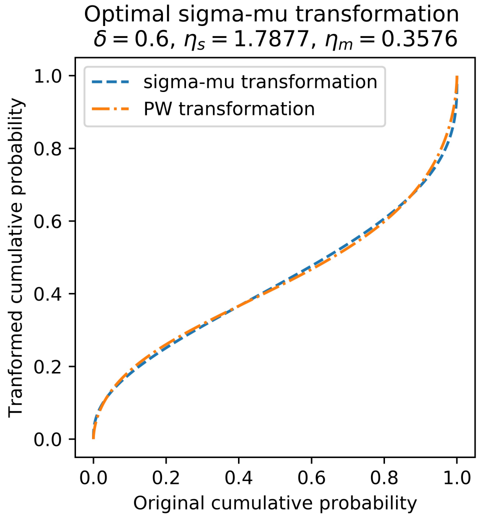


Table 1: Approximation of probability weighting using η_s and η_m

This table shows the optimized pairs of (η_s,η_m), as well as the squared difference, corresponding to different δ. δ takes values from 0.3 to 0.9. The (η_s,η_m) are optimized by minimizing the squared difference metrics defined in Equation (15).

θ	η_s	η_m	squared difference
0.3	3.86119	3.58262	0.00088
0.4	2.88477	1.65702	0.00071
0.5	2.23307	0.77908	0.00033
0.6	1.78775	0.35764	0.00011
0.7	1.47895	0.15139	0.00003
0.8	1.26259	0.05263	0.00001
0.9	1.10953	0.01061	0.00000

2.6 Optimal general contract with sigma-mu transformation

After using sigma-mu transformation to approximate probability-weighting function, I derive the closed form for the optimal general contract in Equation (12). The stock price of the firm follows log-normal distribution:

$$\widetilde{P_T} \sim \ln \mathcal{N}\left(\mu\left(e\right),\ \sigma^2 T\right),$$

where $\mu\left(e\right) = \ln P_0\left(e\right) + \left(r - d\right)T - \frac{\sigma^2}{2}T$. Thus, the objective probability density functions of $\widetilde{P_T}$ is defined as:

$$f\left(P_T \mid e\right) = \frac{1}{P_T\sqrt{2\pi\sigma^2 T}} \exp\left[-\frac{\left(\ln P_T - \mu\left(e\right)\right)^2}{2\sigma^2 T}\right].$$

The parameters of the objective future stock price are transformed according to the sigma-mu transformation, i.e., Equation (14). So the subjective the future stock price is:

$$\begin{aligned}
\widetilde{P_T^A} &= \exp\left\{\sigma_A\sqrt{T}\widetilde{u} + \mu_A\left(e\right)\right\} \\
&= \exp\left\{\eta_s\sigma\sqrt{T}\widetilde{u} + \mu\left(e\right) + \eta_m\sigma\sqrt{T}\right\} \\
&= \exp\left\{\sigma\sqrt{T}(\eta_s\widetilde{u} + \eta_m) + \mu\left(e\right)\right\} \\
&= \exp\left\{\sigma\sqrt{T}\widetilde{v} + \mu\left(e\right)\right\} \\
&= P_0\left(e\right)\exp\left\{\left(r + d\right)T - \frac{\sigma^2}{2}T + \widetilde{v}\sigma\sqrt{T}\right\}
\end{aligned}$$

where $\widetilde{u} \sim \mathcal{N}(0,1)$, and $\widetilde{v} \sim \mathcal{N}(\eta_m, \eta_s^2)$. The agent's subjective $\widetilde{P_T^A}$ is equal to the firm value at time 0 multiplied by a different stochastic factor that follows the log-normal distribution. The CEO's subjective probability density function of $\widetilde{P_T}$ can be written as:

$$g\left(P_T \mid e\right) = \frac{1}{P_T\sqrt{2\pi\sigma_A^2 T}} \exp\left[-\frac{\left(\ln P_T - \mu_A\left(e\right)\right)^2}{2\sigma_A^2 T}\right],$$

where $\mu_A\left(e\right)$ and $\sigma_A\left(e\right)$ are calculated by Equation (14). The partial derivative of $g\left(P_T \mid e\right)$ with respect to e is:

$$g_e\left(P_T \mid e\right) = \underbrace{\frac{1}{P_T\sqrt{2\pi\sigma_A^2 T}} \exp\left[-\frac{\left(\ln P_T - \mu_A\left(e\right)\right)^2}{2\sigma_A^2 T}\right]}_{g(P_T|e)} \left(-\frac{2}{2\sigma_A^2 T}\left(\ln P_T - \mu_A\left(e\right)\right)\right)\left(-\frac{d\mu_A\left(e\right)}{de}\right) \tag{16}$$

$$= g\left(P_T \mid e\right)\left(\frac{\ln P_T - \mu_A\left(e\right)}{\sigma_A^2 T}\right)\frac{d\left(\ln P_0(e) + \left(r_f - d\right)T - \frac{\sigma^2}{2}T + \eta_m\sigma\sqrt{T}\right)}{de} \tag{17}$$

$$= g\left(P_T \mid e\right)\left(\frac{\ln P_T - \mu_A\left(e\right)}{\sigma_A^2 T}\right)\left(\frac{dP_0\left(e\right)/de}{P_0\left(e\right)}\right). \tag{18}$$

So the ratio between $g_e(P_T \mid e)$ and $g_e(P_T \mid e)$ is,

$$\frac{g_e(P_T \mid e)}{g(P_T \mid e)} = \frac{\ln P_T - \mu_A(e)}{\sigma_A^2 T} \cdot \frac{P_0'(e)}{P_0(e)} \tag{19}$$

The ratio between $g(P_T \mid e)$ and $f(P_T \mid e)$ is:

$$\begin{aligned}
\frac{g(P_T \mid e)}{f(P_T \mid e)} &= \frac{P_T\sqrt{2\pi\sigma^2 T}}{P_T\sqrt{2\pi\sigma_A^2 T}} \exp\left[\frac{(\ln P_T - \mu)^2}{2\sigma^2 T} - \frac{(\ln P_T - \mu_A)^2}{2\sigma_A^2 T}\right] \\
&= \frac{\sigma}{\sigma_A} \exp\left[\frac{(\ln P_T - \mu)^2}{2\sigma^2 T} - \frac{(\ln P_T - \mu_A)^2}{2\sigma_A^2 T}\right] \\
&= \frac{\sigma}{\sigma_A} \exp\left[\left(\frac{1}{2\sigma^2 T} - \frac{1}{2\sigma_A^2 T}\right)(\ln P_T)^2 + \left(\frac{\mu_A}{\sigma_A^2 T} - \frac{\mu}{\sigma^2 T}\right)\ln P_T + \left(\frac{\mu^2}{2\sigma^2 T} - \frac{\mu_A^2}{2\sigma_A^2 T}\right)\right]
\end{aligned} \tag{20}$$

Insert Equation (19) and Equation (20) into the generalized optimal contract, i.e., Equation (12):

$$\begin{aligned}
W_T(P_T) &= \left\{\frac{g(P_T)}{f(P_T)}\left[\lambda_o + \lambda_e \frac{g_e(P_T \mid e)}{g(P_T \mid e)}\right]\right\}^{1/\gamma} \\
&= \left\{\frac{\sigma}{\sigma_A} \exp\left[\beta_2 (\ln P_T)^2 + \beta_1 \ln P_T + \beta_0\right]\left[\lambda_o + \lambda_e \frac{\ln P_T - \mu_A(e)}{\sigma_A^2 T} \cdot \frac{P_0'(e)}{P_0(e)}\right]\right\}^{1/\gamma} \\
&= \left\{\frac{\sigma}{\sigma_A} \exp\left[\beta_2 (\ln P_T)^2 + \beta_1 \ln P_T + \beta_0\right]\left[\lambda_e \frac{P_0'(e)}{\sigma_A^2 T P_0(e)} \ln P_T + \lambda_o - \lambda_e \frac{P_0'(e)}{\sigma_A^2 T P_0(e)} \mu_A(e)\right]\right\}^{1/\gamma} \\
&= \left[\exp\left(\beta_2 (\ln P_T)^2 + \beta_1 \ln P_T + \beta_0\right)(\alpha_1 \ln P_T + \alpha_0)\right]^{1/\gamma},
\end{aligned}$$

where

$$\begin{cases}
\alpha_0 = & \frac{\sigma}{\sigma_A}\left[\lambda_o - \lambda_e \frac{P_0'(e)}{\sigma_A^2 T P_0(e)} \mu_A(e)\right] \\
\alpha_1 = & \frac{\sigma}{\sigma_A} \lambda_e \frac{P_0'(e)}{\sigma_A^2 T P_0(e)} \\
\beta_0 = & \frac{\mu^2}{2\sigma^2 T} - \frac{\mu_A^2}{2\sigma_A^2 T} \\
\beta_1 = & \frac{\mu_A}{\sigma_A^2 T} - \frac{\mu}{\sigma^2 T} \\
\beta_2 = & \frac{1}{2\sigma^2 T} - \frac{1}{2\sigma_A^2 T}
\end{cases}$$

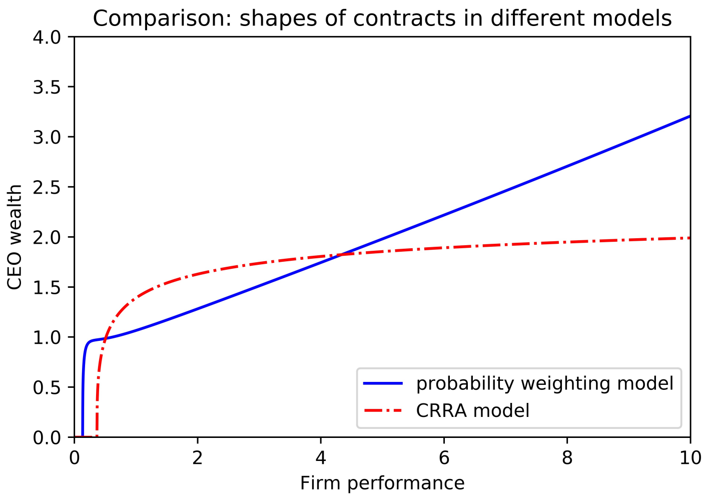
Proposition 1 *The optimal contract that solves the principal-agent problem (6) to (8) given CRRA utility (Equation (2)) and an approximation to the probability weighting feature (Equation (14)) has the following shape:*

$$W_T(P_T) = \begin{cases} \left[\exp\left\{\beta_2 (\ln P_T)^2 + \beta_1 \ln P_T + \beta_0\right\}(\alpha_1 \ln P_T + \alpha_0)\right]^{1/\gamma} & \textit{if } \ln P_T > -\frac{\alpha_0}{\alpha_1} \\ 0 & \textit{if } \ln P_T \leqslant -\frac{\alpha_0}{\alpha_1} \end{cases} \tag{21}$$

The optimal general contract (Equation (21)) is increasing and convex when P_T is large. To see this, I show that $\lim_{P_T \to \infty} \frac{dW_T(P_T)}{dP_T} = \infty$ for all of the CRRA parameters. The proof for this

Figure 3: Shapes of contracts for different models

This figure compares the shapes of analytical optimal general contracts in different models. The red dashed curve shows the optimal contract when CEO has CRRA utility but is not probability weighting. The blue curve shows the optimal contract when CEO has both CRRA utility and is probability weighting. The x-axis is the firm performance, and the y-axis is the CEO wealth at the time T.



feature is in Appendix A. An intuition for the convexity of the contract for large P_T is that the term $\exp\left\{\beta_2 (\ln P_T)^2 + \beta_1 \ln P_T + \beta_0\right\}$ is not only convex when P_T is large, but also dominates the term $(\alpha_1 \ln P_T + \alpha_0)$ in its power. Dittmann and Maug (2007) show that the optimal contract for the CRRA model has the shape $W_T(P_T) = (\alpha_1 \ln P_T + \alpha_0)^{1/\gamma}$.[3] Thus, the exponential component is the key feature why my model can predict a convex contract shape of the contract. Notably, when $\ln P_T \leqslant -\frac{\alpha_0}{\alpha_1}$, the contract should take value 0 because the term under the power $1/\gamma$ is non-positive.

Figure 3 compares the shapes of optimal general contracts in different models. The red dashed curve shows the optimal contract when CEO has CRRA utility but is not probability weighting. The blue curve shows the optimal contract when CEO has both CRRA utility and is probability weighting. We can see that the optimal contract for the CRRA model is globally concave and it becomes flat for high performance. This cannot explain the observed positive option grants. However, the optimal contract for the probability weighting model is convex when performance is high. This well explains the positive option holding in observed CEO compensation contracts.

3. Calibration strategy

In the last section, I show that the optimal general contract features convexity when CEO is probability weighting. This provides theoretical evidence for the existence of positive options in CEO

[3] To be more specific, Dittmann and Maug (2007) show that the optimal general contract for CRRA model is:

$$W_T(P_T) = \begin{cases} (\alpha_1 \ln P_T + \alpha_0)^{1/\gamma} & \text{if } \ln P_T > -\frac{\alpha_0}{\alpha_1} \\ 0 & \text{if } \ln P_T \leqslant -\frac{\alpha_0}{\alpha_1} \end{cases}$$

compensation. I then calibrate this model with observed CEO contracts in the U.S. firms so see whether my model works empirically. This section introduces the methods for model calibration with the observed data sets. The idea behind the calibration is as follows. If the observed contract is optimal, then it must provide an incentive for the CEO to choose the optimal effort level $e*$. Using this assumption, I numerically solve the formulae (6) to (8) to search for a contract which provide the same incentive to the CEO to choose the optimal e^*, but generates lower expected cost for the shareholders. If the observed contract is optimal, then the new contract searched by the principal-agent problem should generate similar shape with the observed contract. Dittmann and Maug (2007) find that when the CEO has CRRA utility, the optimal piecewise linear contracts predict no option holding and negative fixed salary. In this paper, when CEOs are probability weighting, the optimal piecewise linear contracts should contain positive option holding and positive fixed salary.

3.1 Pay-performance sensitivity

Because the production function $P_0(e)$ and the cost function $C(e)$ are not known. I need to first re-write the incentive constraint in another form. Equation (8) can be re-written as:

$$\mathbb{E}_A\left[\frac{dV\left(\widetilde{W_T}\right)}{d\widetilde{W_T}}\cdot\frac{d\widetilde{W_T}}{d\widetilde{P_T}}\cdot\frac{d\widetilde{P_T}}{dP_0}\cdot\frac{dP_0}{de}\right]_{e=e^*}=\left[\frac{dC\left(e\right)}{de}\right]_{e=e^*}$$

$$\Longrightarrow\mathbb{E}_A\left[\frac{dV\left(\widetilde{W_T}\right)}{d\widetilde{W_T}}\cdot\frac{d\widetilde{W_T}}{d\widetilde{P_T}}\cdot\frac{d\widetilde{P_T}}{dP_0}\right]_{e=e^*}=\left[\frac{dC\left(e\right)}{de}\Big/\frac{dP_0}{de}\right]_{e=e^*}\equiv UPPS$$

Because $\frac{dP_0}{de}$ and $\frac{dC(e)}{de}$ have fixed values given the certain effort level e^*, the expression

$$\mathbb{E}_A\left[\frac{dV\left(\widetilde{W_T}\right)}{d\widetilde{W_T}}\cdot\frac{d\widetilde{W_T}}{d\widetilde{P_T}}\cdot\frac{d\widetilde{P_T}}{dP_0}\right]$$

should be constant if the new contract provides the same effort incentive. I denote this as utility-adjusted pay-performance sensitivity (UPPS), which measures how the utility of CEO is reacted to the firm's intrinsic value. Using the analogy of option, this term can also be denoted as the Delta of the contract.

3.2 Calibration model

I restrict myself to the piecewise linear contract. The model corresponding to the principal-agent problem (6) to (8) for my calibration can be formulated as:

$$\min_{\phi,n_s,n_o} \mathbb{E}\left[W_T\left(\widetilde{P_T} \mid \phi, n_s, n_o\right)\right] \tag{22}$$

$$s.t.\quad \mathbb{E}_A\left[V\left(W_T\left(\widetilde{P_T} \mid \phi, n_s, n_o\right)\right)\right] \geqslant \mathbb{E}_A\left[V\left(W_T^o\left(\widetilde{P_T} \mid \phi^o, n_s^o, n_o^o\right)\right)\right] \tag{23}$$

$$UPPS\left[W_T\left(\widetilde{P_T} \mid \phi, n_s, n_o\right)\right] \geqslant UPPS\left[W_T^o\left(\widetilde{P_T} \mid \phi^o, n_s^o, n_o^o\right)\right] \tag{24}$$

where W_T^o stands for the observed piece-wise contract which is determined by fixed salary ϕ^o, number of share holding n_s^o, number of option holding n_o^o. The optimal piecewise linear contract should provide at least the same utility-adjusted pay-performance sensitivity. It should also provide at least the same expected utility as the observed contract so that the CEO will be willing to accept the contract.

To numerically solve the principal-agent problem, Equations (22) to (24) should be written in the integral forms. The wealth of the CEO at the time T for a piecewise linear contract is:

$$\widetilde{W_T} = (\phi + \omega_0)e^{rT} + n_s e^{dT}\widetilde{P_T} + n_o \max\left\{\widetilde{P_T} - K, 0\right\}. \tag{25}$$

The objective function (22) is the expected costs of the contract, which is evaluated with the shareholders' preference. It can be written in the integral form as

$$\begin{aligned}&\mathbb{E}\left[W_T\left(\widetilde{P_T} \mid \phi, n_s, n_o\right)\right]\\ =&\int_{-\infty}^{\infty} W_T dF\left(u\right),\end{aligned}$$

where $F(u)$ is the CDF for random variable $\widetilde{u} \sim \mathcal{N}(0,1)$.

The expected utility of the CEO in the participation constraint (23) can be written in the integral form as

$$\begin{aligned}&\mathbb{E}_A\left[V\left(W_T\left(\widetilde{P_T} \mid \phi, n_s, n_o\right)\right)\right]\\ =&\int_{-\infty}^{\infty} \frac{W_T^{1-\gamma}}{1-\gamma} dG\left(v\right),\end{aligned}$$

where $G(v)$ is the CDF for random variable $\widetilde{v} \sim \mathcal{N}\left(\eta_m, \eta_s^2\right)$, i.e., the distribution after approximation to probability weighting.

To numerically calculate the utility-adjusted pay-performance sensitivity, I need to find explicit expressions for the derivatives $\frac{dV\left(\widetilde{W_T}\right)}{d\widetilde{W_T}}$, $\frac{d\widetilde{W_T}}{d\widetilde{P_T}}$ and $\frac{d\widetilde{P_T}}{dP_0}$. The utility function of the CEO is $V(W_T) = \frac{W_T^{1-\gamma}}{1-\gamma}$, so $V'(W_T) = \frac{dV(W_T)}{dW_T} = W_T^{-\gamma}$. According to Equation (25), the derivation of $\widetilde{W_T}$ w.r.t. $\widetilde{P_T}$

is

$$\frac{d\widetilde{W_T}}{d\widetilde{P_T}} = n_s e^{dT} + n_o \mathbb{I}_{\widetilde{P_T}>K}.$$

Since $\widetilde{P_T} = P_0 \exp\left\{\left(r_f - d - \frac{\sigma^2}{2}\right)T + \widetilde{v}\sigma\sqrt{T}\right\}$, the derivation of $\widetilde{P_T}$ w.r.t P_0 is

$$\frac{d\widetilde{P_T}}{dP_0} = \exp\left\{\left(r_f - d - \frac{\sigma^2}{2}\right)T + \widetilde{v}\sigma\sqrt{T}\right\}.$$

Now I have the the functional forms of the derivatives $\frac{dV\left(\widetilde{W_T}\right)}{d\widetilde{W_T}}$, $\frac{d\widetilde{W_T}}{d\widetilde{P_T}}$ and $\frac{d\widetilde{P_T}}{dP_0}$. Thus, the Utility-adjusted pay-performance sensitivity in the incentive constraint (Equation (24)) can be written as:

$$\begin{aligned} UPPS =& \mathbb{E}_A\left[\frac{dV\left(\widetilde{W_T}\right)}{dP_0}\right] = \mathbb{E}_A\left[\frac{dV\left(\widetilde{W_T}\right)}{d\widetilde{W_T}} \cdot \frac{d\widetilde{W_T}}{d\widetilde{P_T}} \cdot \frac{d\widetilde{P_T}}{dP_0}\right] \\ =& \int_{-\infty}^{\infty} W_T^{-\gamma}\left[n_s e^{dT} + n_o \mathbb{I}_{P_T>K}\right] \exp\left\{\left(r_f - d - \frac{\sigma^2}{2}\right)T + v\sigma\sqrt{T}\right\} dG\left(v\right). \end{aligned}$$

3.3 Numerical solution

The calibration model is numerically solved by computer programs. For each observed CEO contract, the input variables that need to be optimized are fixed salary, share holding, and option holding, e.g., (ϕ^o, n_s^o, n_o^o). The parameters are strike price, maturity time, stock volatility and CEO personal wealth, namely the tuple $(K^o, T^o, \sigma^o, \omega_0^o)$. The notation o indicates the observed contracts. The programs search for the optimal set of (ϕ, n_s, n_o) that minimums the value of Equation (22) with the constraints of Equations (23) and (24). Moreover, there are also bounds for the variables. First, ϕ is set to be larger than minus CEO personal wealth ω_0. This means that the firm can "punish" the CEO when the firm performance is bad, but still ensures that the CEO's wealth at time T will not be negative. Due to the non-negativity of the CEO's final wealth, the utility function of the CEO will always be defined. Second, n_s is set to be between 0 and 1. It means that the stocks held by the CEO cannot exceed the total shares outstanding of the firm. Third, n_o is set to be non-negative, which means that the CEO cannot "sell" stock options of her own firm.

Because the integration is done numerically, the domain cannot be spanned to $-\infty$ and ∞. In my program, the integration domain is between -20 and 20. For a random variable following standard normal distribution, the probability that it goes beyond the range $(-20, 20)$ is smaller than 5.51×10^{-89}. So the error is very small. Because the utility function of the CEO takes different functional forms when $\gamma \neq 1$ and when $\gamma = 1$, I use $\gamma = 1.01$ in the calibration so that the function $V\left(W_T\right) = \frac{W_T^{1-\gamma}}{1-\gamma}$ are always used.

4. Data

The data set for observed CEO compensation contracts is constructed on the basis of the compensation databases from Execucomp. I select the year 2012 as the year of interest. This year provides the biggest number of observations in the sample for my calibration. The selection of the observed contracts follows several criteria. First of all, the executive must show up as the CEO in 2012. Second, the CEO should work in the same firm for the full fiscal years of 2011 and 2012. Third, the CEO should show up in the data set from the year 2007. This criterion is used for calculating the personal wealth of the CEO. It does not require that the CEO work in the same firm during the period.

In this paper, observed CEO pay packages are summarized as a stylized contract that consists of three component: fixed pay, stock, and options. This stylized compensation contract has been used in previous literature, e.g., (Dittmann and Maug, 2007; Dittmann, Yu, and Zhang, 2017). In this three-component view, a CEO contract has a component that is not affected by firm performance, a component that is highly correlated with firm performance, and a component that is correlated with firm performance only when performance exceeds certain thresholds. The stylized CEO compensation contract summarizes the complicated pay package and relates CEO wealth, to only to, the stock price of the firm. This simplifies my calculation. Actual CEO compensation is more complicated. For example, Bizjak, Kalpathy, Li, and Young (2017) shows that in 2012, 37.2% of firms in their sample use relative performance awards. The award paid to CEO is based on firm performance relative to peer companies. In this paper, I do not include relative performance awards into stylized CEO compensation contract for two reasons. First, relative performance awards relate CEO compensation to both firm performance and peer performance. This complicates the wealth function and brings difficulties to solving the non-linear programming. Second, in this paper, I mainly focus on option grants and convexity of CEO compensation contracts. Other long-term incentive packages are of second-order importance and are not considered in the model.

The fixed salary ϕ of the CEO is composed of base salary, annual bonus, non-equity incentive compensation, changes in pension provision, and other compensation. The stock holding n_s is the percent of the non-option shares held by the CEO in 2012. Data on the option holding is obtained from the Outstanding Equity Awards database. This database records all stock option granted after 2006 when the new format was adopted. Each CEO may receive multiple options historically. These options have different numbers n_0, strike prices K, and time to maturity T. That means, in the year 2012, the CEO hold a bunch of a combination of many (n_0^i, K^i, T^i). To calibrate for the optimal contract, I need to find an option that is representative of all options held by the CEO. Thus, I numerically solve for (K, T) from the equation system

$$\begin{cases} n_o BS(P_0, K, T, \sigma, r_f) & = \sum_i n_o^i BS(P_0, K^i, 0.7T^i, \sigma, r_f) \\ n_o \Delta(P_0, K, T, \sigma, r_f) & = \sum_i n_o^i \Delta(P_0, K^i, 0.7T^i, \sigma, r_f), \\ n_0 & = \sum_i n_o^i \end{cases} \tag{26}$$

for the representative option (n_0, K, T). The first equation in equation system (26) indicates that the Black-Scholes value of the representative option should be equal to the aggregated Black-Scholes value of all options currently held by the CEO. The second equation indicates that the Delta value of the representative option should be equal to the aggregated Delta value of all options historically received. The third equation indicates that the number of the representative option should be equal to the total number of the options held by the CEO. The time to maturity is multiplied by 0.7 because the CEOs usually exercises his options before the expiration date (Huddart and Lang, 1996; Carpenter, 1998)[4]. The Black-Scholes value (BS) and the Delta value (Δ) are defined as follows.

$$BS = \mathcal{N}(d_1)P_0 - \mathcal{N}(d_2)Ke^{-r(T)}$$
$$\Delta = \frac{\partial BS}{\partial P_0} = N(d_1),$$

where $d_1 = \frac{1}{\sigma\sqrt{T}}\left[\ln\left(\frac{P_0}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)T\right]$ and $d_2 = d_1 - \sigma\sqrt{T}$. $\mathcal{N}(\cdot)$ is the cumulative normal distribution function.

For each CEO, the equation system (26) is solved once. The number of options n_0 is rescaled by the total outstanding shares of the firm so that it is expressed as a percentage of the firm value. The strike price is multiplied by the total outstanding shares of the firm so that it is comparable to the firm value.

The personal wealth of the CEO is calculated as the sum of the five-year historical fixed salary income received by the CEO from 2007 to 2011, assuming that the CEO did not consume any income received in this period. The tax rate is set to be 42%. The firm value P_0 is the market value of the firm on the last trading date of 2012. The annual standard deviation of the stock return σ is calculated using the stock market performance of the firm in the year 2009, 2010 and 2011. The risk-free rate is the 5-year bond of the US government on the last trading date of 2012. The dividend rate d is the dividend per share of the firm in 2011 obtained from Compustat.

Table 2, Panel A summarizes the main variables in the sample that will be used in the calibration. The sample consists of 622 U.S. CEOs contracts in the year 2012. These CEOs are those who work in the same firm for two years and appear in the database for consecutive five years. The mean firm value is 7,656 million dollars. The mean fixed salary is 3.15 million dollars. A CEO holds on average 1.72% of the firm total shares. The average option holding counts for 0.87% of the firm value. The average age of the CEOs is 57. To see whether these CEOs are representative, Panel B summarizes the sample of 1385 U.S. CEOs appeared in the Compustat database. These CEOs are not in the calibration either because they are not in the same firm for two years, or because they are not in the database from 2007. For these CEOs, representative options and personal wealth are not calculated. We can see that a CEO in the calibration sample holds on average 1.40% of the firm's total shares and 0.7% of the options. The numbers are close to the representative sample for calibration. Moreover, the medians of the variables in the calibration sample are very close to the medians of the variables in the bigger sample. This indicates that my calibration sample is a good representative of all CEO contracts in that year.

[4]In the next version, I will calculate for each CEO how advanced they historically exercise their options in that firm. It is the better measurement of the observed pay timing.

Table 2: Summary statistics: sample of 622 U.S. CEOs

This table summarizes the variables. Panel A summarizes the variables for the sample of 622 U.S. CEOs contracts in the year 2012 that is used in the calibration. Panel B summarizes the variables for the sample of all 1386 U.S. CEOs contracts in the year 2012.

Panel A: sample of 622 U.S. CEOs contracts

Variable	Description	Obs	Mean	Median	Std. Dev.	Min	Max
P_0	Firm value ($m)	622	7656.00	1924.54	18735.94	27.77	220107.40
ϕ	Fixed salary ($m)	622	3.15	2.01	3.66	0.08	37.72
n_s	Shares held (%)	622	1.72	0.47	4.23	0.00	59.93
n_o	Options held (%)	622	0.87	0.57	0.98	0.00	8.64
T	Time to maturity	622	3.74	3.53	1.58	0.24	20.45
ω_0	Personal wealth ($m)	622	6.46	4.88	6.26	0.00	72.51
σ	Yearly standard deviation	622	0.49	0.47	0.18	0.17	1.79
Age	Age of the CEO	622	57.39	57.00	6.62	39.00	85.00

Panel B: Sample of 1386 U.S. CEOs contracts in 2012

Variable	Description	Obs	Mean	Median	Std. Dev.	Min	Max
P_0	Firm value ($m)	1,386	8581.04	1929.81	22384.49	1.89	233999.40
ϕ	Fixed salary ($m)	1,386	2.99	2.02	3.24	0	37.72
n_s	Shares held (%)	1,386	1.40	0.32	3.85	0.00	59.93
n_o	Options held (%)	1,386	0.70	0.38	0.96	0.00	13.31
σ	Yearly standard deviation	1,386	0.51	0.47	0.24	0.15	2.76
Age	Age of the CEO	1,384	56.30	56.00	6.80	35.00	85.00

5. Empirical results

5.1 Optimal linear contract

This section presents the empirical results from the calibration using the sample of 622 U.S. CEO contracts. Table 3 shows the optimal linear contract obtained by numerically solving Equations (22) to (24). Panel A shows the results where CEO is probability-weighting. The sigma-mu transformation function is used to approximate the probability weighting function with parameters $\eta_s = 1.79$ and $\eta_m = 0.36$. These values are the optimal approximation of $\delta = 0.6$. Panel B presents the results where CEO is only CRRA but not probability-weighting. For both panels, γ is the risk aversion parameter, which takes values from 1 to 8. Both tables list the median and the mean value of the optimal fixed salary, stock holding and option holding. The column $n_o > 0$ is the fraction of the optimal contracts over the sample where CEO holds positive options of the firm[5]. The column $\phi > 0$ is the fraction of the optimal contracts where CEO receives positive fixed salary.

The risk-aversion of the typical CEOs are about 1 and 2. The results in Panel A shows that my probability-weighting model predicts 94.53% of positive option holding when γ equals to 1. The fraction of positive option holding decreases when γ increases. It converges to 1.77% when the CEO has a risk aversion parameter $\gamma \geq 5$. On the other hand, my model predicts the high fraction of positive fixed salary for optimal contracts for all γ. The fraction converges to above 99% when $\gamma \geq 5$. The results in Panel A shows that when the CEO is more risk-averse, the optimal piece-wise

[5]Here I define that the CEO a positive fraction of the firm's options if $n_o > 10^{-8}$.

Table 3: Calibration results for probability-weighting model and CRRA model

This table shows the optimal linear contract obtained by numerically solving Equations (22) to (24). Panel A presents the results where CEO is probability-weighting. The sigma-mu transformation function is used to approximate the probability function. $\eta_s = 1.78775$, $\eta_m = 0.35764$ are used in the calibration. These pair of values can be the optimal approximation of $\delta = 0.6$. Panel B presents the results where CEO is only CRRA but not probability-weighting. γ is the risk aversion parameter, which takes values from 1 to 8. Both tables list the median and the mean value of the optimal fixed salary, shares holding and option holding. The column $n_o > 0$ is the fraction of the optimal contracts over the sample where CEO holds positive options of the firm. The column $\phi > 0$ is the fraction of the optimal contracts where CEO receives positive fixed salary.

Panel A: Probability weighting model

		ϕ		n_s		n_o			
γ	obs.	median	mean	median	mean	median	mean	$n_o > 0$	$\phi > 0$
1	622	0.221	0.497	0.000	0.006	0.013	0.030	94.53%	95.82%
2	622	0.070	0.153	0.007	0.019	0.000	0.001	26.05%	84.08%
3	622	0.083	0.168	0.006	0.019	0.000	0.000	6.91%	93.73%
5	622	0.097	0.182	0.005	0.018	0.000	0.000	1.77%	99.36%
8	622	0.101	0.187	0.005	0.017	0.000	0.000	1.77%	99.84%

Panel B: CRRA model

		ϕ		n_s		n_o			
γ	obs.	median	mean	median	mean	median	mean	$n_o > 0$	$\phi > 0$
1	622	0.008	0.073	0.009	0.021	0.000	0.001	15.92%	53.38%
2	621	0.041	0.121	0.007	0.020	0.000	0.000	7.40%	71.22%
3	622	0.062	0.146	0.007	0.019	0.000	0.000	3.70%	84.08%
5	622	0.087	0.171	0.006	0.018	0.000	0.000	0.64%	93.73%
8	622	0.098	0.182	0.005	0.018	0.000	0.000	0.48%	98.87%

linear contracts become more "flat" in shape. It consists of fewer options and more fixed salary. The intuition is that CEO with a high level of risk aversion requires more "safe money" for his compensation packages.

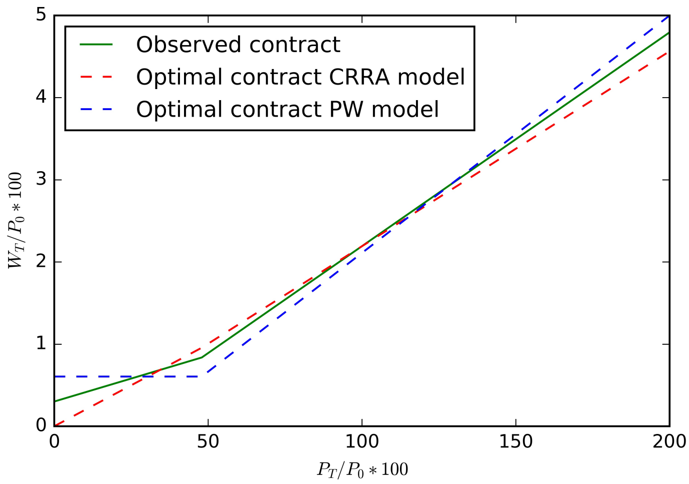
If we compare the probability weighting model in Panel with the traditional CRRA model in Panel B, we can see that the probability weighting model outperforms the CRRA model in predicting a positive option holding of CEOs for all γ. It reflects the convexity of the optimal piece-wise linear contracts. The probability weighting model explains the observed contracts particularly well when the risk-averse parameter γ is low. For example, the probability weighting model predicts 94.53% positive option holding when $\gamma = 1$, while the CRRA model only predicts 15.92% positive option holding.

Moreover, the mean and median option holding in the probability weighting model are higher than those in the CRRA model. The mean and median shares holding are smaller than those in the CRRA model. The median and mean fixed salary is higher in the probability weighting model. The results suggest that the optimal contracts contain more options and more fixed salary in the probability weighting setting.

Figure 4 gives an intuitive comparison between the probability weighting model and the CRRA model. The green curve depicts the shape of an observed contract with $r_f = 0.0197$, $d = 0.008$,

Figure 4: Shapes of contracts in different models

This table gives an comparison between the probability weighting model and the CRRA model. The green curve is the shape of the observed contract where $r_f = 0.0197$, $d = 0.008$, $s = 0.3579$, $T = 3.5813$, $\phi = 0.0763$, $n_s = 0.01086$, $n_o = 0.01486$, $K = 47.96$, and $\omega_0 = 0.2042$. The firm size at time 0. P_0 is re-scaled to 100. The red dashed curve is the optimal piece-wise linear contract with the CRRA model, while the blue dashed curve is the optimal piece-wise linear contract with the probability weighting model.



$s = 0.3579$, $T = 3.5813$, $\phi = 0.0763$, $n_s = 0.01086$, $n_o = 0.01486$, $K = 47.96$, and $\omega_0 = 0.2042$. The total value of the stocks and options for this contract is close to the mean value of the sample. The red dashed curve is the optimal piece-wise linear contract with the CRRA model, while the blue dashed curve is the optimal piece-wise linear contract with the probability weighting model. The observed contract features positive fixed salary and positive option holding. In the figure, the observed contract is convex and has a kink at the strike price of the options. The optimal piecewise linear contract with CRRA model contains only stock holding. The fixed salary and option holding are both non-positive. Thus, the shape of the optimal contract is a straight line. The optimal piece-wise linear contract with probability weighting model predicts positive fixed salary as well as positive option holding. So the shape of the optimal contract is convex and has a kink at the strike price.

5.2 Other values of probability weighting

To see whether my model is robust against the different level of probability weighting, I calibrate the model with different sets of η_s and η_m as approximations of various δ that takes values other than 0.6. I set δ to take values from 0.4 to 0.8. A lower δ means greater probability weighting. When δ is higher, there is the less probability weighting. Table 4 presents the calibration results. Panel A reports the case when δ is equal to 0.4, i.e., the greatest probability weighting, Panel B reports the case when δ is equal to 0.5, Panel C reports the case when δ is equal to 0.7, and Panel D reports the case where δ is equal to 0.8, i.e., the least probability weighting. We can see that the probability model predicts higher option holding than the CRRA model for almost all δ and γ. The only exceptions occur when $\gamma = 3$ and $\delta = 0.8$ or $\delta = 0.9$. This means that my probability weighting model can well explain the positive option holding in observed contracts, i.e., convex contracts, for a range of probability weighting parameters. Moreover, the probability weighting model outperforms the CRRA model in predicting positive fixed salary for all δ. This indicates that when the CEO is moderately probability weighting, my model can explain the observed contracts better than the CRRA in terms of both option holding and fixed salary.

5.3 Robustness checks

In the calibrations above, I use η_s and η_m to approximate the probability weighting parameter δ. Probability weighting exaggerates of the probability of the extreme outcomes and attaches higher dicision weights to them. Therefore, I am more interested in the effect of η_s in shaping the optimal contract. My parametrization allows me to disentangle the effect of a transformed mu, i.e., η_m, and the effect of a transformed sigma, i.e., η_s on the optimal contracts.

As a robustness check, I test whether my calibration results still hold when only the sigma is transformed. To be more specific, I keep $\eta_m = 0$ and use different η_s in the calibration, and see whether η_s alone can explain the observed positive option holding. Table 5 shows the calibration results when the parameter η_m is set to be 0, and the parameter η_s takes values 2 and 3. Panel A presents the results where $\eta_s = 2$ and $\eta_m = 0$. Panel B presents the results where $\eta_s = 3$ and $\eta_m = 0$. Compared with the CRRA model where $\eta_s = 1$ (Table 3, Panel B), we can see that, a larger-than-one η_s alone can explain positive option holding better than the CRRA model for all γ. For example, in the CRRA model, the fraction of positive option holding is 15.92% when $\gamma = 2$. In the contrast, the probability weighting model with $\eta_s = 2$ predicts a fraction of 20.58% and the model with $\eta_s = 3$ predicts a fraction of 27.81%. The probability weighting model also performs better in explaining larger fractions of positive fixed salary.

6. Conclusion

In this paper, I analyze a principal-agent model where the CEO is probability weighting and risk-averse. In order to solve for the closed-form optimal contracts, I use a sigma-mu transformation to approximate the probability weighting function. For each probability weighting parameter δ, I can always find a set of parameters η_s and η_m that transforms the original normal distribution into a new normal distribution with the similar shape. I derive the closed-form optimal general contract which exhibits convexity when firm performance is high. It provides theoretical evidence for the considerable number of options in CEO's compensation packages.

To see whether my model fits the observed contracts well, I then calibrate the model with a wide range of CRRA and Probability weighting parameters using the observed U.S. CEO contracts. I show that the model with probability weighting can explain the shape of the observed contracts better than the normal CRRA model. As a robustness check, I set the η_m to be 0 and only change the value of η_s. I find that η_s alone can also explain positive option holding.

My finding suggests that shareholders exploit probability weighting to provide cheap incentives that

Table 4: Calibration results for probability-weighting model with different parameters

This table shows the optimal linear contract obtained by numerically solving Equations 22) to (24) using the probability-weighting model. The sigma-mu transformation function is used to approximate the probability function. Panel A presents the results where $\eta_s = 2.88477$ and $\eta_m = 1.65702$ as the approximation of $\delta = 0.4$. Panel B presents the results where $\eta_s = 2.23307$ and $\eta_m 0.77908$ as the approximation of $\delta = 0.5$. Panel C presents the results where $\eta_s = 1.47895$ and $\eta_m = 0.15139$ as the approximation of $\delta = 0.7$. Panel D presents the results where $\eta_s = 1.26259$ and $\eta_m = 0.05263$ as the approximation of $\delta = 0.8$. γ is the risk aversion parameter. The γ used for calibration ranges from 1 to 8. All tables list the median and the mean value of the optimal fixed salary, shares holding and option holding. The column $n_o > 0$ is the fraction of the optimal contracts over the sample where CEO holds positive options of the firm. The column $\phi > 0$ is the fraction of the optimal contracts where CEO receives positive fixed salary.

Panel A: $\eta_s = 2.88477$ and $\eta_m = 1.65702$ as the approximation of $\delta = 0.4$

		ϕ		n_s		n_o			
γ	obs.	median	mean	median	mean	median	mean	$n_o > 0$	$\phi > 0$
1	622	0.100	0.000	0.019	0.268	0.001	0.043	99.68%	77.49%
2	622	0.096	0.006	0.000	0.177	0.019	0.004	45.34%	92.77%
3	622	0.091	0.006	0.000	0.177	0.018	0.001	14.31%	98.23%
5	622	0.099	0.005	0.000	0.185	0.018	0.000	3.54%	100.00%
8	622	0.101	0.005	0.000	0.188	0.017	0.000	2.57%	100.00%

Panel B: $\eta_s = 2.23307$ and $\eta_m 0.77908$ as the approximation of $\delta = 0.5$

		ϕ		n_s		n_o			
γ	obs.	median	mean	median	mean	median	mean	$n_o > 0$	$\phi > 0$
1	622	0.204	0.000	0.016	0.458	0.002	0.037	98.23%	93.89%
2	622	0.086	0.007	0.000	0.166	0.019	0.003	37.94%	89.07%
3	622	0.088	0.006	0.000	0.173	0.018	0.000	11.09%	97.11%
5	622	0.098	0.005	0.000	0.183	0.018	0.000	2.57%	99.84%
8	622	0.101	0.005	0.000	0.188	0.017	0.000	1.45%	100.00%

Panel C: $\eta_s = 1.47895$ and $\eta_m = 0.15139$ as the approximation of $\delta = 0.7$

		ϕ		n_s		n_o			
γ	obs.	median	mean	median	mean	median	mean	$n_o > 0$	$\phi > 0$
1	622	0.151	0.363	0.003	0.013	0.006	0.017	81.51%	84.73%
2	622	0.055	0.141	0.007	0.020	0.000	0.001	15.11%	80.87%
3	622	0.078	0.162	0.006	0.019	0.000	0.000	3.38%	91.32%
5	622	0.095	0.179	0.005	0.018	0.000	0.000	0.80%	99.20%
8	622	0.101	0.186	0.005	0.017	0.000	0.000	1.13%	99.84%

Panel D: $\eta_s = 1.26259$ and $\eta_m = 0.05263$ as the approximation of $\delta = 0.8$

		ϕ		n_s		n_o			
γ	obs.	median	mean	median	mean	median	mean	$n_o > 0$	$\phi > 0$
1	622	0.042	0.157	0.008	0.020	0.000	0.004	49.84%	64.79%
2	622	0.047	0.132	0.008	0.020	0.000	0.000	9.32%	76.05%
3	622	0.073	0.157	0.006	0.019	0.000	0.000	2.57%	88.42%
5	622	0.093	0.177	0.005	0.018	0.000	0.000	0.96%	97.43%
8	622	0.099	0.185	0.005	0.017	0.000	0.000	0.80%	99.52%

Table 5: Calibration results for probability-weighting model with different η_s

This table shows the optimal linear contract obtained by numerically solving Equations (22) to (24) using the probability-weighting model. The sigma-mu transformation function is used to approximate the probability function. γ is the risk aversion parameter. The γ used for calibration is from 1 to 8. Panel A presents the results where $\eta_s = 2$ and $\eta_m = 0$. Panel B presents the results where $\eta_s = 3$ and $\eta_m = 0$. All tables list the median and the mean value of the optimal fixed salary, shares holding and option holding. The column $n_o > 0$ is the fraction of the optimal contracts over the sample where CEO holds positive options of the firm. The column $\phi > 0$ is the fraction of the optimal contracts where CEO receives positive fixed salary.

Panel A: $\eta_m = 0$ and $\eta_s = 2$

		ϕ		n_s		n_o			
γ	obs.	median	mean	median	mean	median	mean	$n_o > 0$	$\phi > 0$
1	622	0.194	0.000	0.012	0.388	0.009	0.027	91.96%	94.69%
2	622	0.077	0.007	0.000	0.164	0.019	0.001	20.58%	90.19%
3	622	0.090	0.006	0.000	0.175	0.018	0.000	4.82%	97.59%
5	622	0.099	0.005	0.000	0.185	0.018	0.000	1.93%	99.84%
8	622	0.101	0.005	0.000	0.188	0.017	0.000	1.13%	100.00%

Panel B: $\eta_m = 0$ and $\eta_s = 3$

		ϕ		n_s		n_o			
γ	obs.	median	mean	median	mean	median	mean	$n_o > 0$	$\phi > 0$
1	622	0.184	0.000	0.016	0.327	0.004	0.035	98.55%	96.78%
2	622	0.094	0.007	0.000	0.177	0.019	0.002	27.81%	97.59%
3	622	0.096	0.006	0.000	0.183	0.018	0.000	7.72%	99.84%
5	622	0.101	0.005	0.000	0.187	0.018	0.000	2.25%	100.00%
8	622	0.102	0.005	0.000	0.189	0.017	0.000	1.29%	100.00%

encourage CEOs to exert more effort. It enriches the literature on the optimal CEO compensation contract and provides an alternative theory in explaining the observed options pay in CEO contracts. It is a complement to the loss-aversion theory (Dittmann, Maug, and Spalt, 2010) and the risk-seeking incentive theory (Dittmann, Yu, and Zhang, 2017).

My results can also be linked to a wide range of behavioral finance literature on CEO compensation. Malmendier and Tate (2005) find that overconfident CEOs exercise of their options late. Their results can be linked with the effect of a positive η_m in my paper. On the other hand, a manager is mis-calibrated (overprecise) if she underestimates the volatility of the future return (Ben-David, Graham, and Harvey, 2013). This is equivalent to an $\eta_s < 1$ in my paper. My results suggest that probability weighting effect may exceed the mis-calibration effect in shaping the CEO contracts. The sigma-mu transformation can be related, but not confused with CEO overconfidence. Overconfidence concerns with the CEO's *belief* over the future performance; the distribution of the future performance is clearly known by CEO herself. On the contrary, probability weighting is about *preference*, namely how CEO evaluate the expected utility. She attaches higher psychological weights to extreme outcomes.

In this paper, I do not endogenize project choice for CEOs in order to keep my model concise. Dittmann, Yu, and Zhang (2017) show that when CEOs can select projects, convex contracts provides risk-seeking incentives to CEOs. If probability-weighting CEOs can endogenously choose projects, then they might be even more risk-seeking. The reason is that probability weighting increases the volatility of a normal distribution. When a CEO chooses a project with higher risk σ, the perceived risk $\eta_s\sigma$ is propotionally higher. This increases the perceived value of stock options.

My research can be further extended in the following directions. First, in this paper, I calibrate the model with the optimal piecewise linear contracts. As an improvement, the general optimal non-linear contract can also be calibrated with the observed contracts.

Second, I use the sigma-mu transformation function to approximate the probability weighting function because it results in an closed-form solution for the optimal general contract. It also reduces the difficulties in the calibration. In the future, researches can be done in calibrating the model with the original probability weighting function. This means that the CDF of the random variable after the probability weighting transformation should be numerically calculated.

Third, I use the CRRA utility function instead of the loss-aversion utility function in my model. The reason is that the loss-aversion can explain the convexity of the observed contracts in the central region (Dittmann, Maug, and Spalt, 2010). In the future, the complete prospect theory, i.e., loss-aversion utility function and probability weighting, can be incorporated into the principal-agent model. This model may work well in explaining the convexity of the CEO contracts for the whole domain of the firm performance.

Appendix A: Proof of the convexity for large P_T

To show that the contract is convex when P_T is large, I will show that the slope of the contract tends to be infinitely large when P_T goes to infinity, i.e., $\lim_{P_T\to\infty}\frac{dW_T(P_T)}{dP_T}=\infty$ for all $\gamma>0$. The optimal general contract is:

$$W_T(P_T)=\left[\exp\left(\beta_2\left(\ln P_T\right)^2+\beta_1\ln P_T+\beta_0\right)\left(\alpha_1\ln P_T+\alpha_0\right)\right]^{1/\gamma}.$$

The first-order derivative of W_T w.r.t. P_T is

$$\begin{aligned}\frac{dW_T}{dP_T}=&W_T^{1-\gamma}\left[W_T^{\gamma}\left(2\beta_2\frac{\ln P_T}{P_T}+\beta_1\frac{1}{P_T}\right)+\exp\left\{\beta_2\left(\ln P_T\right)^2+\beta_1\ln P_T+\beta_0\right\}\frac{\alpha_1}{P_T}\right]\\ =&W_T\left(2\beta_2\frac{\ln P_T}{P_T}+\beta_1\frac{1}{P_T}\right)+W_T^{1-\gamma}\exp\left\{\beta_2\left(\ln P_T\right)^2+\beta_1\ln P_T+\beta_0\right\}\frac{\alpha_1}{P_T}\\ =&\left[\exp\left(\beta_2\left(\ln P_T\right)^2+\beta_1\ln P_T+\beta_0\right)\left(\alpha_1\ln P_T+\alpha_0\right)\right]^{1/\gamma}\left(2\beta_2\frac{\ln P_T}{P_T}+\beta_1\frac{1}{P_T}\right)\\ &+\left[\exp\left(\beta_2\left(\ln P_T\right)^2+\beta_1\ln P_T+\beta_0\right)\left(\alpha_1\ln P_T+\alpha_0\right)\right]^{\frac{1-\gamma}{\gamma}}\exp\left\{\beta_2\left(\ln P_T\right)^2+\beta_1\ln P_T+\beta_0\right\}\frac{\alpha_1}{P_T}\end{aligned}$$

All terms with lower powers (little-o) are omitted in calculation of the limit, e.g., $\ln P_T$ is omitted if $(\ln P_T)^2$ is in the presence. All positive coefficients, such as β_2 and α_1 are also omitted because they do not affect the evaluation. Note that $\beta_2=\frac{1}{2\sigma^2}-\frac{1}{2\sigma_A^2}=\frac{1}{2\sigma^2}\left(1-\frac{1}{\eta_s^2}\right)>0$. Thus, the limit of the first-order derivative of W_T w.r.t. P_T equals:

$$\lim_{P_T \to \infty} \frac{dW_T}{dP_T} = \lim_{P_T \to \infty} \left[\exp\left\{ (\ln P_T)^2 \right\} \ln P_T \right]^{1/\gamma} \left(\frac{\ln P_T}{P_T} \right) + \underbrace{\left[\exp\left\{ (\ln P_T)^2 \right\} \ln P_T \right]^{\frac{1-\gamma}{\gamma}} \exp\left\{ \beta_2 (\ln P_T)^2 \right\} \frac{1}{P_T}}_{positive}$$

The second component must be positive, so we focus only on the first component:

$$\begin{aligned} & \lim_{P_T \to \infty} \left[\exp\left\{ (\ln P_T)^2 \right\} \ln P_T \right]^{1/\gamma} \left(\frac{\ln P_T}{P_T} \right) \qquad \text{(Let } x = \ln P_T\text{)} \\ = & \lim_{x \to \infty} \frac{\left[\exp\left\{ x^2 \right\} x \right]^{1/\gamma} x}{\exp\{x\}} \\ = & \lim_{x \to \infty} \exp\left\{ \frac{\frac{1}{\gamma} \ln\left[\exp\left\{ x^2 \right\} \right] + \frac{1}{\gamma} \ln x + \ln x}{x} \right\} \\ = & \lim_{x \to \infty} \exp\left\{ \frac{\frac{1}{\gamma} x^2}{x} \right\} \\ = & \infty \end{aligned}$$

The slope of the $W_T(\cdot)$ tends to be infinite when P_T is infinitely large. This indicates a convex shape of the optimal general contract.

Chapter III

The real costs of CEO compensation: the effect of behindness aversion of employees

> "Wide pay gaps between CEOs and other employees are associated with higher employee turnover, which can adversely affect a company's performance and thereby shareowner interests."
>
> Investors and investor organizations collectively representing $3 trillion in assets under management in a letter to the SEC in support of the pay ratio disclosure.

1. Introduction

Some of the strongest opposition against high and increasing CEO pay comes from rank and file employees, in particular from employees within the same firm. It is difficult to explain this phenomenon with normative preferences, because formally CEOs are employees and paid by shareholders, so regular employees should, in principle, not object to them being overpaid. A potential explanation is that workers envy CEOs their higher pay, i.e., workers suffer disutility from the gap between their own pay and the CEO's pay. Formally, such preferences are called inequality aversion or, more precisely, behindness aversion (see Fehr and Schmidt (1999) and Neilson and Stowe (2010); for empirical evidence see Card, Mas, Moretti, and Saez (2012)).[6]

We provide a principal-agent model where the principal designs a contract with two agents: the CEO and the employee who is behindness averse. The employee represents all employees in the firm. We find that the wage of the employee is an increasing function of the wage of the CEO. The reason is that employees experience an additional disutility when the CEO pay is raised, so that the firm raises the employees' pay to compensate them for this disutility and to prevent them from leaving the firm. We take this prediction to the data and ask whether an employee truly compares himself to the CEO.[7]

There can be direct and indirect channels through which CEO compensation affects employee wages. Through the direct channel, workers observe the compensation of CEOs from published reports. They derive dis-utility directly from comparison. This means that workers near the bottom of the hierarchy are more sensitive to increases in the CEO wage because the wage gap is larger. Another

[6] It is widely accepted that the feeling of happiness does not only rely on someone's own material payoff, but also on the payoff of others. Schmitt and Marwell (1972) show that subjects withdraw from profitable experiments if they receive inequitable payoffs. Using data on British workers, Clark and Oswald (1996) show that the satisfaction levels of workers are negatively related to their comparison wage rates. Akerlof and Yellen (1990) show that the fair wage of workers is not only determined by the market clearing wage, but also by the comparison with salient others. Besides, Bolton and Ockenfels (2000) find that a simple model where someone's true payoff consists of her own pecuniary and own relative payoff explains many laboratory experiments.

[7] This paper adds another behavioral bias to Edmans, Gabaix, and Jenter (2017) who survey executive compensation.

channel is indirect, which can also be referred to as a trickle-down effect. Top managers compare their wages to the CEO, and their dis-utility is compensated. Regular managers compare their wages to top managers, lower managers to regular managers, and regular employees compare their wages to lower managers. The effect of high CEO compensation gradually passes down to regular employees. Akerlof and Yellen (1990) argue that a possible reference group to which employees can compare their wages are agents with a higher income within the firm which is consistent with the indirect channel.

The main challenge in studying the effect of high CEO compensation on workers' pay is the availability of data. We construct a matched CEO-employee panel data set for German firms by combining a data set on the CEO compensation with a data set on employee wages. Data on the CEO compensation is hand-collected from firms' annual reports. Data on employee wages comes from the Research Data Center (FDZ) of the German Federal Employment agency (BA). This agency has established a complete record of employee wages in German establishments since 1975 (for East Germany since 1992). The matched panel data set contains more than 200,000 establishment-year observations, and is available from 2000 to 2011. This unique data set enables us to test several hypotheses on the relationship of CEO compensation and employee wages.

We find evidence that higher CEO compensation is positively related to employee wages across firms and across time. When the CEO compensation increases by 1%, the median employee wage increases by 0.04%. This finding does not only hold in cross-sectional regressions but also when we control for time invariant unobserved characteristics of the firm and the establishment. To further alleviate potential endogeneity concerns, we adopt the difference-in-difference setting. In this analysis, we find that when CEO compensation becomes publicly observable, employees receive significantly higher wages. We also implement the triple-diff-in-diff approach where we find that a higher CEO-board wage gap results in a higher increase in employees' pay upon disclosure. Moreover, using CEO abnormal compensation, we show that paying more than the fair wage to CEOs increases employees' envy, while paying less than the fair wage to CEOs mitigates employees' envy.

The introductory quote ("Wide pay gaps between CEOs and other employees are associated with higher employee turnover, which can adversely affect a company's performance and thereby share-owner interests.") refers to Wade, O'Reilly, and Pollock (2006) who show that CEO overpayment is related to higher turnover for other managers (see also Bloom and Michel (2002)). What does the relation look like for rank-and-file workers? The investors from the introductory quote assume that employees are behindness averse, compare themselves to the CEO, and if the disutility becomes too large, draw the consequences and resign from the job. We do instead argue (with the principal-agent model) that the firm anticipates the behindness aversion of the employees, pays them a larger wage, and thereby prevents employee turnover. The data allows us to measure turnover. We show that an increased wage for the employees overcompensates for their behindness aversion and the employee turnover probability decreases in CEO pay. It is not surprising that highly paid CEOs might be able to drive down employee turnover.

These findings have far reaching consequences for executive compensation. Behindness aversion drives up the costs of executive compensation by increasing employee wages. Any additional dollar paid to the CEO for providing incentives also leads to higher employee wages to compensate employees for their (perceived) losses from behindness aversion. The average CEO in our sample receives €2.6 million a year. If a firm increases its pay by 1% (=€26,000) for the average CEO, then the firm will pay an additional compensation of €14.4 to the median employee with an average annual salary of €35,000. For the average firm in our sample with 50,000 employees, this sums up to €720,000 per year, increasing the total wage bill by €746,000 per year.

We show that regular employee wages rise with lagged CEO compensation. This could be explained by productivity dynamics and rent-extraction: In phases where productivity is high, the pressure on wages decreases and all wages are increasing. We do several tests in the paper to reduce this concern: First, we include ROA and market-to-book ratio as control variables which help capture changes in productivity. Second, we introduce industry × year and state × year fixed effects into the regression which filter out industry and state shocks. Third, changes in productivity cannot explain our differences-in-difference results.

The model closest to ours is Dur and Glazer (2008) who analyze the agency problem and optimal contracts when the employee feels envy toward his principal. They show that envy tightens the employees' participation constraint and causes higher pay or a lower workload. The authors also show that workers and firms can benefit from profit-sharing programs because they reduce the expected disutility from envy.

There exist a few empirical studies which examine the relation between CEO compensation and employee wages or productivity. Cronqvist, Heyman, Nilsson, Svaleryd, and Vlachos (2009) work with Swedish data and relate managerial entrenchment to the wages of regular employees. They find that CEOs with more control pay higher employee wages, especially for employees close to the CEO (geographically and hierarchically). They argue that CEOs derive private benefits from treating colleagues in their vicinity nicely. Wade, O'Reilly, and Pollock (2006) regress the CEO compensation on CEO's personal traits and firm variables, and use the residuals as a proxy for CEO over- or underpayment. They show that CEO overpayment is related to higher pay for other managers (see also Bloom and Michel (2002)). Faleye, Reis, and Venkateswaran (2013) calculate the wage gap between the CEO and rank-and-file employees. They investigate the determinants of the pay gap, but they fail to find any significant effect of an increased pay gap on employee productivity except for firms where the tournament incentives are high. This paper is related to the literature which tries to understand what types of firms exhibit more pay inequality (see, for example, Mueller, Ouimet, and Simintzi (2017)). Lin, Schmid, and Sun (2016) investigate how employee representation on corporate boards of German firms affects executive compensation. They find that more employee representation increases executive compensation and employment protection of workers. However, they do not investigate employee wages. Our paper is - to the best of our knowledge - the first to show that there exists a positive relation between CEO and rank-and-file employee pay, and we ascribe this relation to the behindness aversion of employees.

The paper is organized as follows: Section 2 discusses the principal-agent model. Section 3 presents the data. Section 4 documents the relation between CEO compensation and employee wages consistent with our employee-behindness-aversion hypothesis. We also present evidence for the causality of CEO compensation on employee wages and robustness checks in Section 4. Section 5 contains our employee turnover results, and Section 6 concludes the paper.

2. Principal-agent model

We model a principal (i.e., the firm) who contracts with two agents: the CEO and one employee. The employee represents all employees in the firm. The employee is behindness averse:

$$U\left(W_T^w, W_T^c\right) = V^w\left(W_T^w\right) - \alpha S\left(W_T^c - W_T^w\right), \tag{27}$$

where W_T^w is the employee's wage, W_T^c is the CEO's wage, $V^w(\cdot)$ is a risk-averse utility function, $S(\cdot)$ an inequality function, and α the inequality parameter.[8] The employee's effort is observable and contractible. Hence, employees only add a participation constraint to the principal's problem. The CEO is rational and risk-averse. Her effort is not observable, so she adds a participation constraint and an incentive compatibility constraint to the principal problem. Exerting effort e leads to private costs $C(e)$ that are increasing and convex in e.

The principal proposes a contract that is signed by the CEO and the worker at time $t = 0$. After that, the CEO makes her effort decision e. At time $t = T$, the consequences of the CEO's effort become apparent in the distribution of the firm's stock price $g(P_T \mid e)$. The principal maximizes:

$$\begin{aligned}
\max_{e, W_T^c(\cdot), W_T^w(\cdot)} \int_o^\infty & \quad \left(P_T - W_T^c - W_T^w\right) g\left(P_T \mid e\right) dP_T \\
s.t. \int_o^\infty & \quad V^c\left(W_T^c\right) g\left(P_T \mid e\right) dP_T - C(e) \geq \overline{U}^c \\
\int_o^\infty & \quad V^c\left(W_T^c\right) g_e\left(P_T \mid e\right) dP_T - C'(e) = 0 \\
\int_o^\infty & \quad \left[V^w\left(W_T^w\right) - \alpha S\left(W_T^c - W_T^w\right)\right] g\left(P_T \mid e\right) dP_T \geq \overline{U}^w,
\end{aligned}$$

where $V^c(\cdot)$ is the CEO's utility function and $\overline{U}^c$and $\overline{U}^w$ are the outside options of the CEO and the employee, respectively. In Appendix A, we prove the following Proposition:

Proposition 1: If $S(\cdot)$ is convex, the employee's wage increases with the CEO's wage:

$$\frac{dW_T^w}{dW_T^c} = \frac{\alpha S''\left(W_T^c - W_T^w\right)}{\alpha S''\left(W_T^c - W_T^w\right) - V^{w''}\left(W_T^w\right)} > 0.$$

When $S(.)$ is concave, then further assumptions are needed.

A limitation of our model is that it does not have an incentive compatibility constraint (IC) for employees. Grund and Sliwka (2005) and Neilson and Stowe (2010) feature an additional (IC) constraint for two identical agents that are inequality averse when analyzing tournament structures.

[8]The employee always earns less than the CEO, so $W_T^c - W_T^w$ never becomes negative. Therefore, we need not specify an extra parameter if the employee is ahead of the CEO. The results are the same notwithstanding if the employee is compassionate (i.e., dislikes being ahead) or competitive (i.e., likes being ahead).

They discriminate two different effects. On the one hand, an agent will work harder if she is envious (incentive effect). On the other hand, the more inequality averse the agent is, the more the principal needs to compensate the negative utility from inequality (participation effect). These papers find that agents with inequality aversion exert higher efforts than those who are purely self-interested under certain tournament structures. Faleye, Reis, and Venkateswaran (2013) find evidence for this result.

3. Data

The sample contains all companies included in the two main German stock market indices, DAX and MDAX, between 2000 and 2011. We hand collect data on executive compensation and corporate governance from annual reports and Hoppenstedt company profiles. We do not include non-listed firms, because information on executive compensation is usually unavailable. Stock market data comes from *Datastream* and balance sheet and accounting data from *Worldscope*.

3.1 Workers' compensation

Employment and wage data at the establishment level is obtained from the Institute of Employment Research (IAB). The IAB is the research organization of the German federal employment agency, the Bundesagentur für Arbeit (BA). The BA collects worker and employer contributions to unemployment insurance and distributes unemployment benefits. All German businesses are required to report detailed information on employment and wages to the BA.[9] Individual-level data is aggregated at the establishment level, made anonymous, and offered for scientific use by the IAB (the Establishment History Panel). An establishment is any facility having a separate physical address, such as a factory, service station, restaurant, or office building. The IAB offers detailed establishment level data on industry, location, employment, employee education, age, nationality, and wages, and provides this data in the form of establishment-level statistics, such as sums, medians, and quartiles on wages and employment according to different classifications and breakdowns.

IAB does not have a firm identifier, which is why manual matching is necessary. At our request, the IAB matched our sample of listed firms with their establishment-level database using an automatic procedure, based on company name and address information (city, zip code, street, and house number). Additionally, we provided the IAB with names of major subsidiaries listed in the annual reports of our sample firms in 2006. All cases not unambiguously matched by the automatic matching procedure are checked by hand to avoid mismatching. The matching was performed for 2004, 2005, and 2006. Firms are dropped if they do not exist during the period 2004 through 2006. All establishments are matched only once to our sample firms. This matching procedure does not

[9] German establishments are required to report salaries of their employees up to an upper earnings limit (social security contribution ceiling) that is annually adjusted (West German states: €52,800 in 2000 up to €66,000 in 2011. East German states: €43,600 in 2000 up to €57,600 in 2011). When this limit has been reached, establishments are only required to report the ceiling. In our data set, we delete 5.23% of the observations because the median average wage for the establishment was equal to the ceiling value for the respective year.

Table 6: Sample

This table displays the number of firms and establishments in the sample for each year between 2000 and 2011.

Year	Firms	Establishments
2000	35	3,486
2001	47	7,261
2002	59	8,329
2003	66	16,471
2004	98	20,814
2005	100	23,783
2006	99	25,767
2007	97	24,436
2008	95	21,310
2009	83	19,246
2010	84	16,924
2011	84	15,607

allow us to identify changes in establishment ownership after 2006.[10] Thus, if an establishment is acquired before 2004 or sold to another firm after 2006, it will be treated as if it belonged to the matched firm after the acquisition or before the sale. This will blur the match between firms and establishments and potentially lead to an attenuation bias working against finding significant results. Table 6 provides an overview of our matching process.

While fiscal years of German firms are mostly from January to December, establishment years for IAB data are from July to June. Therefore, we lead all variables from *Worldscope* by six months relative to IAB years. Effectively, we assign year-end values from *Worldscope* to June 30 information on employment and wages of the same year.

3.2 CEO compensation

We hand-collected data on compensation for CEOs and other members of the management board from firms' annual reports. Before 2006, most firms only disclose the total compensation of the management board as a whole. Only a few firms reported the individual compensation. From 2006, the German Corporate Governance Code required firms to disclose the individual compensation of members of the management board in their annual reports.[11] Hence, data on individual compensation for the management board is available for most firms after 2006. If a firm discloses the individual compensation, we record the payment for each executive, while for all other firms the total compensation for the management board is recorded.

[10] At the time of matching establishments to firms, establishment data was not available for 2007 and subsequent years.

[11] The German Corporate Governance Code (2006), Clause 4.2.4, requires that "The total compensation of each member of the Management Board is to be disclosed by name, divided into non-performance-related, performance-related, and long-term incentive components, unless decided otherwise by the General Meeting by three quarters majority." This means that the disclosure of the compensation of each member of the management board is mandatory from 2006 as long as the general annual meeting has not decided otherwise with three quarters majority. Compare that to the German Corporate Governance Code (2005), Clause 4.2.4, "Compensation of the members of the Management Board shall be reported in the Notes of the Consolidated Financial Statements subdivided according to fixed, performance related and long-term incentive components." According to the Code, the word "shall" is used as a recommendation but not a regulation.

Managerial compensation consists of several components: fixed salary, remuneration in kind, annual bonus, and compensation from long-term incentive programs. The long-term incentive programs include stock options, stock appreciation rights, and other stock based instruments. All these separate components of compensation are recorded if available. Our principle variable *CEO total* is the aggregate compensation that is mentioned in the Table "Board of Management Compensation - Aggregate Compensation" in the annual report. This is the aggregate value of the realized cash and bonus payments and the promised long-term compensation. It seems natural that the employees focus on this value because this is also usually reported in the press. Some executives in our panel data set join or leave the management board during the year. Their remuneration is then adjusted for the period in office to make them comparable to the standard annual compensation.

Table 7 presents summary statistics and variable definitions for firm-level variables (Panel A) and establishment-level variables (Panel B). The average firm year in our sample has sales of €15.8 billion, which shows that our sample mostly consists of large firms. The average CEO has a total annual compensation of €2.6 million and is 54 years old. The average median annual gross wage of full-time employees for our sample is €35,167.

3.3 Institutional setting

Historically, wages in German firms were mostly set through collective bargaining agreements between trade unions and employers' associations. However, in the last three decades, a major shift away from industry-level agreements has taken place. Hassel (1999) reports that in 1995, 53.4% of the plants were covered by industry-level wage agreements, 8.2% by firm-level agreements, and 38.4% were not covered at all. Although their sample may not be fully comparable to that of Hassel (1999), Addison, Bryson, Teixeira, Pahnke, and Bellmann (2010) report that only 47.3% of the German plants had industry-level agreements in 2000, a number that drops to 35.4% by 2008. Firm-level agreements were almost stable with 2.5% in 2000 and 2.7% in 2008, whereas the plants not covered by any collective bargaining agreement increased from 50.1% in 2000 to 61.9% in 2008. Over the same period, unionization also decreased considerably in Germany. Based on survey data, Schnabel and Wagner (2007) estimate it to be about 33% in 1992, declining to around 20% in 2004. If industry- or firm-level agreements exist, these agreements are binding for all workers as German law forbids discriminatory wage policies that disadvantage non-union members.

As a reaction to the declining popularity of collective bargaining agreements, trade unions and employers' associations are allowed so-called opening-clauses. Since the mid-1980s, labor regulation (including wage setting) has become increasingly flexible even for firms covered by collective bargaining agreements. Many areas of regulation are no longer determined at the industry level. Instead, works councils at the establishment level directly negotiate agreements with employers (Ellguth, Gerner, and Stegmaier (2012)). In particular, large firms (as in our sample) make use of these opening-clauses. Hassel and Rehder (2001) show that 55 of the 120 biggest companies in Germany negotiated a firm-level pact that deviates from the industry-level agreement during the 1990s.

However, even if firms do not use an opening-clause, they are free to deviate from the collective bar-

Table 7: Summary statistics

Panel A displays definitions and descriptive statistics for the main firm-level variables used in our analysis. Panel B displays definitions and descriptive statistics for the main establishment-level variables used in our analysis.

Panel A: Summary statistics for main firm-level variables

Variable name	Definition	Mean	Std.	Obs.
Compensation				
CEO total	Annual total compensation of the CEO	2,564,779	2,395,251	555
CEO cash	Annual cash income of the CEO	2,002,491	1,657,132	555
Board total	Average annual compensation for all management board members: total compensation for the board / board size	1,411,901	1,096,341	939
Other total	Average annual total compensation for management board members excluding the CEO	1,421,626	1,073,914	554
CEO premium	ln(CEO total - Other total)	13.43	1.21	524
CEO pay ratio	CEO total / Other total	1.85	1.29	554
CEO-board ratio	CEO total / Board total - 1	0.48	0.64	555
CEO characteristics				
CEO tenure	Time since first appointed as the CEO (year)	6.41	6.00	536
CEO ownership	=1 if the CEO holds more than 1% of the firm outstanding shares	0.01	0.07	551
CEO switch	=1 if another person takes over the CEO position	0.09	0.29	555
CEO age	Age of CEO (in years)	54.05	6.91	527
CEO out-hiring	=1 if the CEO is hired from outside the firm	0.43	0.50	536
Firm-level characteristics				
ROA	Return on asset	0.10	0.12	910
ROE	Return on equity	0.34	0.30	910
Market to book ratio	Market to book ratio	2.33	2.45	931
Size (millions)	Total sales of the firm	15,844	27,976	924
Leverage	Total debt / total asset	0.63	0.20	932
# Firm employees	Number of employees working for the firm in Germany	49,899	90,643	934
Employee risk	Standard deviation of change in number of employees at the firm level	0.13	0.09	935
Union	=1 if one of the major German labor unions has representatives on the firm's supervisory board	0.95	0.22	939
Disclosure	=1 if the compensation of the CEO is disclosed in annual reports	0.59	0.49	939
Stock return	Total annual stock return calculated using the return index provided by Datastream	0.15	0.49	838
Board size	Number of members on the executive board	4.74	2.11	939
R&D to sales	R&D to sales ratio	9.58	42.07	623

Panel B: Summary statistics for main establishment-level variables

Variable name	Definition	Mean	Std.	Obs.
Wage structure				
Wage	Median gross average daily wage for full-time employees × 365	35,167	13,428	158,545
Q1 wage	First quartile gross average daily wage for full-time employees × 365	31,678	12,554	163,531
Q3 wage	Third quartile gross average daily wage for full-time employees × 365	37,301	13,967	142,865
Employee structure				
# Establishment employees	Total number of full-time employees at the establishment	64.79	691.72	203,434
Female %	Proportion of full-time female employees	0.43	0.36	167,296
Low qualified %	Proportion of full-time low-qualified employees	0.04	0.12	167,296
Qualified %	Proportion of full-time median-qualified employees	0.73	0.33	167,296
Highly qualified %	Proportion of full-time high-qualified employees	0.08	0.19	167,296
German %	Proportion of German employees	0.97	0.10	167,296
Manager %	Proportion of managers	0.03	0.13	167,296
White-collar %	Proportion of white-collar workers	0.61	0.46	167,296
Employees age	Median age of full-time employees at the establishment level	41.46	8.29	203,434
Other variables				
Close to head	=1 if the establishment is located in the same federal state as the firm's headquarter	0.18	0.38	203,434
Outflow	Outflow of employees$_t$ / # Establishment employees$_{t-1}$	0.21	0.22	76,616
Outflow white-collar	Outflow of white-collar employees$_t$ / # Establishment employees$_{t-1}$	0.14	0.22	76,616
Inflow	Inflow of employees$_t$ / # Establishment employees$_{t-1}$	0.42	5.68	76,616
Inflow of white-collar	Inflow of white-collar employees$_t$ / # Establishment employees$_{t-1}$	0.29	4.33	76,616
Industry	2-digit NACE code (economic division) of the the establishment (edition: 2003)			
State	Federal state where the establishment is located			

gaining agreement as long as they pay wages above the level stipulated in the agreement. Collective bargaining agreements only determine minimum standards. Jung and Schnabel (2011) show that more than 43% of the establishments covered by a collective agreement pay wages above the level stipulated in the collective agreement. For these 43% of the establishments, average actual wages exceed wages that were stipulated by the collective bargaining agreement by about 10%. Both numbers increase with the size of an establishment, i.e. positive deviations are more likely for the large firms in our data set. Taken together, these studies show that wage setting is rather flexible (in both directions) at the firm level in Germany.

4. The relation between CEO compensation and employee wages

4.1 Baseline results

We start by analyzing the relation between CEO compensation and employee wages using the following baseline regression model:

$$ln(Wage)_{ijt} = \alpha_t + \alpha_k + \alpha_s + \beta ln(CEO\ total)_{jt-1} + \gamma X_{ijt-1} + \varepsilon_{ijt} \qquad (28)$$

The dependent variable, $ln(Wage)_{ijt}$, is the logarithm of the median annual wage in establishment i and year t, where j indexes firms. $ln(CEO\ total)_{jt-1}$ is the logarithm of the CEO's total compensation over the prior year $t-1$. In our benchmark regressions, we control for year fixed effects, α_t, industry fixed effects of the establishment, α_k, and state fixed effects, α_s. X_{ijt-1} is a vector of control variables, which include establishment-level variables such as number, median age, qualifications, and the nationality of employees, and firm-level variables such as profitability, size, leverage, CEO ownership, and tenure. All explanatory variables are lagged by one year. We run fixed effects regressions and use White (1980) robust standard errors that allow for clustering at the firm level.

Table 8 presents our results. Specification (1) only includes industry, state, and year fixed effects. The following specifications slowly build the full model. First, adding establishment level controls in specification (2) and then stepwise firm-level controls (specifications (3) to (5)). Across all specifications, we observe that firms that pay their CEOs more also pay significantly higher wages to their other employees. In specifications (3) to (5), we also include ROA and market-to-book ratio which control for firm productivity.

In specification (6), which includes observations after 2005 and the full set of control variables, the coefficient for *ln(CEO total)* is 0.041 (t = 2.93). This result means that if CEO compensation increases by 1%, the median employee's wage increases by 0.04%. This effect is economically sizable. The average CEO in our sample receives €2.6 million a year. If a firm increases its pay by 1% (=€26,000) for the average CEO, then the firm will pay an additional compensation of €14.4 to the median employee with an average annual salary of €35,000. For the average firm in our sample with 50,000 employees that sums up to €720,000 per year, this increases the total wage bill by €746,000 per year.

Table 8: CEO compensation and employee wages: Regression results

This table presents results for regressions with the log median annual wage of full-time employees as the dependent variable. All independent variables are lagged by one year. See Table 7 for a detailed overview of variable definitions. In specification (6), we consider the observations after 2005 only. We use the White (1980) robust standard errors clustered at the firm level. The t-statistics are reported below the estimates. ***, ** and * indicate that the value is significantly different from zero at the 1%, 5% and 10% levels.

Dependent variable:	ln(Wage)					
	(1)	(2)	(3)	(4)	(5)	(6)
ln(CEO total)	0.051**	0.046**	0.039***	0.037***	0.038***	0.041***
	2.14	2.22	3.6	3.71	3.71	2.93
ROA			-0.303*	-0.16	-0.109	-0.219*
			-1.82	-1.2	-0.81	-1.87
Price to book ratio			-0.019*	-0.020*	-0.024**	-0.025**
			-1.81	-1.97	-2.16	-2.13
ln(Size)			0.009	-0.03	-0.021	-0.039
			0.5	-1.17	-0.83	-1.42
Leverage			-0.133	-0.195*	-0.268*	-0.279*
			-1.31	-1.83	-1.92	-1.81
Union				0.121*	0.127*	0.153**
				1.91	1.93	2.32
ln(# Firm employees)				0.038*	0.032	0.049**
				1.78	1.56	2.15
Employee risk				0.485**	0.505**	0.609**
				2.25	2.42	2.62
CEO ownership					-0.009	0.007
					-0.19	0.14
CEO tenure					0.003*	0.004**
					1.84	2.07
ln(# Estab. Employees)		0.045***	0.044***	0.044***	0.044***	0.050***
		3.65	3.67	3.73	3.72	4.19
Female %		-0.280***	-0.262***	-0.274***	-0.274***	-0.277***
		-3.73	-3.68	-3.7	-3.67	-3.54
low qualified %		-0.011	-0.013	-0.011	-0.008	-0.018
		-0.19	-0.24	-0.2	-0.15	-0.33
Qualified %		0.220***	0.220***	0.213***	0.212***	0.212***
		5.48	5.79	5.84	5.78	5.25
Highly qualified %		0.504***	0.513***	0.509***	0.510***	0.503***
		11.2	12.49	13.37	13.41	12.9
German %		0.256***	0.252***	0.251***	0.255***	0.250***
		5.33	5.25	5.38	5.43	5.05
Manager %		0.017	0.024	0.04	0.043	0.047
		0.63	0.88	1.52	1.52	1.44
White collar %		0.167***	0.161***	0.175***	0.176***	0.190***
		7.71	6.85	8.33	8.36	9.1
Employee age		0.004**	0.004**	0.004**	0.004**	0.004**
		2.49	2.55	2.6	2.58	2.21
Close to head		0.031**	0.028**	0.029**	0.030**	0.031**
		2.41	2.23	2.2	2.24	2.48
Industry, state, and year FE	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.52	0.618	0.61	0.613	0.615	0.622
Number of observations	108363	106341	103961	103960	103581	68356

These results are obtained after controlling for observable characteristics known to influence employee wages. In particular, we control for establishment and firm size, employee characteristics, profitability, leverage, and union presence. As expected, employee wages are higher when employees are better educated, older, German, male, work in larger establishments or firms, have a higher risk of losing their jobs, the leverage is lower, a union member has a board seat, and the establishment is close to the headquarter (see, for example, Cronqvist, Heyman, Nilsson, Svaleryd, and Vlachos (2009), Brown and Medoff (1989)). In an untabulated robustness check, we adjust all nominal variables for inflation and find very similar results.

4.2 Difference-in-difference regression

The German regulation on mandatory disclosure of the CEO compensation was publicly discussed in 2003 and 2004, enacted by the federal parliament in 2005 and became effective in 2006. This regulation does not directly affect employee wages, but it changes the channel through which the employees observe their firm's CEO compensation. Before the regulation came into effect, the employees could generally only observe the aggregate remuneration of all members on the management board. After the regulation was adopted by the firms, the employees have been able to directly observe the CEO compensation, which is on average 44% more than the compensation of an average management board member in 2006. Under the employee-behindness-aversion hypothesis, we expect that employees in those firms that disclose their CEO compensation for the first time feel more disadvantaged and are paid more. Thus, we regard the change in policy as a natural experiment.

In the difference-in-difference setting, we select the firms that disclose their CEO compensation before 2003 (i.e., {2000, 2001, 2002}) as the control group. And we regard those firms that do not disclose their wage before 2003 but do disclose it in the year 2006 as the treatment group.[12] The independent variable *Treatment* equals 1 when an observation is in the treatment group, and *Post-2006* equals 1 when the year is in or after 2006. Table 9 presents the results in specifications (1) and (2). The coefficients on *Treatment* × *Post-2006* in both specifications are statistically significant at the 1% level. The results are also economically significant. When firms are required to make their CEO compensation publicly observable, they pay 11.5% higher wages to their employees. This value may seem large as compared to our baseline regression, where we found a coefficient on *ln(CEO pay)* of 0.041. The most likely explanation for this difference is the selection bias.

Selection bias: One concern about the difference-in-difference setting is the assumption of a random formation of the treatment and control groups. Before 2006, firms could choose whether to disclose their CEO compensation or not. From 2006, firms are required to disclose their CEO compensation, unless otherwise decided by the annual general meeting by a three quarters majority. Therefore, our difference-in-difference method might suffer from a potential selection bias: firms that did not expect any strong effects from publishing CEO salaries on employee wages might have self-selected into the control group and disclosed individual salaries before this was required by

[12]Our analysis in 8 assumes that CEO compensation is always available and discards the observations without CEO compensation. The difference-in-difference regression uses all of the data notwithstanding if CEO compensation is available or not.

Table 9: Difference-in-difference and triple-diff-in-diff regressions

This table presents results for regressions in a difference-in-difference setting (specifications 1 and 2) and a triple-diff-in-diff setting (specifications 3 and 4) with the log median annual wage of full-time employees as the dependent variable. All independent variables are lagged by one year. See Table 7 for a detailed overview of variable definitions. The control group contains the firms which disclose the compensation of the CEO before 2003 (i.e., {2000, 2001, 2002}). The treatment group contains the firms which do not disclose the compensation of the CEO before 2003, but do disclose it in the year 2006. The independent variable *Treatment* equals 1 when an observation is in the treatment group. *Post*-2006 equals 1 when the year is in or after 2006. *CEO-board ratio* is the percentage that the CEO earns more than the board average compensation. We use the White (1980) robust standard errors clustered at the firm level. The *t*-statistics are reported below the estimates. ***, ** and * indicate that the value is significantly different from zero at the 1%, 5% and 10% levels.

Dependent variable:	ln(Wage)			
	(1)	(2)	(3)	(4)
Treatment × Post-2006 × CEO-board ratio			0.303***	0.348***
			2.92	3.03
Treatment × Post-2006	0.124***	0.115***	0.000	-0.021
	3.26	3.28	0.01	-0.36
Treatment × CEO-board ratio			-0.279***	-0.307***
			-3.32	-3.35
Post-2006 × CEO-board ratio			-0.170**	-0.229***
			-2.39	-2.65
Treatment	-0.029	-0.037	0.075***	0.075**
	-1.04	-0.99	2.77	2.23
CEO-board ratio			0.232***	0.273***
			3.33	3.48
ROA	-0.124	-0.143	-0.088	-0.108
	-1.05	-1.3	-0.59	-0.78
Price to book ratio	-0.025**	-0.020**	-0.034***	-0.028***
	-2.61	-2.22	-3.35	-2.94
ln(Size)	0.024*	0.009	0.022	0.015
	1.86	0.32	1.49	0.48
Leverage	-0.452***	-0.432***	-0.420***	-0.398***
	-3.53	-3.2	-3.15	-2.91
Union		0.108*		0.124*
		1.88		1.87
ln(#Firm Employees)		0.006		-0.006
		0.19		-0.18
Employee risk		0.172		0.127
		0.7		0.51
ln(#Branch Employees)	0.053***	0.053***	0.057***	0.057***
	4.55	4.59	4.21	4.22
Female %	-0.382***	-0.387***	-0.379***	-0.384***
	-9.21	-9.51	-8.29	-8.73
Low qualified %	-0.008	-0.008	-0.014	-0.014
	-0.11	-0.12	-0.19	-0.2
Qualified %	0.196***	0.191***	0.238***	0.240***
	4.71	4.77	3.94	4
Highly qualified %	0.510***	0.509***	0.511***	0.509***
	12.62	12.92	10.36	10.7
German %	0.222***	0.222***	0.194***	0.188***
	4.9	4.93	4.27	4.35
Manager %	0.05	0.053	0.071**	0.074**
	1.61	1.64	2.24	2.27
White-collar %	0.176***	0.180***	0.160***	0.163***
	5.38	5.36	4.98	4.95
Employee age	0.003*	0.003*	0.003	0.003
	1.77	1.8	1.46	1.49
Close to head	0.017	0.018	0.018	0.019
	1.3	1.38	1.36	1.42
Time, region, industry dummies	Yes	Yes	Yes	Yes
Adjusted R^2	0.619	0.62	0.631	0.632
Number of observations	82751	82751	65517	65517

the regulation. This leaves those firms which expected stronger effects on employee wages for the treatment group. Therefore, the estimated 11.5% treatment effect is probably overestimating the average effect on firms. However, the null-hypothesis is that employee wage and CEO compensation are independent from one another and this independence is clearly rejected.

We provide two additional tests to support our conjecture that CEO compensation influences employee wages.

Triple diff-in-diff: Under the behindness aversion hypothesis, employees' envy should be increasing in the wage gap between CEO and other management board members. We expect that the increase in employees' wage is positively related to the CEO-board wage ratio upon disclosure of the CEO compensation. To test this presumption, we adopt a difference-in-difference-in-difference (triple diff-in-diff) approach. *CEO-board ratio* is the percentage by which CEO compensation exceeds the average management board compensation for a given year. Table 9 presents the results in specifications (3) and (4). The coefficients on *Treatment* × *Post-2006* × *CEO-board ratio* are both statistically significant at the 1% level. So we indeed find that, after the new disclosure regulation has become effective, employee wages increase more in firms with a relatively higher CEO pay. The results are also economically significant. When firms start disclosing their CEO compensation publicly after the regulation change, they pay 0.35% higher wages to their employees if the *CEO-board ratio* increases by one percentage point.

Parallel trends: If our presumption that the disclosure of CEO pay led to an increase in employee wages is correct, we would expect no significant differences between the treatment and the control group before 2006 (parallel trends assumption) and an increase in the difference between both groups afterwards. We include yearly interaction effects with the treatment dummy in our regression. All independent variables are lagged by one year, so we lose the year 2011. We use 2010 as our base year. Table 10 presents the results. The coefficients on the yearly interaction effects become significantly different from zero after 2007, i.e., *Treatment* × *2007*, *Treatment* × *2008*, and *Treatment* × *2009* are significant at the 1% level. Over the years 2007 to 2009, the coefficients increase and the results become more significant. This might imply that the increase in workers' wages is rather gradual. Moreover, the insignificant coefficients on yearly interaction effects from 2000 to 2004 (except 2002) imply that the parallel-trend assumption holds.

The difference-in-difference analysis also helps us answer the question of whether the positive relationship between CEO pay and workers' wage is driven by workers' envy or by the CEO's compassion. Because the CEO always knows the wage of workers, the disclosure of the CEO pay does not affect the CEO's compassion towards normal workers. In contrast, workers do not always know the CEO pay, so the disclosure of the CEO pay will increase the workers' envy towards the CEO. The results from a difference-in-difference analysis confirm that the increase in the workers' pay is driven by the workers' envy.

In sum, we interpret these results as strong evidence for a causal effect of CEO compensation on employee wages consistent with the existence of the employee's behindness aversion. These diff-in-diff results cannot be explained by the production dynamics hypothesis which states that the pressure on wages decreases if productivity is high.

Table 10: Difference-in-difference with yearly interaction terms

This table presents results for regressions in a difference-in-difference setting with yearly interaction terms. The dependent variable is the log median annual wage of full-time employees. All independent variables are lagged by one year. See Table 7 for a detailed overview of variable definitions. The control group contains the firms which disclose the compensation of the CEO before 2003 (i.e., {2000, 2001, 2002}). The treatment group contains the firms which do not disclose the compensation of the CEO before 2003, but do it disclose in the year 2006. The independent variable *Treatment* equals 1 when an observation is in the treatment group. *Post*-2005 equals 1 when the year is in or after 2005. We use the White (1980) robust standard errors clustered at the firm level. The *t*-statistics are reported below the estimates. ***, ** and * indicate that the value is significantly different from zero at the 1%, 5% and 10% levels.

Dependent variable:	ln(Wage)	
	(1)	(2)
Treatment × 2000	0.042	0.051
	0.84	0.99
Treatment × 2001	-0.084*	-0.065
	-1.8	-1.3
Treatment × 2002	-0.109**	-0.103**
	-2.32	-2.29
Treatment × 2003	-0.076	-0.069
	-1.23	-1.2
Treatment × 2004	-0.034	-0.029
	-0.79	-0.68
Treatment × 2005	0.029	0.035
	0.9	1.02
Treatment × 2006	0.043	0.045*
	1.6	1.67
Treatment × 2007	0.088***	0.087***
	3.43	3.14
Treatment × 2008	0.113***	0.111***
	4.97	4.94
Treatment × 2009	0.147***	0.144***
	6.14	5.56
Treatment	0.015	0
	0.39	-0.01
ROA	-0.114	-0.132
	-0.97	-1.22
Price to book ratio	-0.025***	-0.020**
	-2.67	-2.28
ln(Size)	0.024*	0.009
	1.86	0.3
Leverage	-0.451***	-0.431***
	-3.5	-3.17
Union		0.107*
		1.87
ln(#Firm Employees)		0.007
		0.2
Employee risk		0.174
		0.71
Establishment variables	Yes	Yes
Industry, state, and year FE	Yes	Yes
Adjusted R^2	0.619	0.62
Number of observations	82751	82751

4.3 Unobservables and fixed effects

A specific concern might be that the relationship between CEO compensation and employee wages is driven by firm-level or establishment-level unobservables. We address this concern using three approaches in Table 11. First, we include firm fixed effects instead of industry fixed effects. Second, we include both firm and industry fixed effects, which is possible since the industry differs across establishments. Finally, we control for establishment fixed effects. These tests are demanding on the data because the wages of both the CEO and the workers are rather sticky. The results are nevertheless reassuring. While we lose economic significance, coefficients of *ln(CEO total)* are, on average, about 50% smaller; however, the statistical significance remains intact.

Another concern may be that there are CEO-level unobservables. We investigate this possibility by adding CEO and CEO-firm fixed effects. Once more, we observe a reduction in economic significance but the statistical significance is largely unaffected.

Finally, we might be concerned that the reason why CEO compensation affects employees' wages might be driven by unobservable time-varying factors at the industry level, the firm level, or the state level. We include industry × year fixed effects, firm × year fixed effects, and state × year fixed effects. We find that neither economic nor statistical significance is materially affected. This analysis suggests that time-varying industry level, firm level and state level unobservables are not inducing our results.

4.4 Abnormal CEO compensation

It is possible that employees feel more envy towards CEOs who receive an abnormally high compensation as compared to similar peers. To address this concern, we use CEO abnormal compensation as an additional explanatory variable. CEO abnormal compensation is defined as the difference between actual and expected CEO compensation. Our hypothesis is: If the CEO abnormal compensation is positive, i.e., the CEO earns more than what she deserves to get, the employee is more behindness averse. If the CEO abnormal compensation is negative, i.e., the CEO earns less than what she deserves to get, the employee is less behindness averse.

The analysis takes three steps. In the first step, we calculate the CEO expected compensation. We adopt the model used by Gillan, Hartzell, and Parrino (2009). The CEO expected total compensation is predicted by regressing the log CEO total compensation on log firms' total assets, the ratio of EBIT to assets (ROA), the ratio of assets to firm value (book-to-market ratio), CEO tenure, as well as year and industry (2-digit SIC) fixed effects. In the second step, we calculate the log CEO abnormal compensation. The log CEO abnormal total compensation is the difference between the actual log total compensation and the expected log total compensation. In the third step, we regress log workers' median wage on log CEO abnormal compensation and our standard set of control variables. Table 12 shows a statistically significant effect (1% level) of CEO abnormal total compensation on employee wages. A 1% increase of *CEO abnormal total* results in a 0.03% increase in the median employee's wage.

Table 11: Different sets of fixed effects

This table presents results for regressions with the log median annual wage of full-time employees as the dependent variable. All independent variables are lagged by one year. See Table 7 for a detailed overview of variable definitions. The table displays results for regressions with different sets of fixed effects. The unreported control variables are the same as in specifications (1), (2), (3), (4), and (5) of Table 8. We use the White (1980) robust standard errors clustered at the firm level. The t-statistics are reported below the estimates. ***, ** and * indicate that the value is significantly different from zero at the 1%, 5% and 10% levels.

Dependent variable:	ln(Wage)				
	(1)	(2)	(3)	(4)	(5)
Firm, year and state fixed effects					
ln(CEO total)	0.019**	0.018**	0.015**	0.016**	0.016**
	2.14	2.45	2.45	2.45	2.36
Adjusted R^2	0.401	0.553	0.550	0.550	0.550
Number of observations	108,589	106,538	104,150	104,149	103,770
Firm, industry, year and state fixed effects					
ln(CEO total)	0.018*	0.017*	0.017**	0.017**	0.017**
	1.91	1.93	2.17	2.10	2.16
Adjusted R^2	0.566	0.646	0.638	0.638	0.638
Number of observations	108,363	106,341	103,961	103,960	103,581
Establishment and year fixed effects					
ln(CEO total)	0.015**	0.016**	0.013*	0.013**	0.013*
	2.18	2.26	1.98	2.00	1.96
Adjusted R^2	0.922	0.926	0.927	0.927	0.927
Number of observations	108,589	106,538	104,150	104,149	103,770
CEO, year and state fixed effects					
ln(CEO total)	0.021**	0.022***	0.020***	0.020***	0.020***
	2.03	2.70	3.25	3.21	3.22
Adjusted R^2	0.401	0.553	0.550	0.550	0.550
Number of observations	108,547	106,496	104,108	104,107	103,770
CEO × firm, year and state fixed effects					
ln(CEO total)	0.021**	0.022***	0.020***	0.020***	0.020***
	2.03	2.70	3.25	3.21	3.22
Adjusted R^2	0.401	0.553	0.550	0.550	0.550
Number of observations	108,589	106,538	104,150	104,149	103,770
Industry × year and state fixed effects					
ln(CEO total)	0.053**	0.048**	0.039***	0.037***	0.037***
	2.04	2.07	3.23	3.29	3.27
Adjusted R^2	0.524	0.620	0.612	0.615	0.617
Number of observations	108,363	106,341	103,961	103,960	103,581
State × year and firm fixed effects					
ln(CEO total)	0.018*	0.017**	0.015**	0.016**	0.015**
	1.97	2.31	2.3	2.36	2.27
Adjusted R^2	0.404	0.554	0.552	0.552	0.552
Number of observations	108,589	106,538	104,150	104,149	103,770

Table 12: CEO abnormal total compensation

This table presents results for regressions with the log median annual wage of full-time employees as the dependent variable. The independent variable is the logarithmic CEO abnormal total compensation. We measure the CEO abnormal total compensation using the method adopted by Gillan et al. (2009). The logarithmic CEO abnormal total compensation is the difference between the CEO actual logarithmic compensation and the expected logarithmic total compensation calculated by regressing the logarithmic CEO total compensation on logarithmic firms' total assets, the ratio of EBIT to assets (ROA), the ratio of assets to firm value (book-to-market ratio), CEO tenure, the two-digit SIC of the firm, and the year of the observation. All independent variables are lagged by one year. The establishment variables are the same as in specification (5) of Table 8. See Table 7 for a detailed overview of variable definitions. We use the White (1980) robust standard errors clustered at the firm level. The *t*-statistics are reported below the estimates. ***, ** and * indicate that the value is significantly different from zero at the 1%, 5% and 10% levels.

Dependent variable:	ln(Wage)		
	(1)	(2)	(3)
ln(CEO abnormal total compensation)	0.028***	0.027**	0.033***
	3.16	2.36	3.89
ROA	-0.317*	-0.179	-0.107
	-1.81	-1.29	-0.79
Market to book ratio	-0.018	-0.018*	-0.023**
	-1.62	-1.82	-2.04
ln(Size)	0.022	-0.016	-0.006
	1.27	-0.61	-0.23
Leverage	-0.119	-0.204*	-0.261*
	-1.11	-1.82	-1.88
Union		0.126*	0.121*
		1.93	1.84
ln(# Firm employees)		0.037*	0.031
		1.7	1.44
Employee risk		0.561**	0.550***
		2.61	2.71
CEO ownership			-0.006
			-0.12
CEO tenure			0.004*
			1.99
ln(# Establishment employees)	0.044***	0.044***	0.044***
	3.64	3.7	3.72
Female %	-0.262***	-0.274***	-0.275***
	-3.66	-3.66	-3.66
low qualified %	-0.016	-0.012	-0.01
	-0.29	-0.22	-0.19
Qualified %	0.220***	0.213***	0.212***
	5.8	5.83	5.81
Highly qualified %	0.515***	0.512***	0.511***
	12.31	13.21	13.37
German %	0.255***	0.253***	0.256***
	5.31	5.44	5.48
Manager %	0.022	0.04	0.043
	0.79	1.47	1.5
White collar %	0.159***	0.174***	0.176***
	6.71	8.31	8.43
Employee age	0.004**	0.004**	0.004**
	2.54	2.57	2.59
Close to head	0.028**	0.029**	0.029**
	2.24	2.21	2.22
Industry, state, and year FE	Yes	Yes	Yes
Adjusted R^2	0.609	0.613	0.614
Number of observations	103664	103663	103581

Table 13: Timing and alternative measures of CEO compensation

This table presents results for regressions with the log median annual wage of full-time employees as the dependent variable. Panel A displays results for regressions with different time lags for independent variables: (1) no lag, (2) all independent variables are lagged by 1 year (baseline specification), (3) all independent variables are lagged by 2 years, (4) all independent variables are lagged by 3 years, (5) only *ln(CEO total)* is lagged by 1 year, (6) only *ln(CEO total)* is lagged by 2 years. Panel B displays results for regressions when alternative measures of executive compensation are used as independent variables. All independent variables are lagged by one year. See Table 7 for a detailed overview of variable definitions. We use the White (1980) robust standard errors clustered at firm level. The t-statistics are reported below the estimates. ***, ** and * indicate that the value is significantly different from zero at the 1%, 5% and 10% levels.

Panel A: Different time lags

Dependent variable:	ln(Wage)					
	(1)	(2)	(3)	(4)	(5)	(6)
ln(CEO total)	0.021*					
	1.68					
ln(CEO total) (t-1 for all)		0.038***				
		3.71				
ln(CEO total) (t-2 for all)			0.032***			
			3.36			
ln(CEO total) (t-3 for all)				0.039***		
				3.65		
ln(CEO total) (t-1)					0.025**	
					2.27	
ln(CEO total) (t-2)						0.022**
						2.06
Firm variables	Yes	Yes	Yes	Yes	Yes	Yes
Establishment variables	Yes	Yes	Yes	Yes	Yes	Yes
Industry, state, and year FE	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.617	0.615	0.606	0.598	0.624	0.62
Number of observations	125173	103581	84815	67029	107880	89112

4.5 Timing and alternative measures of CEO compensation

In order to better understand the relationship between CEO and employee compensation, we analyze different time lags of CEO total compensation. The most salient measure of CEO compensation for employees should be the total compensation from the last fiscal year, because that number is published during year t. The hypothesis is: If the firm anticipates the behindness aversion of its employees and therefore offers them an increased wage proactively, $ln(CEO\ total)_{t-1}$ will have the largest impact. If instead lengthy negotiations between employees and the firm take place, then a higher order lag of CEO compensation may be more relevant. Table 13 Panel A shows exactly the hypothesized result. $ln(CEO\ total)_{t-1}$ exhibits the highest t-statistic independent of whether we use lagged or contemporaneous control variables. In fact, the contemporaneous CEO compensation only has a marginally significant impact on employee wages. It is consistent with the idea that the firm anticipates the behindness aversion of its rank-and-file workers and that no lengthy negotiations are needed.

In a second step, we analyze the impact of different measures of executive compensation. If the correlation between CEO and employee compensation were to be mainly driven by unobservables (e.g., some dimension of profitability not captured by our other controls, i.e., *ROA*, or *Market to book ratio*), we would expect a similar correlation between average board or other executive compensation and employee wages. However, as shown by Panel B of Table 13, this is not the case. Other executives than the CEO (*ln(Other total)*) are insignificantly related to employee wages. CEO compensation has significantly more explanatory power than alternative measures of executive compensation. The explanatory power even increases if we use the CEO premium (i.e., the difference between CEO and average other executive compensation).

4.6 Subsamples

In Table 13, we have seen that employee wages react more strongly to increases in CEO compensation than, for example, average management board compensation. However, before 2006, firms were not required to publish management board compensation individually. Even today, the German Corporate Governance Code still allows that management board compensation is not disclosed at the individual level, if the annual general meeting approves the non-publication with a three-quarter majority. This means that employees cannot observe the CEO's compensation for many German firms before 2006 and for some after 2006. If inequality aversion is indeed the driver for the reason that CEO compensation affects employees' wages, we expect two effects from the regulatory change in 2006: (1) for firms that do not disclose management compensation individually before and after 2006, the impact of *ln(Board total)* on employee wages is largely unchanged and similar to the impact of *ln(CEO total)* for firms that disclose CEO compensation; (2) for firms that disclose management compensation individually after 2006, *ln(Board total)* becomes insignificant. That is exactly what we observe in our subsample analysis in Table 14. Only looking at firms that do not disclose management compensation individually (column (1)), we find a positive and significant coefficient on *ln(Board total)*. The economic effect is cut by more than half and statistical significance disappears, if we look at the sample of firms disclosing management compensation individually (column (2)). Both effects are even more pronounced for the non-disclosing firms before 2006 (column 4) and the disclosing firms after 2006 (column 5). This result is also confirmed when we use both measures *ln(other total)* and *ln(CEO total)* in the same regression (columns (3) and (6)). These findings suggest that employees benchmark their own salaries towards the most salient management compensation figure available. If compensation is disclosed individually, employees seem to only compare their wage to the CEO's compensation but not to that of other executives. If the CEO compensation is not available, the closest proxy, average management board compensation, is used as a benchmark. In sum, these empirical patterns lend strong support to the hypothesis that behindness aversion is an important driver in setting wages for rank-and-file employees.

4.7 Wage changes

To further test whether our model is robust, we ask whether the increases in employee wages are associated with the increases in CEO pay. The regression of changes on changes removes the effect of time-invariant unobservables. The change of the independent variable *ln(CEO total)* is the annual growth rate in CEO total compensation, and the change of the explanatory variable *ln(Wage)* is the annual growth rate in employees' wages. In order to rule out the possibility that the increase in employees' wages is driven by fast growing establishments, we drop the observations where the

Table 14: Subsample analysis

This table presents results for regressions with the log median annual wage of full-time employees as the dependent variable. The "No disclosure" sample includes all establishment-year observations of firms not disclosing the individual CEO compensation in a given year, i.e., the sample consists of firms that only disclose the aggregated compensation of all members in the management board in a given year. The "Disclosure" sample includes all establishment-year observations of firms disclosing the individual CEO compensation in a given year. Since 2006, the German Corporate Governance Code requires firms to disclose the individual compensation of all management board members. The German Corporate Governance Code still allows the firm not to disclose management board compensation individually, if the firm's annual general meeting approves the non-publication with a three-quarter majority. All independent variables are lagged by one year. The firm variables and establishment variables are the same as in specification (5) of Table 8. See Table 7 for a detailed overview of variable definitions. We use the White (1980) robust standard errors clustered at the firm level. The *t*-statistics are reported below the estimates. ***, ** and * indicate that the value is significantly different from zero at the 1%, 5% and 10% levels.

Dep. variable:	ln(Wage)					
Sample:	No disclosure all	Disclosure all	Disclosure all	No disclosure before 2006	Disclosure after 2006	Disclosure after 2006
	(1)	(2)	(3)	(4)	(5)	(6)
ln(board total)	0.058**	0.021		0.075**	0.016	
	2.15	1.64		2.56	0.88	
ln(CEO total)			0.099***			0.111***
			3.88			3.82
ln(other total)			-0.080***			-0.101***
			-2.78			-2.8
Firm variables	Yes	Yes	Yes	Yes	Yes	Yes
Establishment variables	Yes	Yes	Yes	Yes	Yes	Yes
Industry, state, year FE	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.569	0.612	0.614	0.569	0.62	0.622
Number of obs.	21287	103960	103960	18746	68442	68442

yearly growth rate of the number of employees in an establishment is above the 95% percentile. Specifications (2) to (5) in Table 15 show that the coefficients on *CEO total increase* are still statistically significant at the 10% level. The coefficient can be interpreted as: when the annual growth rate of the CEO total compensation is increased by 1 percentage point, then the annual growth rate of the employees' wages will increase by 0.004 percentage points.

4.8 Additional controls

While we control for a number of variables in the regressions in Table 8, other, potentially unobservable, variables may be driving our results. To minimize any such concerns, we report further results with additional firm-level controls in Table 16. First in column (1), we add the firm's annual *Stock return* as an alternative measure of firm performance. The results show that employees' wages are hardly influenced by stock returns above and beyond what is already captured in our other control variables. The statistical and economic significance of the coefficient on *ln(CEO total)* is not affected. Second, we add *Board size*, defined as the number of members on the executive board. Once more, we do not observe any significant effect on employees' wages or the coefficient on *ln(CEO total)*. Third, we add additional CEO characteristics: (1) *CEO switch* equals one if a new CEO is appointed in t-1; (2) *CEO age* in years; (3) *CEO out-hiring* equals one if the CEO is recruited from outside the firm. None of these CEO characteristics has a significant influence on employee pay. Fourth, we add *ln(R&D to sales)*. The results in Table 16 show that firms with higher R&D expenditures (relative to sales) pay lower employee wages. However, the coefficient on *ln(CEO total)* is hardly affected even though we lose more than 60% of the observations. Fifth, we split up the dummy variable *Union* into four dummy variables to control separately for the influence of the four largest German unions (IG Metall, ver.di, IG BCE, and IG BAU). We find that firms with one of the four largest unions on their supervisory board pay their employees more on average but there is no significant difference between these four unions. The coefficient for *ln(CEO total)* is not materially affected.

5. Employee turnover

Wade, O'Reilly, and Pollock (2006) show that CEO overpayment is related to a higher turnover for other managers (see also Bloom and Michel (2002)). What does the relation look like for rank-and-file workers? One important task of the CEO is to keep the employee turnover low because excessive turnover can result in shareholder value losses. We test this hypothesis using the employee inflow/outflow data provided by the IAB. We define two employee turnover variables: (1) *Outflow* as $Outflow\ of\ employees_t/\#Employees_{t-1}$ and (2) *Inflow* as $Inflow\ of\ employees_t/\#Employees_{t-1}$.

Table 17 presents the results. We observe that *Outflow* and *Inflow* are negatively correlated with CEO compensation. This result holds for all employees and for the subsample of white-collar employees. However, it is only significant for *Outflow* at the 5% level. This finding implies that employees are, on average, overcompensated for their behindness aversion, thus leading to a reduction

Table 15: Increase in CEO compensation and increase in employee wages

This table presents results for regressions with the annual increase in the median annual wage of full-time employees as the dependent variable. All independent variables are lagged by one year. See Table 7 for a detailed overview of variable definitions. The variable *CEO total increase* is the annual increase in CEO total compensation. We only use the observations where the yearly changes of the numbers of employees at the establishment level is below the 95% percentile. We use the White (1980) robust standard errors clustered at the firm level. The t-statistics are reported below the estimates. ***, ** and * indicate that the value is significantly different from zero at the 1%, 5% and 10% levels.

Dependent variable:	Wage increase				
	(1)	(2)	(3)	(4)	(5)
CEO total increase	0.005	0.004*	0.004*	0.004*	0.004*
	1.61	1.68	1.74	1.73	1.68
ROA			-0.04	-0.032	-0.026
			-1.42	-1.27	-1.11
Market to book ratio			0.002	0.002	0.001
			1.15	1.45	0.92
ln(Size)			0.001	0.000	0.001
			0.83	0.04	0.33
Leverage			0.008	0.003	0.006
			0.53	0.18	0.3
Union				0.014*	0.011
				1.98	1.59
ln(# Firm employees)				0.000	-0.001
				0.04	-0.17
Employee risk				0.044	0.042
				1.13	1.09
CEO ownership					0.020*
					1.99
CEO tenure					0.000
					0.67
ln(# Estab. employees)		-0.002***	-0.002***	-0.002***	-0.002***
		-4.74	-4.83	-5.1	-5.17
Female %		0.026***	0.024***	0.023***	0.023***
		3.6	3.14	3.03	2.97
Low qualified %		0.023*	0.022*	0.023*	0.023*
		1.88	1.89	1.84	1.88
Qualified %		-0.010***	-0.009***	-0.010***	-0.010***
		-3.58	-3.76	-3.85	-3.84
Highly qualified %		-0.015**	-0.016**	-0.016**	-0.016**
		-2.17	-2.27	-2.3	-2.32
German %		-0.025***	-0.023**	-0.024**	-0.024**
		-2.83	-2.59	-2.58	-2.6
Manager %		-0.029***	-0.030***	-0.029***	-0.028***
		-4.78	-4.51	-4.5	-4.54
White-collar %		0.007***	0.008***	0.009***	0.010***
		3.02	2.86	3.5	4.27
Employee age		-0.001***	-0.001***	-0.001***	-0.001***
		-4.29	-4.32	-4.35	-4.21
Close to head		-0.000	-0.000	-0.000	-0.000
		-0.07	-0.19	-0.12	-0.06
Industry, state, and year FE	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.002	0.005	0.005	0.006	0.006
Number of observations	85,660	85,660	85,546	85,546	85,348

Table 16: Robustness checks with additional controls

This table presents results for regressions with the log median annual wage of full-time employees as the dependent variable using additional control variables: (1) Stock return, (2) board size, (3) additional CEO characteristics, (4) R&D to sales, and (5) union variables. All independent variables are lagged by one year. See Table 7 for a detailed overview of the variable definitions. We use the White (1980) robust standard errors clustered at the firm level. The t-statistics are reported below the estimates. ***, ** and * indicate that the value is significantly different from zero at the 1%, 5% and 10% levels.

Dependent variable:	ln(Wage)				
	(1)	(2)	(3)	(4)	(5)
ln(CEO total)	0.038***	0.039***	0.036***	0.034**	0.039***
	2.99	3.91	3.23	2.03	4.21
Stock return	-0.008				
	-0.58				
Board size		0.005			
		0.80			
CEO switch			0.002		
			0.14		
CEO age			0.001		
			0.40		
CEO out-hiring			-0.016		
			-0.65		
ln(R&D to sales)				-0.019**	
				-2.29	
IGBAU					0.179**
					2.57
IGBCE					0.109*
					1.77
IGMetall					0.115
					1.48
Verdi					0.133*
					1.81
Firm variables	Yes	Yes	Yes	Yes	Yes
Establishment variables	Yes	Yes	Yes	Yes	Yes
Industry, state, and year FE	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.605	0.615	0.615	0.496	0.615
Number of observations	100,112	103,581	103,574	40,952	103,581

in turnover because the outside options are relatively less attractive.

6. Conclusion

In this paper, we document a strong positive effect of CEO compensation on the wages of rank-and-file employees. This pattern is not explained by established determinants of employee wages and is unlikely to be caused by unobservables at the industry, firm, CEO, establishment, or state level. Difference-in-difference analysis, triple-diff-in-diff, and the analysis of CEO abnormal compensation suggest a causal interpretation of our findings. The evidence is most consistent with firms paying higher wages to their employees in order to compensate them for the disutility caused by the pay gap to the CEO. The most likely driver underlying this phenomenon is behindness aversion of employees. This paper also shows that a highly paid CEO keeps the employee turnover low.

An interesting implication would be a cross-country comparison. In countries that experience higher inequality aversion the executive pay (all else equal) should be lower. Gabaix and Landier (2008) test this using the World Value Survey. However, they have only 17 observations and find insignificant results. A promising research project might be to have a larger data set that enables the researcher to sufficiently control for all known effects on pay levels, e.g., firm size.

Our evidence of employee behindness aversion implies that managerial compensation incurs additional "inequality costs", which need to be taken into consideration when determining the optimal contracts for both CEOs and employees. Any additional dollar paid to the CEO for providing incentives also leads to higher employee wages to compensate employees for their (perceived) losses from behindness aversion. These costs must be taken into account by shareholders to arrive at the real cost of CEO compensation.

Appendix A: Proof of the Proposition 1

The Lagrangian is:

$$\begin{aligned}
\mathcal{L} \;=\; & \int_o^\infty \left(P_T - W_T^c - W_T^w\right) g\left(P_T \mid e\right) dP_T \\
& + \lambda_{PCC}\left(\int_o^\infty V^c\left(W_T^c\right) g\left(P_T \mid e\right) dP_T - C\left(e\right) - \overline{U}^c\right) \\
& + \lambda_{ICC}\left(\int_o^\infty V^c\left(W_T^c\right) g_e\left(P_T \mid e\right) dP_T - C'\left(e\right)\right) \\
& + \lambda_{PCW}\left(\int_o^\infty \left[V^w\left(W_T^c\right) - \alpha S\left(W_T^c - W_T^w\right)\right] g\left(P_T \mid e\right) dP_T - \overline{U}^w\right)
\end{aligned}$$

Table 17: Employee turnover

This table presents results for regressions with different employee turnover variables as dependent variables. All independent variables are lagged by one year. See Table 7 for a detailed overview of the variable definitions. We use the White (1980) robust standard errors clustered at the firm level. The t-statistics are reported below the estimates. ***, ** and * indicate that the value is significantly different from zero at the 1%, 5% and 10% levels.

Dependent variable:	Outflow	Outflow white-collar	Inflow	Inflow white-collar
	(1)	(2)	(3)	(4)
ln(CEO total)	-0.023**	-0.016**	-0.041	-0.045
	-2.59	-2.61	-0.58	-0.66
ROA	-0.028	-0.038	-1.18	-1.123
	-0.44	-0.8	-1.04	-1.06
Price to book ratio	0.007	0.006*	0.106	0.108
	1.39	1.84	1.05	1.07
ln(Size)	0.008	-0.005	0.019	0.067
	0.68	-0.63	0.14	0.53
Leverage	-0.094	-0.062*	-1.277	-1.274
	-1.64	-1.76	-1.55	-1.57
Union	0.018	0.009	0.14	0.142
	0.82	0.48	0.61	0.63
ln(# Firm employees)	0.011	0.021**	0.092	0.03
	1.43	2.62	0.74	0.27
Employee risk	-0.158	-0.008	0.607	0.674
	-1.41	-0.08	0.74	0.84
CEO ownership	0.015	0.046*	-0.155	-0.203
	0.4	1.77	-0.47	-0.63
CEO tenure	-0.001	-0.001	-0.003	-0.002
	-1.05	-1.17	-0.54	-0.41
ln(# Estab. Employees)	-0.041***	-0.029***	-0.307***	-0.245***
	-22.7	-9.08	-4.48	-3.51
Female %	-0.043	-0.008	-0.336***	-0.242***
	-1.59	-0.9	-4.51	-3.26
Low qualified %	0.01	0.031*	0.234	0.095
	0.4	1.97	0.66	0.38
Qualified %	-0.031***	-0.01	0.021	-0.007
	-3.16	-0.84	0.12	-0.04
Highly qualified %	-0.028	-0.068***	1.154**	1.049**
	-1.56	-3.91	2.45	2.31
German %	-0.155***	-0.069***	-0.846*	-0.702
	-3.56	-4.8	-1.77	-1.46
Manager %	-0.007	-0.004	0.453	0.484
	-0.27	-0.18	0.73	0.81
White collar %	0.055	0.325***	0.174	0.314***
	1.6	23.51	1.49	2.69
Employee age	-0.002*	-0.001	0.007*	0.006*
	-1.78	-0.99	1.98	1.91
Close to head	0.006	0.005	0.185	0.18
	1.34	1.35	1.36	1.35
Establishment and year FE	Yes	Yes	Yes	Yes
Adjusted R^2	0.131	0.395	0.015	0.014
Number of observations	56922	56922	57601	57601

To use the implicit function theorem, we define

$$G(W_T^c, W_T^w) = \frac{d\mathcal{L}}{dW_T^w} = -1 + \lambda_{PCW}\left[V^{w\prime}\left(W_T^w\right) + \alpha S'\left(W_T^c - W_T^w\right)\right].$$

$$\begin{aligned} \frac{\partial G}{\partial W_T^w} &= \lambda_{PCW}\left[V^{w\prime\prime}\left(W_T^w\right) - \alpha S''\left(W_T^c - W_T^w\right)\right] \\ \frac{\partial G}{\partial W_T^c} &= \lambda_{PCW}\alpha S''\left(W_T^c - W_T^w\right) \\ \Rightarrow \frac{dW_T^w}{dW_T^c} &= -\frac{\partial G/\partial W_T^c}{\partial G/\partial W_T^w} \\ &= \frac{\alpha S''\left(W_T^c - W_T^w\right)}{\alpha S''\left(W_T^c - W_T^w\right) - V^{w\prime\prime}\left(W_T^w\right)} \end{aligned}$$

This expression is positive if $S(\cdot)$ is convex, which proves Proposition 1.

Chapter IV

Wage gap and stock returns

1. Introduction

In recent years, pay inequality between managers and workers has received increasing attention from academics, regulators, and the media[13]. Notably, on October 17, 2015, the Securities and Exchange Commission adopted a new rule, effective from January 1, 2017, that requires U.S. companies to disclose the ratio of CEO pay to the median employee wage. In this regard, on May 11, 2016, the New York Times pointed out that *"(t)he strong case for the rule (...) keeps getting stronger"*, providing support to the idea that corporations should *"rein in"* the difference in pay between managers and workers.

Interestingly, the populist anger that meets high executive premia does not seem to be justified by the facts at hand, at least on purely economic grounds. Recent research shows that high pay inequality does not hinder firm performance (Cronqvist, Heyman, Nilsson, Svaleryd, and Vlachos, 2009), and actually seems to improve it (Faleye, Reis, and Venkateswaran, 2013; Mueller, Ouimet, and Simintzi, 2017), which is in line with the conjecture that a larger wage gap reflects higher managerial skills. The stock market, however, does not seem to price this information correctly. In a study on British companies, Mueller, Ouimet, and Simintzi (2017) show that stocks with low pay inequality yield negative risk-adjusted returns, and argue that this might be due to the inability of arbitrage forces to observe the wage gap. In this paper, we shed new light on this issue, both theoretically and empirically.

We propose an asset pricing model with unsophisticated traders and short-sales constraints, in which the optimal wage gap between the CEO and rank-and-file workers increases with managerial effort. In equilibrium, we show that firms with low wage gaps should be overpriced. Using a unique data set on German firms' employee compensation, we provide strong support for the model's predictions. We find that a long-short portfolio of stocks with high and low wage gaps, respectively, yields positive and robust risk-adjusted returns. We also show that the results are partly driven by the fact that some investors exhibit inequality-averse preferences. The empirical evidence suggests that high wage gap firms are indeed better managed, but investors do not fully incorporate this information into the stock price.

In our theoretical analysis, we consider a three-period economy. At time zero, the representative firm appoints a manager to hire workers and carry out a project. The firm engages in efficient contracting, and chooses the optimal level of managerial effort by solving the following trade-off: high effort costs more, but improve both the firm's productivity, through a more efficient employment of resources, and the outcome of wage negotiations with workers, through a tougher attitude in bargaining. At time one, the firm seeks funding on the stock market to pay workers' wages and start production.

[13] See, e.g., Graef (1991), Pfeffer and Langton (1993), Main, O'Reilly, and Wade (1995), Bertrand and Mullainathan (2000), Bertrand and Mullainathan (2001), Bloom and Michel (2002), and Wade, O'Reilly, and Pollock (2006).

At time two, the firm liquidates.

Investors can be either sophisticated (arbitrageurs) or unsophisticated (noise traders), where the latter face short-sales constraints[14]. Arbitrageurs use the wage gap as a signal for managerial effort, and have unbiased evaluations. Noise traders, instead, do not recognize the informativeness of the wage gap. Their evaluation of low wage gap stocks is then above the fundamental value, and vice versa for high wage gap stocks. However, the presence of short-sales constraints only enables them to take long positions on low wage gap stocks, thereby inflating prices. In the final period, sophisticated investors arbitrage the mispricing away, generating lower returns. The error correction should be particularly pronounced for stocks with the lowest wage gaps.

In the empirical analysis, we take these predictions to the data. The major hurdle to overcome in this respect is the lack of publicly available data on rank-and-file workers' compensation. In the U.S., disclosure of workers' wages before January 1, 2017, was only discretionary, with low coverage. In this paper, we overcome this issue by using the "Establishment History Panel" database, maintained and made available by the German Federal Employment Agency. This is a unique data set that contains the annual gross wage for all workers in Germany from January 2001 to December 2011, together with information on the establishment they work for.[15]

Given the nature of workers' compensation data, one important point to address is whether investors are able to observe the wage gap. In Germany, we argue that this is possible through at least three channels. First, companies are under the scrutiny of labor unions, who release information on compensation schemes. Second, companies with extreme wage gaps are subject to extensive media coverage, which constitutes public information (see, e.g., Wade, O'Reilly, and Pollock (2006)). Third, some sophisticated investors may have preferential access to firm-specific information, including managers and workers' pay.

To construct our test portfolios, we divide stocks into pay inequality quantiles. Then, we rebalance them at the beginning of each year, and define high and low wage gap portfolios, respectively, as the stocks that lie at the top and the bottom 30%, 20%, and 10% of the distribution. We define the dependent variable as monthly excess returns on the top, bottom, and top-minus-bottom portfolios, calculated during the year following portfolio formation. To estimate risk-adjusted returns, we run time series regressions of the CAPM and Carhart (1997) four-factor model, using the European risk factors from Kenneth French's website.

We find strong support for the model's predictions. In the European four-factor model, we find that the top 30% wage gap stocks earn an insignificant alpha of 0.10% per month, while the bottom 30% wage gap stocks yield negative and significant risk-adjusted returns of 0.90% per month. This results in a profitable long-short investment strategy. Consistent with the hypothesis that the error correction is larger for extreme wage gaps, we find that the long-short strategy becomes more profitable when considering the 20% portfolios (1.30% per month) and the 10% portfolios (1.80%

[14]One way to make this distinction practical is to think of them as hedge funds and mutual funds, respectively (Chen, Hong, and Stein, 2002).

[15]While the panel is publicly available, employees of the German Federal Employment Agency provided an anonymous match between 100 large firms and establishments, making it accessible only to us.

per month). This empirical pattern is robust to a number of empirical specifications.[16].

To provide further evidence that it is indeed unsophisticated investors that drive the results, we identify companies that are difficult to evaluate and/or arbitrage. Notable examples are those with low market capitalization and extreme book-to-market ratios (Baker and Wurgler, 2006; Baker and Wurgler, 2007). With lower arbitrage activity, we expect the mispricing of the wage gap to be particularly pronounced among these stocks. Consistent with the conjecture, we find that the negative risk-adjusted returns of low wage gap stocks are entirely concentrated around small stocks (i.e., below-median market capitalization) and stocks with extreme book-to-market ratios (i.e., top and bottom 30%).

Next, we acknowledge the fact that some individuals may exhibit inequality aversion Fehr and Schmidt (1999). We conjecture that this type of preferences could also characterize investor behavior, and then affect the evaluation of the wage gap. A recent Eurosif study suggests that this might indeed be the case for German investors.[17] First, Germany is one of the European countries that invest the most in SRI stocks (only second to France), and labor rights constitutes one of the dimensions through which firms are evaluated. Second, there is large participation of institutional investors with a focus on societal values. Pension funds account for nearly half (48%) of all assets under management invested in SRI, followed by religious institutions and charities (24%), and foundations (12%). In light of this, it is plausible that some of these investors may indeed be inequality-averse, either to cater to their clientele's preferences, or because they may actually be prone to such a bias themselves.[18]

To look into this hypothesis from a theoretical perspective, we introduce inequality-averse traders in the model. We define them as an additional type of investors who dislike pay inequality within firms, and therefore exhibit what we refer to as an inequality-aversion bias. In particular, they acknowledge the economic meaning of the wage gap, but also consider non-monetary arguments when investing (e.g., society's welfare), which leads them to evaluate low wage gap firms in a more favorable way than they do high wage gap firms. If more investors exhibit this bias, the overpricing of low wage gap stocks increases.

The presence of inequality-averse traders has three testable implications. In our model, an increase in noise trader demand increases the overpricing of low wage gap stocks, due to binding short-sales constraints (see, e.g., Stambaugh, Yu, and Yuan (2012)). On the other hand, however, it also dilutes the impact of inequality-averse demand, which leads to less overpricing. These two effects are in opposite directions, and the latter dominates if the inequality-aversion bias is large enough.

Our first test, then, is to analyze the relation between noise trader demand and the returns on our arbitrage strategy. We identify times of large noise trader demand as those in which beginning-

[16]We find similar results when considering the German four-factor model, cash rather than total CEO compensation, the wage gap of the other members of the board excluding the CEO, and all establishments rather than only those located in the same state as the company headquarters. We also form the test portfolios in April rather than January, to acknowledge the fact most companies issue their reports at the end of the first quarter, and find analogous results. Finally, we estimate panel regressions from Brennan, Chordia, and Subrahmanyam (1998), and find again a positive relation between wage gap and stock returns.

[17]http://www.eurosif.org/wp-content/uploads/2016/11/SRI-study-2016-HR.pdf.

[18]See, e.g., Hong and Kostovetsky (2012), and DeVault, Sias, and Starks (2014), for a discussion of biases in institutional investor behavior.

of-period investor sentiment is high (Yu and Yuan, 2011; Stambaugh, Yu, and Yuan, 2012). We capture the level of sentiment of German investors by considering the consumer confidence index for Germany. We find that low wage gap stocks are overpriced only in times of low sentiment, while the mispricing disappears in times of high sentiment. For high wage gap stocks, risk-adjusted returns continue to be close to zero in times of both high and low sentiment. As a result, arbitrage returns are higher when noise trader demand is low. This pattern lends support to the inequality-aversion story.

Second, we argue that if some investors exhibit a bias related to pay inequality, then the effect of such a bias should be stronger when the wage gap is salient, i.e., when it is disclosed by firms and covered by the media. This would be a peculiar trait of inequality-averse traders, because arbitrageurs are unbiased, and noise traders simply neglect the wage gap at all times during the year. To test for this, we identify April as the month in which the wage gap is typically disclosed and covered by the press. The increase in attention towards pay inequality should then lead inequality-averse investors to trade more. Consistent with the conjecture, we find that our long-short strategy yields positive risk-adjusted returns only in the second quarter, and mostly in the month of April.

Third, we argue that if inequality aversion is a genuine behavioral bias, then it should not be related to economic fundamentals. In particular, it should not matter whether high CEO pay is justified or not by the manager's contribution to the firm's value. To test this hypothesis, we decompose CEO pay into an economically justified and an unjustified component, following Rouen (2017), where the justified component is a linear combination of firm characteristics. As noise traders neglect the wage gap but observe the other fundamentals of the firm, their pricing error should be confined to the unjustified wage gap. The inequality-aversion bias, instead, should apply to both the justified and the unjustified wage gap. Our results show that the mispricing is symmetrically present in both specifications of CEO pay, which lends support to the inequality-aversion hypothesis.

Our story hinges on the assumption that there is a positive relation between the wage gap and managerial effort. High pay inequality, however, might also be the result of rent extraction. In this case, the fundamental value of high wage gap firms is lower than that of low wage gap firms. Since unsophisticated investors do not distinguish between these two types, firms with a high wage gap should exhibit overpricing, and therefore negative risk-adjusted returns. Under efficient contracting, instead, this mechanism applies to low wage gap firms. Our empirical results on risk-adjusted returns lend support to the efficient contracting story.

Overall, the evidence we provide is consistent with the idea that the wage gap is not fully priced in the stock market. Our paper sheds new light on this issue in three ways. First, we provide a theoretical framework to analyze the economic meaning of the wage gap. Second, we show that the overpricing of low wage gap stocks is driven by unsophisticated investors, because pay inequality is priced correctly in stocks where arbitrage forces are likely to be at play. Third, we find that investors trade in a way that is consistent with inequality-averse preferences. To the best of our knowledge, this is the first paper that finds evidence that inequality aversion is a type of behavior among investors.

Our findings speak to a more general literature on market learning. Previous work shows that the market is unable to correctly evaluate intangibles, such as R&D (Lev and Sougiannis, 1996; Chan, Lakonishok, and Sougiannis, 2001), advertising (Chan, Lakonishok, and Sougiannis, 2001), patent citations (Deng, Lev, and Narin, 1999), software development costs (Aboody and Lev, 1998). Similarly, Edmans (2011) finds that investors do not fully incorporate public information on employee satisfaction into the stock price. We show that pay inequality is a further mechanism.

Our findings speak to the literature on market learning. Previous work shows that the market is unable to correctly evaluate intangibles, such as R&D (Lev and Sougiannis, 1996; Chan, Lakonishok, and Sougiannis, 2001), advertising (Chan, Lakonishok, and Sougiannis, 2001), patent citations (Deng, Lev, and Narin, 1999), software development costs (Aboody and Lev, 1998). Similarly, Edmans (2011) shows that investors do not fully incorporate public information on employee satisfaction into the stock price. These stocks, indeed, all trade below their fundamental values, and earn positive risk-adjusted returns in the long-run. In this paper, we unveil a different mechanism. The market evaluates high wage gap stocks correctly but overvalues low wage gap stocks, and the mispricing comes from the trading activity of unsophisticated investors rather than arbitrageurs.

This paper also contributes to the literature that studies the impact of values on investor behavior (Grinblatt and Keloharju, 2001; Bhattacharya and Groznik, 2008; Morse and Shive, 2011; Kaustia and Torstila, 2011). Previous research shows that investors consider nonmonetary variables in their trading strategies, such as moral issues (Hong and Kacperczyk, 2009), or political affiliation (Hong and Kostovetsky, 2012). Our results extend these findings by providing evidence that investors, much like the general public, dislike pay inequality within firms.

The rest of the paper is organized as follows. Section 2 introduces the model. Section 3 illustrates the data. Section 4 explains the methodology. Section 5 discusses the empirical results. Section 6 concludes.

2. Model

We consider a three-period economy. At time zero, the representative firm appoints a manager who exerts effort e to negotiate workers' wages w, hire L workers, and carry out a project of size K. The manager receives a salary W based on his effort. At time one, the firm seeks funding on the stock market to pay workers' wages and start production. At time two, the firm liquidates and pays off investors.

Under efficient contracting, the firm maximizes profits by eliciting the level of effort that is also individually optimal for the manager:

$$\max_{e,L} \pi(e, L) = y(e, L) - w(e)L - W(e), \tag{29}$$

where $y(e, L) \equiv \theta(e)K^{1-\alpha}L^{\alpha}$ is the firm's output, with $\theta(e)$ indicating productivity. The firm's choice of optimal managerial effort (e^*) reflects the following trade-off: high effort costs more, ($W^{'}(e) > 0$), but also makes the firm's employees more valuable ($\theta^{'}(e) > 0$), and yields better outcomes in salary negotiations ($w^{'}(e) < 0$). In Appendix A, we show that a high wage gap between CEO and regular workers is an indication of high effort and high profitability.

The CEO type varies across firms. The stock market then evaluates managers using the wage gap as a signal for effort. Investors have unit mass and can be either sophisticated (arbitrageurs) or unsophisticated (noise traders). As such, they can be thought of as hedge funds and mutual funds, respectively, where the latter face short-sales constraints (Chen, Hong, and Stein, 2002).

We identify two types of unsophisticated investors: noise traders and inequality-averse traders, whose populations are of size λ_U and λ_I, respectively, with $\lambda_U, \lambda_I \in (0, 1)$ and $\lambda_U + \lambda_I < 1$. The former neglect the wage gap, but do take into account other fundamental information about the firm. The latter acknowledge the economic meaning of the wage gap, but dislike pay inequality within firms. Therefore, they exhibit an inequality-aversion bias: their evaluation of low wage gap firms is higher than the evaluation of arbitrageurs, and vice versa for high wage gap firms.

We consider a stock market from Hong and Sraer (2013), as investor i is risk-neutral and solves:

$$\max_{n_{ij}} \left(n_{ij}(E_i(v_j) - p_j) - \frac{1}{2}\frac{n_{ij}^2}{\gamma} \right), \tag{30}$$

where n_{ij} is the number of shares traded by investor of type i in stock j, $E_i(v_j)$ is investor i's subjective evaluation of the stock's cash flow, p_j is the price of stock j, and γ captures transaction costs.[19] The discount rate is set to zero, without loss of generality.

Following Chen, Hong, and Stein (2002), we define returns as the difference between price and fundamental value. Given the presence of short-sales constraints, the equilibrium price depends on whether the true value of profits per unit of capital, $\phi(e)$, is above or below the flat prior ϕ.

Propostion 1. Wage gap and stock returns.

a) *For high wage gap stocks ($\phi(\bar{e}) > \phi$), short-sales constraints are binding. Therefore, given supply q, high wage gap stocks are correctly priced*:

$$p^*(\bar{e}) = \phi(\bar{e})K - \frac{q}{\gamma}, \tag{31}$$

then returns on high wage gap stocks only reflect compensation for transaction costs:

$$r^*(\bar{e}) = \frac{q}{\gamma}. \tag{32}$$

b) *For low wage gap stocks ($\phi(\underline{e}) < \phi$), short-sales constraints do not bind. The equilibrium price is then*:

$$p^*(\underline{e}) = \phi(\underline{e})K + \lambda_U(\phi - \phi(\underline{e}))K + \lambda_I(\phi - \phi(\underline{e}))bK - \frac{q}{\gamma}, \tag{33}$$

where $b > 0$ is a parameter that captures the inequality-aversion bias. This implies lower re-

[19] A type of transaction cost that is characterized by such a convex function is the bid-ask spread, as larger trades are typically associated with more unfavorable price movements.

turns:

$$r^*(\underline{e}) = \frac{q}{\gamma} \underbrace{-\lambda_U(\phi - \phi(\underline{e}))K - \lambda_I(\phi - \phi(\underline{e}))bK}_{<0}, \tag{34}$$

c) *Stocks with high wage gaps outperform stocks with low wage gaps*:

$$r^*(\bar{e}) - r^*(\underline{e}) = \lambda_U(\phi - \phi(\underline{e}))K + \lambda_I(\phi - \phi(\underline{e}))bK. \tag{35}$$

Proof: See Appendix A.

Proposition 2. Inequality-averse investors.

a) *An increase in noise trader demand has an ambiguous effect on the return differential between high and low wage gap stocks*:

$$\frac{d\left(r^*(\bar{e}) - r^*(\underline{e})\right)}{d\lambda_U} = (\phi - \phi(\underline{e}))K - \left|\frac{d\lambda_I}{d\lambda_U}\right| (\phi - \phi(\underline{e}))bK. \tag{36}$$

The derivative is positive if $b < \hat{b} \equiv \left(\left|\frac{d\lambda_I}{d\lambda_U}\right|\right)^{-1}$*, and negative otherwise.*

b) *The difference in valuations between stocks with high and low wage gaps is ambiguous in the presence of inequality-averse investors*:

$$p^*(\bar{e}) - p^*(\underline{e}) = \underbrace{(\phi(\bar{e}) - \phi)K + (1 - \lambda_U)(\phi - \phi(\underline{e}))K}_{>0} \underbrace{-\lambda_I(\phi - \phi(\underline{e}))bK}_{<0}. \tag{37}$$

The difference in valuations between high and low wage gap stocks in Equation (37) reflects two opposite effects. The first effect incorporates the trading of arbitrageurs and noise traders, which leads to a positive difference between the prices of high and low wage gap stocks. The second effect is the presence of inequality-averse traders, whose demand increases the price for low wage gap stocks. This additional effect can mitigate or eliminate the price differential altogether.

2.1 Alternative settings

We acknowledge that the wage gap can also reflect two alternative stories. First, there might be no relation between the wage gap and managerial effort. In this case, the wage gap bears no information on the quality of management. This implies that noise traders, who neglect pay inequality, evaluate firms correctly, and therefore the overpricing of low wage gap stocks only comes from inequality-averse demand. As a result, stock returns follow a pattern that is essentially identical to the one from the efficient contracting story.

Second, a high wage gap might be the result of rent extraction. For example, an entrenched CEO might set a higher wage gap. In this case, pay inequality is negatively related to the quality of management. If so, high wage gap firms exhibit a lower fundamental value due to agency issues, while low wage gap firms are efficient. Arbitrageurs recognize this information. Inequality-averse traders do too, but trade against high pay inequality. Noise traders continue to neglect the wage gap, thereby evaluating all firms equally. As a result, it is firms with a high wage gap, rather than a low one, that exhibit overpricing.

These three scenarios can be distinguished by testing our assumption that the relation between wage gap and firm performance is positive.

2.2 Testable implications

Using Propositions 1a, 1b, and 1c above, we derive the following hypotheses on stock returns:

Hypothesis 1a. Stocks with high wage gaps yield zero risk-adjusted returns.

Hypothesis 1b. Stocks with low wage gaps yield negative risk-adjusted returns. The magnitude of the mispricing decreases with the wage gap.

Hypothesis 1c. A portfolio of stocks with a long-short position in high and low wage gap stocks, respectively yields positive risk-adjusted returns. The profitability of the trading strategy increases with the difference in wage gaps.

These hypotheses reflect the fact that short-sales constraints should be binding for high wage gap stocks, but not for low wage gap stocks. In particular, the lower the wage gap, the greater the pricing error.

From Proposition 2a, we derive:

Hypothesis 2. If investors exhibit high inequality-aversion, an increase in noise trader demand is followed by a decrease in the return differential between high and low wage gap stocks.

The intuition behind Hypothesis 2 is as follows. In standard asset pricing models, an increase in noise trader demand is followed by greater overpricing, and therefore higher arbitrage returns. This result also finds strong empirical support (see, e.g., Stambaugh et al., 2012). In the presence of inequality-averse investors, however, we show that there is an additional mechanism. An increase in noise trader demand also dilutes the impact of inequality-averse demand, which implies a decrease in overpricing and arbitrage returns. The dilution effect dominates if the inequality-aversion bias is large enough.

In the empirical analysis, we take these predictions to the data.

3. Data

The main hurdle in studies of workers' wages is the lack of publicly available data. In this paper, we overcome this issue by using the "Establishment History Panel" database, maintained and made available by the German Federal Employment Agency (Bundesargentur für Arbeit, BfA). This is a unique data set that reports the annual gross wage for all rank-and-file workers in Germany, along with information on the local establishment they work for. The data set was then matched with stock market data for 100 large firms from the two stock market indices DAX and MDAX in years 2004 to 2006 by the agency itself.[20]

The database also contains individual characteristics such as nationality, age, gender, qualification, and type of work. While the complete database contains all these variables for each employee, the data set made available to researchers aggregates these variables across all workers at the establishment level. Our data set then contains the median and quartiles of the wage distribution in any given establishment, but not the wage of each individual worker.

In addition to workers' wages, we complement our data set using company-level accounting and stock market data from Worldscope and Datastream, and CEO compensation data from the companies' annual reports. The sample period is from January 2001 to December 2011, and therefore includes the recent German reform on executive compensation disclosure. Before the reform, listed companies only had to report the aggregate pay of their key corporate executives. Since corporations were not keen on providing information on individual managers' pay, the Federal Government of Germany passed a regulation in 2005 that made such disclosure mandatory.

The number of firms at the start of the sample period is 66, peaking to 100 in 2005, and slowly decreasing to 84 by 2011. On average, the number of firms in our sample is 95 per year. The number of establishments follows a similar pattern, starting at 16,471, peaking to 25,767 in 2006, only to slowly decrease to 15,607 by the end of the sample period.

The industry breakdown shows that the most represented sector in the sample is post and telecommunication (24%), followed by financial intermediation (17%), and retail trade (9%). The distribution of establishments by states, on the other hand, shows that the most represented regions are "Nordrhein-Westfalen" (18%), followed by Bayern (16%), and Baden-Württemberg (12%). A significant proportion of establishments is located in the same state as the firms' headquarters (18%).

The data set includes 146 CEOs and 734 executive board members overall. The representative CEO in our sample is 54 years old, with a tenure of approximately six years, and an average annual compensation of €2.62 million, of which €2.06 million in cash (see Table 18). CEO turnover is relatively low (9%), and a substantial portion of CEOs is hired inside the firm (43%). The average management board includes five members. Each member receives an average total compensation of €1.45 million per year, of which €1.17 million in cash. The pay of non-CEO managers is highly correlated with that of the CEO (80%).

The average firm employs 52,124 workers overall, and around 60 full-time employees per establishment, with a median age of 41 years. The overwhelming majority of such workers is German (97%). Most of them are highly qualified (73%), and have a white-collar job (61%). Women are well-represented in the sample (43%). In 95% of the cases there is a labor union within the firm.

The median wage at the establishment level for full-time rank-and-file workers is €35,167 per year,

[20] At the time of matching establishments to firms, establishment data was not available for 2007 and subsequent years.

Table 18: Summary statistics

Summary statistics for board (Panel A) and firm-level (Panel B) variables in our sample. CEO and managers' pay are defined as the total annual compensation, including cash and stocks. Wage gaps are calculated with respect to workers wages, which we measure as the average annual wage paid in establishments that are located in the same state as the firm's headquarters, weighted by the number of employees in each establishment. The statistics for compensation and total assets are expressed in euros. The sample period is from January 2002 to December 2011.

Panel A

Variable	Observations	Mean	Std. Deviation	Min	Max
CEO pay	6,161	2.62	2.42	0.02	16.60
CEO cash	6,161	2.06	1.68	0.02	12.90
Manager pay	6,149	1.45	1.08	0.02	8.17
Manager cash	6,149	1.17	0.79	0.00	6.87
CEO wage gap (logs)	5,981	3.89	1.01	-0.61	8.42
Manager wage gap (logs)	5,969	3.41	0.88	-0.40	7.80
CEO cash wage gap (logs)	5,981	3.71	0.95	-0.61	7.73
Manager cash wage gap (logs)	5,933	3.26	0.77	1.56	7.33

Panel B

Variable	Observations	Mean	Std. Deviation	Min	Max
ROA	9,353	4.51	8.44	-46.39	78.81
ROE	9,403	7.06	40.97	-724.97	444.25
Tobin's q	9,427	2.32	2.55	0.21	28.57
Total assets	9,415	57.87	207.67	29.78	2,193.95
Employees	9,439	52,124	93,462	2	536,350
Returns	13,144	-0.0019	0.1231	-1.8971	0.8450

while the first and third quartiles of the distribution are, respectively, €31,678 and €37,301. Therefore, the average ratio between CEO and workers' pay is 73.6, which is high, but still six times lower than the average wage gap in the U.S. (see, e.g., (Faleye, Reis, and Venkateswaran, 2013)). For non-CEO managers, the average wage gap is 40.8.[21]

4. Methodology

Following previous literature, we define the wage gap as the log-difference between managers' pay and rank-and-file workers' wages. In our baseline specifications, we primarily define managerial compensation as the overall pay of the CEO, including both the variable and the fixed component. For robustness, we alternatively define it as the overall pay of the other members of the management board. On the other hand, we define workers' compensation in a given company as the average establishment-level median wage, weighted by the number of employees in each establishment.[22]

To construct our test portfolios, we rank all stocks in pay inequality quantiles, as measured in the previous year. Then, we rebalance them at the beginning of each year, and define high and low wage gap portfolios, respectively, as the stocks that lie at the top and the bottom 30%, 20%, and 10% of the distribution. We define our dependent variable as excess returns on the top portfolios, excess

[21] For our summary statistics, we consider the wage gap without logs. In the empirical analysis, however, we express the wage gap in log-differences.

[22] All the results that follow are robust to using a simple arithmetic average across establishments.

returns on the bottom portfolios, and returns on the top-minus-bottom portfolios. We consider equally-weighted returns all throughout.[23].

In our baseline specifications, we follow Cronqvist, Heyman, Nilsson, Svaleryd, and Vlachos (2009) and calculate the wage gap only using the establishments that are located in the same state as the firm's headquarters. In addition to their argument on proximity and ease of interaction between the management and employees, we acknowledge the fact that political and economic heterogeneity across German states may create noise in our wage gap estimates. With this restriction, we make sure that managers and workers face the same local government and costs of living.

To test the model's assumption that high wage gap firms are better managed, we estimate the following panel regressions:

$$y_{it} = \alpha_i + \alpha_t + \beta_1 w_{it-1} + \gamma S_{it} + \epsilon_{it}, \tag{38}$$

where y_{it} is the return on assets or Tobin's q for company i in year t, α_i and α_t are firm and year fixed effects, w_{it-1} is the wage gap of stock i in the previous year, and S_{it} is the logarithm of the company's number of employees as a proxy for size. Standard errors are clustered by firm and year.

To test the hypotheses from the model, we primarily run time-series regressions using European financial data from Kenneth French's website.[24] Our main test equation is Carhart's four-factor model, which allows us to control for a number of well-known risk-factors:

$$R_{it} = \alpha_i + \beta_i MKT_t + s_i SMB_t + h_i HML_t + u_i UMD_t + \epsilon_{it}, \tag{39}$$

where the dependent variable is the excess returns on portfolios of stocks with high ($i = H$) or low ($i = L$) wage gaps, or the returns on the long-short portfolio ($i = H - L$). The independent variables are excess returns on the market portfolio (MKT), and the size (SMB), book-to-market (HML), and momentum factor (UMD), respectively. The intercept captures abnormal risk-adjusted returns. In the analysis below, we also report the estimates from simple CAPM regressions (including the market factor only). Standard errors are calculated following Newey and West (1987) to correct for heteroskedasticity and serial correlation.

Hypothesis 1a implies no pricing errors for high wage gap stocks, and therefore $\alpha_H = 0$. Hypothesis 1b predicts that low wage gap stocks trade at a premium, i.e., $\alpha_L < 0$. In turn, Hypothesis 1c implies $\alpha_{H-L} > 0$. Following the theoretical guidance, we estimate the equation for all three thresholds (30%, 20%, 10%), and expect the mispricing to become stronger as we move towards the tails of the distribution.

Of particular importance is also the coefficient of the size factor, because pay inequality seems to be related to firm size. The idea is that executive ability is worth more to firms that own a larger amount of resources. This mechanism leads to "assortative matching", where better managers are

[23]The results that follow are weaker for value-weighted returns, as large firms are less affected by mispricing (Baker and Wurgler, 2006).

[24]All the results that follow are robust to using German risk factors from the Humboldt University of Berlin.

hired by larger firms (Tervio, 2008; Gabaix and Landier, 2008). Therefore, we need to make sure that any difference in returns between high and low wage gap stocks does not simply reflect a size premium, but rather constitutes a separate effect. Following this line of reasoning, we expect the following loadings on the size factor: $s_H < 0$, $s_L > 0$, $s_{H-L} < 0$.

To test Hypothesis 2, we consider investor sentiment as a proxy for the size of noise trader demand (Yu and Yuan, 2011; Stambaugh, Yu, and Yuan, 2012). High sentiment then identifies an increase in λ_U in Equation (35), which allows us to empirically test Equation (36). We capture the level of sentiment of German investors by considering the consumer confidence index for Germany. For robustness, we also consider consumer confidence for the euro area and the OECD. Given the high persistence of all three indices, we consider them in changes. To identify periods with high or low sentiment, for each of the indices we construct a dummy variable that takes on value one if the index has increased in value over the previous month, and zero otherwise. Then, we re-estimate Equation (39) conditioning on high and low sentiment periods:

$$R_{it} = \alpha_i + \delta_i D_{t-1} + \beta_i MKT_t + s_i SMB_t + h_i HML_t + u_i UMD_t + \epsilon_{it}, \tag{40}$$

where D_{t-1} represents the sentiment dummy. If investors exhibit a large inequality-aversion bias, high sentiment should be followed by less overpricing for low wage gap stocks, which implies a decrease in arbitrage returns. Therefore, we should observe $\delta_L > 0$ and $\delta_{H-L} < 0$.

In robustness checks, we also estimate the following panel regressions:

$$R_{it} = \beta_0 + \beta_1 d_{it-1} + \delta^{'} Z_{it} + \epsilon_{it}, \tag{41}$$

where R_{it} is the return on stock i in month t; d_{it-1} is a dummy variable that takes on value one if the wage gap of stock i in the previous year is among the top 30% of the distribution; and Z_{it} is a vector of firm characteristics from Brennan et al. (1998), which includes: size (defined as the log of market capitalization at the end of month $t-2$), the log of the book-to-market ratio (calculated each July and held constant through the following June), the ratio of dividends in the previous fiscal year to market value at calendar year-end (calculated each July and held constant through the following June), the log of cumulative returns over months $t-3$ through $t-2$, months $t-6$ through $t-4$, and months $t-12$ through $t-7$, the log of the dollar volume of trading in the stock in month $t-2$, and the log of the stock price at the end of month $t-2$. To compare extreme wage gaps, we leave out the stocks that lie in the middle 40% of the wage gap distribution. In this specification, the model's predictions imply $\beta_1 > 0$.

5. Empirical results

Our analysis proceeds as follows. First, we provide empirical support to the idea that a high wage gap represents compensation for better management skills. In the second subsection, we estimate our baseline regressions from Equation (39). In the third subsection, we look into whether the results are driven by sophisticated or unsophisticated traders. In the fourth subsection, we distinguish noise traders from inequality-averse traders, estimating Equation (40). The fifth and final subsection provides results for a variety of alternative specifications, including panel regressions from 41.

5.1 ROA and valuations

As a preliminary test, we check the assumption of the model that high wage gap firms are run better than low wage gap firms. This is an important test also to rule out the alternative hypotheses that the wage gap might be uninformative or represent agency issues. To this purpose, we estimate Equation (38) for returns on assets. The results are in Table 19, Panel A. We find that the coefficient of the wage gap is positive and significant. The results then provide support to the idea that high wage gap firms are indeed better managed.

Proposition 2b predicts that the difference in valuations between high and low wage gap stocks is not clear if there are inequality averse traders. To test for this, we estimate Equation (38) using Tobin's q as a dependent variable. The results are in Table 19, Panel B. We find that the coefficient of the wage gap is not significant. This pattern is consistent with the presence of inequality-averse demand.

5.2 Wage gap between CEO and workers

In our first batch of time series regressions, we consider the wage gap between CEO and workers. The results are in Table 20. In Panel A, we construct the portfolios using the 30% threshold. In columns (1) and (2), we consider the high-inequality portfolio. We find that the regression constant is small, negative, and marginally significant in the CAPM (-0.6%), but close to zero both in magnitude and significance in the full four-factor model (-0.1%). The coefficient of size, instead, is negative and highly significant (-0.767), which shows that high-inequality stocks tend to co-move positively with the returns on large firms. This is consistent with the idea that pay inequality is positively related to firm size.

In columns (3) and (4), we consider the low-inequality portfolio. We find that the alpha is indeed negative and highly significant both in the CAPM (-1.2%) and in the four-factor model (-0.9%). Among the other regressors, the size factor has a positive but insignificant coefficient (0.192), which is consistent with the assortative matching hypothesis. In columns (5) and (6), we consider the long-short portfolio on top and bottom pay inequality stocks, respectively. We find that the alpha is positive and significant both in the CAPM (0.7%) and in the four-factor model (0.8%).

In Panel B, we construct the portfolios using the 20% threshold. The results are similar, and become statistically stronger. For high wage gap stocks, the regression constant is no longer significant in the CAPM (-0.8%), and drops to zero both in magnitude and significance in the full four-factor model (0.1%). The alpha on the low inequality portfolio is negative and highly significant both in the CAPM (-1.6%) and in the four-factor model (-1.2%). As a result, the alpha on the long-short portfolio is positive and significant both in the CAPM (0.8%) and in the four-factor model (1.3%).

Table 19: Wage gap, firm performance, and valuations

Panel regressions of firms' returns on assets (Panel A), and valuations (Panel B), defined as Tobin's q, on the wage gap between CEO and workers, lagged one year and expressed in logs, and the logarithm of the firm's employees. CEO compensation is measured as total annual pay, including cash and stocks, in columns (1) and (2), and as cash only in columns (3) and (4). Workers' pay is measured as the annual wage paid in establishments that are located in the same state as the firm's headquarters, calculated as a weighted average across establishments, where the weights are represented by the number of employees in each establishment. All specifications include firm and year fixed effects, and standard errors are clustered by firm and year. Data on workers' wages are from the "Establishment History Panel" database. Company-level accounting and stock market data from Worldscope and Datastream, and managerial compensation data from the companies' annual reports. Observations are monthly. The sample period is from 2002 to 2011. t statistics are below the coefficients, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Panel A

Dep. variable: ROA	**Total pay**		**Cash pay**	
	(1)	(2)	(3)	(4)
Wage gap	0.025**	0.025**	0.027**	0.027**
	2.03	2.01	2.19	2.20
Employees		-0.016		-0.015
		-0.53		-0.50
Constant	0.037	0.201	0.029	0.185
	0.43	0.61	0.35	0.52
Firm FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Clustering	Firm-Year	Firm-Year	Firm-Year	Firm-Year
Adj. R-squared	0.726	0.727	0.727	0.727
Observations	505	505	505	505

Panel B

Dep. variable: Tobin's q	**Total pay**		**Cash pay**	
	(1)	(2)	(3)	(4)
Wage gap	-0.268	-0.264	-0.267	-0.267
	-0.52	-0.52	-0.47	-0.49
Employees		-1.049		-1.056
		-1.16		-1.16
Constant	13.843***	24.771*	13.794***	24.832*
	3.21	1.93	2.94	1.84
Firm FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Clustering	Firm-Year	Firm-Year	Firm-Year	Firm-Year
Adj. R-squared	0.623	0.628	0.623	0.628
Observations	507	507	507	507

Table 20: CEO-workers wage gap

CAPM and Carhart's (1997) four-factor model regressions of equally-weighted returns on a portfolio of stocks with high inequality between CEO and workers' pay (columns 1 and 2), a portfolio of stocks for which such inequality is low (columns 3 and 4), and a portfolio with a long position in high-inequality stocks and a short position in low-inequality stocks (columns 5 and 6). Pay inequality is defined as the log-difference in earnings between CEO and workers. CEO compensation is calculated as the overall annual pay, including cash and stocks, while workers' pay is measured as the average annual wage paid in establishments that are located in the same state as the firm's headquarters, weighted by the number of employees in each establishment. To construct our test portfolios, we rank all stocks in pay inequality quantiles as measured in the previous year, and rebalance them at the beginning of each year. In any given year, we define high and low wage gap stocks, respectively, as the stocks that lie at the top and the bottom 30% (Panel A), 20% (Panel B), and 10% (Panel C) of the distribution. All the regressors are from Kenneth French's website and refer to Europe. Observations are monthly, and the sample period is from January 2002 to December 2011. t statistics are below the coefficients, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Panel A

30% Threshold	(1) H	(2) H	(3) L	(4) L	(5) H−L	(6) H−L
Alpha	-0.006*	-0.001	-0.012***	-0.009**	0.007*	0.008**
	-1.85	-0.18	-2.67	-2.26	1.87	2.34
MKT	1.048***	0.998***	0.907***	0.814***	0.141*	0.184***
	5.54	9.59	5.58	6.60	1.89	3.62
SMB		-0.767**		0.192		-0.960**
		-2.08		0.65		-2.23
HML		-0.393*		-0.249		-0.144
		-1.91		-0.72		-0.60
UMD		-0.260*		-0.379**		0.119**
		-1.96		-2.46		2.19
Adj. R-squared	0.627	0.692	0.539	0.576	0.024	0.176
Observations	120	120	120	120	120	120

Panel B

20% Threshold	(1) H	(2) H	(3) L	(4) L	(5) H−L	(6) H−L
Alpha	-0.008	0.001	-0.016***	-0.012***	0.008*	0.013***
	-1.59	0.26	-3.22	-2.96	1.76	2.81
MKT	1.128***	1.060***	0.905***	0.831***	0.222	0.230**
	3.58	6.32	5.06	6.03	1.05	2.41
SMB		-1.444**		0.368		-1.812***
		-2.25		1.29		-2.82
HML		-0.831**		-0.406		-0.425
		-2.35		-1.10		-1.60
UMD		-0.452**		-0.407***		-0.045
		-2.16		-2.77		-0.47
Adj. R-squared	0.446	0.594	0.484	0.529	0.024	0.268
Observations	120	120	120	120	120	120

Panel C

10% Threshold	(1) H	(2) H	(3) L	(4) L	(5) H−L	(6) H−L
Alpha	0.000	0.002	-0.016***	-0.016***	0.017***	0.018***
	0.14	0.62	-2.83	-3.60	3.44	6.48
MKT	0.819***	0.703***	0.843***	0.712***	-0.024	-0.009
	9.79	7.74	4.07	4.64	-0.12	-0.05
SMB		-0.582***		0.775**		-1.357***
		-3.60		2.36		-3.80
HML		0.363		0.270		0.093
		1.47		0.78		0.35
UMD		-0.185		-0.238***		0.053
		-1.49		-2.60		0.57
Adj. R-squared	0.530	0.571	0.347	0.376	-0.009	0.112
Observations	108	108	108	108	108	108

The result become even stronger in Panel C, where we construct the portfolios using the 10% threshold. The regression constant for the high inequality portfolio is not significant and close to zero both in the CAPM (0.0%) and the four-factor model (0.2%). The alpha on the low inequality portfolio is negative and highly significant both in the CAPM (-1.6%) and in the four-factor model (-1.6%). Therefore, the arbitrage portfolio yields positive and significant risk-adjusted returns both in the CAPM (1.7%) and in the four-factor model (1.8%).

Overall, the evidence is in line with the model predictions. Consistent with Hypotheses 1a, 1b, and 1c, low wage gap stocks trade at a premium, high wage gap stocks are priced correctly, and the arbitrage portfolio yields a positive and significant alpha, which is entirely driven by the short leg of the investment strategy. Also, the negative coefficient of size provides support to the assortative matching hypothesis.

5.3 Informed vs. uninformed traders

To provide further evidence that it is indeed unsophisticated investors that drive the results, we identify categories of firms that are difficult to evaluate and/or arbitrage. Examples of such companies are those of small size, defined as low market capitalization, and with extreme book-to-market ratios (Baker and Wurgler, 2006; Baker and Wurgler, 2007). With lower arbitrage activity, we expect the mispricing of the wage gap to be particularly pronounced among these stocks. To test this conjecture, we double-sort our stocks in portfolios formed on these measures.

We start from firm size. In the first sort, we distinguish between stocks with above- and below-median market capitalization, respectively. In the second sort, we identify stocks with top and bottom 30% wage gaps. To avoid using market prices for both sorts, we consider CEO cash compensation for these tests. Due to little firm coverage at the beginning of the sample period, we construct these portfolios from 2005.

The results are in Table 21. For small stocks (Panel A), we find that an empirical pattern that is strikingly similar to our estimates from the full sample. Risk-adjusted returns are positive but not significant for high wage gap stocks (1.1%), negative and significant for low wage gap stocks (0.7%), and positive and highly significant for the arbitrage portfolio (1.8%). For large stocks (Panel B), instead, none of these coefficients are significant.

Next, we consider the book-to-market ratio. We form our first sort using stocks with middle 40% and extreme 30% book-to-market, respectively. As above, we identify stocks with top and bottom 30% wage gaps in the second sort, and again begin our analysis in 2005.

The estimates are in Table 22. For extreme book-to-market stocks (Panel A), we find that risk-adjusted returns are near-zero for high wage gap stocks (0.2%), negative and significant for low wage gap stocks (0.9%), and positive and highly significant for the arbitrage portfolio (1.1%). For middle book-to-market stocks, instead, these coefficients are all outside the rejection region.

The results are then in line with our conjecture.

Table 21: CEO-worker wage gap: portfolios double-sorted on wage gap and market capitalization

CAPM and Carhart's (1997) four-factor model regressions of equally-weighted returns on a portfolio of stocks with high inequality between CEO and workers' pay (columns 1 and 2), a portfolio of stocks for which such inequality is low (columns 3 and 4), and a portfolio with a long position in high-inequality stocks and a short position in low-inequality stocks (columns 5 and 6). In Panel A we consider stocks whose market capitalization is below the median, while in Panel B we consider stocks whose market capitalization is above the median. Pay inequality is defined as the log-difference in earnings between CEO and workers. CEO compensation is calculated as the cash salary, while workers' pay is measured as the average annual wage paid in establishments that are located in the same state as the firm's headquarters, weighted by the number of employees in each establishment. To construct our test portfolios, we rank all stocks in pay inequality quantiles as measured in the previous year, and rebalance them at the beginning of each year. In any given year, we define high and low wage gap stocks, respectively, as the stocks that lie at the top and the bottom 30% of the distribution. The regressors are from Kenneth French's website and refer to Europe. Observations are monthly, and the sample period is from January 2005 to December 2011. t statistics are below the coefficients, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Panel A

Small stocks	(1) H	(2) H	(3) L	(4) L	(5) H−L	(6) H−L
Alpha	0.010	0.011	-0.008**	-0.007**	0.018**	0.018***
	1.19	1.60	-2.24	-2.19	2.21	2.69
MKT	1.016***	0.991***	0.631***	0.499***	0.385***	0.493***
	7.42	14.18	6.34	6.25	5.50	4.11
SMB		0.399		0.182		0.217
		1.24		0.62		0.81
HML		-0.321		0.493*		-0.814
		-0.61		1.70		-1.50
UMD		-0.284**		-0.116		-0.168**
		-2.16		-0.99		-1.98
Adj. R-squared	0.551	0.561	0.446	0.463	0.150	0.175
Observations	84	84	84	84	84	84

Panel B

Large stocks	(1) H	(2) H	(3) L	(4) L	(5) H−L	(6) H−L
Alpha	-0.002	-0.002	-0.006	-0.005	0.004	0.002
	-0.93	-0.94	-0.78	-0.66	0.62	0.45
MKT	0.773***	0.748***	0.934***	0.822***	-0.161	-0.074
	13.48	12.31	5.72	6.54	-1.17	-0.70
SMB		-0.298***		0.383**		-0.682***
		-2.59		2.02		-4.37
HML		0.186		0.351		-0.165
		0.94		0.95		-0.58
UMD		0.023		-0.129		0.152
		0.20		-0.96		0.92
Adj. R-squared	0.683	0.684	0.59	0.598	0.034	0.100
Observations	84	84	84	84	84	84

Table 22: CEO-worker wage gap: portfolios double-sorted on wage gap and book-to-market ratio

CAPM and Carhart's (1997) four-factor model regressions of equally-weighted returns on a portfolio of stocks with high inequality between CEO and workers' pay (columns 1 and 2), a portfolio of stocks for which such inequality is low (columns 3 and 4), and a portfolio with a long position in high-inequality stocks and a short position in low-inequality stocks (columns 5 and 6). In Panel A we consider stocks whose book-to-market ratio lies in the extreme (i.e., top and bottom) 30% of the distribution, while in Panel B we consider stocks whose book-to-market ratio lies in the middle 40%. Pay inequality is defined as the log-difference in earnings between CEO and workers. CEO compensation is calculated as the cash salary, while workers' pay is measured as the average annual wage paid in establishments that are located in the same state as the firm's headquarters, weighted by the number of employees in each establishment. To construct our test portfolios, we rank all stocks in pay inequality quantiles as measured in the previous year, and rebalance them at the beginning of each year. In any given year, we define high and low wage gap stocks, respectively, as the stocks that lie at the extreme 30% of the distribution. The regressors are from Kenneth French's website and refer to Europe. Observations are monthly, and the sample period is from January 2005 to December 2011. t statistics are below the coefficients, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Panel A

Extreme B/M	(1) H	(2) H	(3) L	(4) L	(5) H−L	(6) H−L
Alpha	0.001	0.002	-0.011***	-0.009**	0.012***	0.011***
	0.70	0.80	-2.77	-2.33	3.79	3.58
MKT	0.789***	0.791***	0.675***	0.553***	0.114	0.238***
	12.59	13.29	5.33	6.62	1.23	3.30
SMB		-0.341**		-0.074		-0.268
		-2.34		-0.28		-1.18
HML		-0.068		0.422		-0.490*
		-0.33		1.26		-1.75
UMD		-0.058		-0.144		0.086
		-0.47		-1.26		1.14
Adj. R-squared	0.646	0.646	0.455	0.466	0.015	0.041
Observations	84	84	84	84	84	84

Panel B

Middle B/M	(1) H	(2) H	(3) L	(4) L	(5) H−L	(6) H−L
Alpha	-0.003	-0.003	-0.005	-0.004	0.002	0.001
	-1.19	-0.92	-0.89	-1.00	0.40	0.41
MKT	0.856***	0.762***	0.851***	0.738***	0.005	0.024
	13.16	11.43	6.12	5.68	0.05	0.17
SMB		-0.069		1.063***		-1.131***
		-0.57		6.61		-7.62
HML		0.419		0.345		0.075
		1.63		1.10		0.23
UMD		-0.048		-0.101		0.053
		-0.45		-0.95		0.43
Adj. R-squared	0.671	0.675	0.521	0.591	-0.012	0.170
Observations	84	84	84	84	84	84

5.4 Inequality aversion

To tease out which category of unsophisticated (i.e., noise and inequality-averse) traders drives the results, we proceed in three ways. First, we propose a formal test of Hypothesis 2, by estimating Equation (40). The intuition is as follows. Noise trader demand is high if the market has positive sentiment (Yu and Yuan, 2011; Stambaugh, Yu, and Yuan, 2012). Without inequality-averse traders, Equation (36) implies that an increased presence of noise traders results in an increased return differential between high and low wage gap stocks. With inequality-averse traders, instead, the return differential decreases when the inequality bias is large.

The results are in Table 23. In Panel A, we consider the consumer confidence index for Germany. In columns (1) and (2), we find that high wage gap stocks exhibit no significant variation in risk-adjusted returns in times of high sentiment, either in the CAPM (-0.2%) or the four-factor model (1.1%). In columns (3) and (4), low wage gap stocks exhibit an interesting pattern. In times of low sentiment, they yield negative and highly significant risk-adjusted returns (-2.1%; -2.0%). In times of high sentiment, instead, their risk-adjusted returns increase significantly (1.6%; 2.3%), which implies less overpricing. In columns (5) and (6), the long-short portfolio exhibits a similar pattern. Risk-adjusted returns are positive and significant in times of low sentiment (1.7%; 1.4%), only to decrease sharply in times of high sentiment (-1.8%; -1.1%).

In Panels B and C we consider consumer confidence for the euro area and the OECD, respectively, and obtain similar results. The alpha of high wage gap stocks continues to be close to zero all throughout. Low wage gap stocks exhibit negative and significant risk-adjusted returns in times of low sentiment, but most of the overpricing goes away in times of high sentiment. Symmetrically, the long-short portfolio yields positive and significant risk-adjusted returns in times of high sentiment, but the alphas significantly decrease in times of high sentiment. The negative coefficient of sentiment in specifications (5) and (6) across all three Panels indicates that we have inequality-averse traders.

Second, we argue that if some investors exhibit a bias related to pay inequality, we expect such a bias to be stronger when the wage gap is salient, i.e., when it is disclosed by firms and covered by the media. This is a peculiar implication of inequality aversion, as noise traders neglect the wage gap at all times during the year, while arbitrageurs always trade on it. To test for this, we identify April as the month in which the wage gap is typically disclosed and covered by the press[25]. The increase in attention towards pay inequality should then lead to more trading of inequality-averse investors. To test for this, we split our regression constant into two sets of dummy variables that take on value one for each quarter and calendar month, respectively.

Note that the tests from our arbitrage portfolio are particularly important here. Calendar anomalies that are not specific to pay inequality should equally affect high and low wage gap stocks, and then cancel out in the long-minus-short portfolio. On the contrary, the attention effect from the inequality aversion story, if present, should affect high and low wage gap stocks in different ways, and therefore should not wash away in the arbitrage strategy.

[25]For 91 firms, the fiscal year-end is on December 31, and the annual meeting is generally in April. Only 9 firms have fiscal years that do not coincide with the calendar year.

Table 23: CEO-workers wage gap and sentiment

CAPM regressions (columns 1, 3, and 5) and Carhart's (1997) four-factor model regressions (columns 2, 4, and 6) of equally-weighted returns on a portfolio of stocks with high inequality between CEO and workers' pay (columns 1 and 2), a portfolio of stocks for which such inequality is low (columns 3 and 4), and a portfolio with a long position in high-inequality stocks and a short position in low-inequality stocks (columns 5 and 6). The regressions include a dummy variable that equals one if sentiment has increased over the previous month and zero otherwise, where sentiment is defined as the measure of consumer confidence published by the OECD for Germany (Panel A), the eurozone area (Panel B), and the OECD area (Panel C). Pay inequality is defined as the log-difference in earnings between CEO and workers. CEO compensation is calculated as the overall annual pay, including cash and stocks, while workers' pay is measured as the average annual wage paid in establishments that are located in the same state as the firm's headquarters, weighted by the number of employees in each establishment. To construct our test portfolios, we rank all stocks in pay inequality quantiles as measured in the previous year, and rebalance them at the beginning of each year. In any given year, we define high and low wage gap stocks, respectively, as the stocks that lie at the top and bottom 30% of the distribution. All the regressors are from Kenneth French's website and refer to Europe. Observations are monthly, and the sample period is from January 2002 to December 2011. t statistics are below the coefficients, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Panel A	(1) H	(2) H	(3) L	(4) L	(5) H−L	(6) H−L
Sentiment	-0.002	0.011	0.016*	0.023***	-0.018**	-0.011**
	-0.12	1.28	1.76	2.81	-2.48	-1.97
Alpha	-0.005	-0.007	-0.021***	-0.020***	0.017***	0.014**
	-0.82	-0.90	-4.97	-3.66	2.87	2.47
Controls	Y	Y	Y	Y	Y	Y
Adj. R-squared	0.624	0.694	0.546	0.593	0.049	0.182
Observations	120	120	120	120	120	120

Panel B	(1) H	(2) H	(3) L	(4) L	(5) H−L	(6) H−L
Sentiment	-0.006	0.004	0.019***	0.021***	-0.026***	-0.016**
	-0.62	0.68	2.74	4.14	-3.36	-2.55
Alpha	-0.002	-0.003	-0.024***	-0.020***	0.022***	0.017***
	-0.38	-0.52	-4.85	-4.81	3.92	3.01
Controls	Y	Y	Y	Y	Y	Y
Adj. R-squared	0.625	0.690	0.549	0.588	0.078	0.194
Observations	120	120	120	120	120	120

Panel C	(1) H	(2) H	(3) L	(4) L	(5) H−L	(6) H−L
Sentiment	-0.004	0.010	0.018**	0.020**	-0.023***	-0.010*
	-0.39	1.28	2.22	2.29	-2.93	-1.95
Alpha	-0.003	-0.006	-0.022***	-0.019***	0.019***	0.013***
	-0.70	-1.03	-4.59	-3.33	3.56	2.85
Controls	Y	Y	Y	Y	Y	Y
Adj. R-squared	0.624	0.693	0.548	0.587	0.067	0.179
Observations	120	120	120	120	120	120

Table 24: CEO-workers wage gap: quarter breakdown

CAPM and Carhart's (1997) four-factor model regressions of equally-weighted returns on a portfolio of stocks with high inequality between CEO and workers' pay (columns 1 and 2), a portfolio of stocks for which such inequality is low (columns 3 and 4), and a portfolio with a long position in high-inequality stocks and a short position in low-inequality stocks (columns 5 and 6). Pay inequality is defined as the log-difference in earnings between CEO and workers. CEO compensation is calculated as the overall annual pay, including cash and stocks, while workers' pay is measured as the average annual wage paid in establishments that are located in the same state as the firm's headquarters, weighted by the number of employees in each establishment. To construct our test portfolios, we rank all stocks in pay inequality quantiles as measured in the previous year, and rebalance them at the beginning of each year. In any given year, we define high and low wage gap stocks, respectively, as the stocks that lie at the top and the bottom 30% of the distribution. We replace the regression constant with a set of dummies that take on value one for each quarter. All the regressors are from Kenneth French's website and refer to Europe. Observations are monthly, and the sample period is from January 2002 to December 2011. t statistics are below the coefficients, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(1) H	(2) H	(3) L	(4) L	(5) H−L	(6) H−L
1^{st} Quarter	0.005	0.029	0.003	0.009	0.001	0.020*
	0.61	1.53	0.42	0.80	0.16	1.88
2^{nd} Quarter	-0.009	-0.007	-0.022***	-0.023***	0.012**	0.016***
	-0.92	-0.92	-2.93	-2.76	2.14	3.59
3^{rd} Quarter	-0.013***	-0.009	-0.018*	-0.015*	0.005	0.007
	-2.70	-1.60	-1.72	-1.71	0.57	0.97
4^{th} Quarter	0.002	-0.002	-0.006	0.003	0.008	-0.004
	0.19	-0.19	-0.74	0.25	0.69	-0.43
MKT	1.054***	1.013***	0.920***	0.810***	0.133*	0.202***
	5.27	8.78	5.46	6.96	1.68	3.38
SMB		-0.955**		0.16		-1.115***
		-2.29		0.43		-2.87
HML		-0.419**		-0.222		-0.197
		-2.13		-0.64		-0.87
UMD		-0.284**		-0.443***		0.158*
		-2.15		-2.89		1.68
Ad. R-squared	0.62	0.706	0.539	0.592	0.018	0.197
Observations	120	120	120	120	120	120

The results are in Tables 24 and 25. Consistent with our conjecture, we find that the overpricing of low wage gap stocks takes place especially during the second quarter, where alpha is positive and significant (2.3%). As a consequence, the long-short strategy exhibits positive and highly significant risk-adjusted returns only in that period (1.6%). The results are similar for the calendar month breakdown, as the long-short strategy yields a positive and highly significant alpha only in the month of April (3.1%).

Third, we argue that if inequality aversion is a genuine behavioral bias, then it should not be related to economic fundamentals. In particular, it should not matter whether high CEO pay is justified or not by the manager's contribution to the firm's value. To test this hypothesis, we follow Rouen (2017) and decompose CEO pay into an economically justified and an unjustified component, where the former is a linear combination of firm characteristics. In particular, we regress log CEO pay on log total assets, the market-to-book ratio, the earnings-to-assets ratio, CEO tenure, and industry and year fixed effects, and define the fitted values and the residuals as justified and unjustified pay, respectively. As noise traders neglect the wage gap but observe the other fundamentals of the firm,

Table 25: CEO-workers wage gap: calendar month breakdown

CAPM and Carhart's (1997) four-factor model regressions of equally-weighted returns on a portfolio of stocks with high inequality between CEO and workers' pay (columns 1 and 2), a portfolio of stocks for which such inequality is low (columns 3 and 4), and a portfolio with a long position in high-inequality stocks and a short position in low-inequality stocks (columns 5 and 6). Pay inequality is defined as the log-difference in earnings between CEO and workers. CEO compensation is calculated as the overall annual pay, including cash and stocks, while workers' pay is measured as the average annual wage paid in establishments that are located in the same state as the firm's headquarters, weighted by the number of employees in each establishment. To construct our test portfolios, we rank all stocks in pay inequality quantiles as measured in the previous year, and rebalance them at the beginning of each year. In any given year, we define high and low wage gap stocks, respectively, as the stocks that lie at the top and the bottom 30% of the distribution. We replace the regression constant with a set of dummies that take on value one for each calendar month. All the regressors are from Kenneth French's website and refer to Europe. Observations are monthly, and the sample period is from January 2002 to December 2011. t statistics are below the coefficients, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(1) H	(2) H	(3) L	(4) L	(5) H−L	(6) H−L
January	0.01	0.042*	0.012	0.015	-0.002	0.027*
	0.92	1.70	0.90	0.86	-0.11	1.69
February	0.000	0.016	-0.004	0.004	0.004	0.012
	-0.07	1.29	-0.48	0.41	0.36	1.16
March	-0.008	-0.004	-0.020**	-0.017**	0.012	0.014
	-0.98	-0.41	-2.53	-2.38	1.20	1.25
April	-0.003	-0.002	-0.025	-0.033**	0.022*	0.031***
	-0.25	-0.25	-1.29	-2.32	1.96	2.88
May	-0.017	-0.016	-0.021	-0.021	0.004	0.005
	-1.03	-1.33	-1.08	-1.07	0.4	0.46
June	-0.007	0.004	-0.007	0.001	0.000	0.002
	-0.68	0.41	-0.42	0.09	0.00	0.21
July	-0.023**	-0.022***	-0.039***	-0.036***	0.016*	0.015
	-2.47	-2.72	-3.31	-4.21	1.78	1.28
August	-0.009	-0.009	-0.008	-0.012	-0.002	0.004
	-0.84	-0.87	-0.37	-0.72	-0.08	0.26
September	-0.027	-0.02	-0.015	-0.004	-0.012	-0.016
	-1.18	-1.33	-0.71	-0.26	-0.83	-1.29
October	0.016	0.003	-0.013	-0.003	0.028	0.007
	0.63	0.21	-1.09	-0.21	1.20	0.38
November	0.016	0.01	0.01	0.016	0.006	-0.005
	1.04	0.71	0.59	0.87	0.37	-0.43
December	-0.017	-0.007	-0.021	-0.007	0.004	0.000
	-1.34	-0.91	-1.50	-0.54	0.29	0.01
MKT	1.053***	1.018***	0.946***	0.818***	0.106	0.200***
	5.94	9.66	6.64	7.33	1.51	3.89
SMB		-0.998***		0.103		-1.101***
		-2.62		0.27		-2.92
HML		-0.397**		-0.15		-0.247
		-2.21		-0.48		-1.18
UMD		-0.264**		-0.442***		0.178***
		-2.14		-2.92		3.19
Adj. R-squared	0.617	0.703	0.531	0.581	-0.010	0.167
Observations	120	120	120	120	120	120

their pricing error should be confined to the unjustified wage gap. For inequality-averse traders, instead, the bias should apply to both the justified and the unjustified wage gap.

The results are in Table 26. In Panel A, we consider the justified wage gap. We find risk-adjusted returns are not significant and close to zero for high wage gap stocks (0.1%), while negative and significant for low wage gap stocks (0.7%), and positive and significant for the arbitrage portfolio (0.8%). In Panel B, we consider the unjustified wage gap. Similarly, we find risk-adjusted returns are not significant and close to zero for high wage gap stocks (0.2%), negative but not significant for low wage gap stocks (0.4%), while positive and marginally significant for the arbitrage portfolio (0.7%). The results lend support to the behavioral bias hypothesis.

5.5 Robustness checks

Next, we test a variety of alternative specifications, in which we consider: non-CEO managers, cash-only compensation, portfolio formation in April rather than January, all establishments rather than same-state establishments only, and panel regressions.

In Table 27, we repeat the analysis for non-CEO managers. The results are similar, both in magnitude and significance. We find that high wage gap stocks yield near-zero risk-adjusted returns for the portfolios with 30% (0.1%), 20% (-0.1%), and 10% (0.2%) thresholds. On the contrary, low wage gap stocks yield negative and significant risk-adjusted returns, increasing in absolute value across the thresholds of 30% (-0.8%), 20% (-1.0%), and 10% (-1.9%). As a result, the long-short portfolios yield positive and significant alphas, as hypothesized, increasing from 0.9% for the 30% portfolios to 2.1% for the 10% portfolios.

Among the other coefficients, as expected, the size factor is the one which retains the highest explanatory power. High wage gap stocks exhibit a negative and significant loading across the 30% (-1.200), 20% (-1.318), and 10% (-0.551) thresholds. For low wage gap stocks, the loading on size is again positive but not significant for the 30% (0.351) and 20% (0.336) portfolios, while it is positive and highly significant for the 10% portfolios (0.866). The difference between these two coefficients, however, is negative and significant for all thresholds, which is in line with the assortative matching hypothesis.

One potential concern is that the variable component of CEO pay makes the wage gap a function of the firm's past performance. To allay this concern, we redefine pay inequality using managers' cash compensation only. The results are in Table 28, and look quite similar. The alpha on high wage gap stocks is close to zero for the portfolios with 30% (0.1%), 20% (0.3%), and 10% (-0.1%) thresholds. Low wage gap stocks, instead, yield negative and significant risk-adjusted returns, increasing in absolute value across the thresholds of 30% (-0.8%), 20% (-1.0%), and 10% (-1.7%). As a result, the long-short portfolios yield positive and significant risk-adjusted returns increasing from 0.9% for the 30% portfolios to 1.6% for the 10% portfolios.

Since company reports are typically released around the end of the first quarter, one additional concern is that our portfolios are formed too early. To address this issue, we move the formation period from January to April. The results are in Table 29, and again hardly change. The alpha

Table 26: CEO-worker wage gap: justified vs. unjustified CEO pay

CAPM and Carhart's (1997) four-factor model regressions of equally-weighted returns on a portfolio of stocks with high inequality between CEO and workers' pay (columns 1 and 2), a portfolio of stocks for which such inequality is low (columns 3 and 4), and a portfolio with a long position in high-inequality stocks and a short position in low-inequality stocks (columns 5 and 6). Pay inequality is defined as the log-difference in earnings between CEO and workers. CEO compensation is calculated as the overall annual pay, including cash and stocks, while workers' pay is measured as the average annual wage paid in establishments that are located in the same state as the firm's headquarters, weighted by the number of employees in each establishment. Following Rouen (2017), we regress log CEO pay on log total assets, the market-to-book ratio, the earnings-to-assets ratio, CEO tenure, and industry and year fixed effects. Then, we define the fitted values as justified pay, and the residuals as unjustified pay. In turn, we construct an economically justified wage gap (Panel A) and an unjustified one (Panel B). To construct our test portfolios, we rank all stocks in pay inequality quantiles as measured in the previous year, and rebalance them at the beginning of each year. In any given year, we define high and low wage gap stocks, respectively, as the stocks that lie at the top and the bottom 30% of the distribution. The regressors are from Kenneth French's website and refer to Europe. Observations are monthly, and the sample period is from January 2005 to December 2011. t statistics are below the coefficients, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Panel A

Justified wage gap	(1) H	(2) H	(3) L	(4) L	(5) H−L	(6) H−L
Alpha	-0.004	0.001	-0.009**	-0.007**	0.005	0.008**
	-1.44	0.45	-2.11	-1.97	1.26	2.49
MKT	1.091***	0.978***	0.938***	0.835***	0.153	0.143**
	5.78	9.58	8.52	10.48	1.24	2.42
SMB		-0.848***		0.347		-1.195***
		-2.65		1.38		-3.20
HML		-0.28		-0.044		-0.236*
		-1.22		-0.17		-1.84
UMD		-0.377***		-0.325***		-0.051
		-3.12		-3.56		-0.92
Adj. R-squared	0.631	0.719	0.599	0.630	0.023	0.237
Observations	120	120	120	120	120	120

Panel B

Unjustified wage gap	(1) H	(2) H	(3) L	(4) L	(5) H−L	(6) H−L
Alpha	-0.003	0.002	-0.012**	-0.004	0.008**	0.007*
	-0.75	0.63	-1.99	-1.11	2.19	1.79
MKT	1.140***	1.032***	1.157***	1.066***	-0.016	-0.034
	6.45	10	4.91	6.58	-0.24	-0.43
SMB		-0.576		-0.330		-0.246
		-1.59		-1.14		-1.01
HML		-0.297		-0.669		0.372
		-1.41		-1.59		1.26
UMD		-0.389***		-0.520***		0.131
		-3.69		-2.93		1.35
Adj. R-squared	0.669	0.731	0.602	0.680	-0.008	0.023
Observations	120	120	120	120	120	120

Table 27: Managers-workers wage gap

CAPM and Carhart's (1997) four-factor model regressions of equally-weighted returns on a portfolio of stocks with high inequality between managers' and workers' pay (columns 1 and 2), a portfolio of stocks for which such inequality is low (columns 3 and 4), and a portfolio with a long position in high-inequality stocks and a short position in low-inequality stocks (columns 5 and 6). Pay inequality is defined as the log-difference in earnings between managers and workers. Managerial compensation is calculated as the overall annual pay of the members of the board, excluding the CEO, while workers' pay is measured as the average annual wage paid in establishments that are located in the same state as the firm's headquarters, weighted by the number of employees in each establishment. To construct our test portfolios, we rank all stocks in pay inequality quantiles as measured in the previous year, and rebalance them at the beginning of each year. In any given year, we define high and low wage gap stocks, respectively, as the stocks that lie at the top and the bottom 30% (Panel A), 20% (Panel B), and 10% (Panel C) of the distribution. All the regressors are from Kenneth French's website and refer to Europe. Observations are monthly, and the sample period is from January 2002 to December 2011. t statistics are below the coefficients, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Panel A

30% Threshold	(1) H	(2) H	(3) L	(4) L	(5) H−L	(6) H−L
Alpha	-0.007	0.001	-0.010**	-0.008**	0.003	0.009**
	-1.60	0.30	-2.48	-2.02	0.68	2.15
MKT	1.193***	1.107***	0.900***	0.810***	0.293	0.297**
	4.47	6.80	6.43	7.67	1.40	2.41
SMB		-1.200**		0.351		-1.551**
		-2.04		1.42		-2.54
HML		-0.689*		-0.131		-0.559**
		-1.89		-0.50		-2.36
UMD		-0.459***		-0.328***		-0.131
		-3.06		-3.10		-1.45
Adj. R-squared	0.547	0.673	0.569	0.601	0.066	0.333
Observations	120	120	120	120	120	120

Panel B

20% Threshold	(1) H	(2) H	(3) L	(4) L	(5) H−L	(6) H−L
Alpha	-0.010**	-0.001	-0.015***	-0.010**	0.005	0.010*
	-2.07	-0.12	-2.70	-2.37	1.11	1.87
MKT	1.207***	1.128***	0.906***	0.815***	0.301	0.312***
	4.03	6.88	4.90	5.64	1.51	3.14
SMB		-1.318**		0.336		-1.653**
		-1.99		1.12		-2.36
HML		-0.743**		-0.370		-0.373
		-2.12		-0.97		-1.38
UMD		-0.455**		-0.433***		-0.023
		-2.27		-2.94		-0.22
Adj. R-squared	0.482	0.603	0.471	0.517	0.051	0.245
Observations	120	120	120	120	120	120

Panel C

10% Threshold	(1) H	(2) H	(3) L	(4) L	(5) H−L	(6) H−L
Alpha	-0.000	0.002	-0.018***	-0.019***	0.018***	0.021***
	-0.02	0.48	-3.15	-4.55	3.61	5.88
MKT	0.855***	0.705***	0.697***	0.649***	0.158	0.056
	11.86	9.20	3.97	5.04	0.90	0.35
SMB		-0.551***		0.866***		-1.416***
		-2.84		3.67		-5.26
HML		0.477**		0.021		0.457
		2.19		0.06		1.16
UMD		-0.225*		-0.112		-0.114
		-1.94		-0.98		-0.64
Adj. R-squared	0.539	0.590	0.320	0.354	0.010	0.163
Observations	108	108	108	108	108	108

Table 28: CEO-workers cash wage gap

CAPM and Carhart's (1997) four-factor model regressions of equally-weighted returns on a portfolio of stocks with high inequality between CEO and workers' pay (columns 1 and 2), a portfolio of stocks for which such inequality is low (columns 3 and 4), and a portfolio with a long position in high-inequality stocks and a short position in low-inequality stocks (columns 5 and 6). Pay inequality is defined as the log-difference in earnings between CEO and workers. CEO compensation is calculated as the annual cash salary, while workers' pay is measured as the average annual wage paid in establishments that are located in the same state as the firm's headquarters, weighted by the number of employees in each establishment. To construct our test portfolios, we rank all stocks in pay inequality quantiles as measured in the previous year, and rebalance them at the beginning of each year. In any given year, we define high and low wage gap stocks, respectively, as the stocks that lie at the top and the bottom 30% (Panel A), 20% (Panel B), and 10% (Panel C) of the distribution. All the regressors are from Kenneth French's website and refer to Europe. Observations are monthly, and the sample period is from January 2002 to December 2011. t statistics are below the coefficients, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Panel A

30% Threshold	(1) H	(2) H	(3) L	(4) L	(5) H−L	(6) H−L
Alpha	-0.004	0.001	-0.012**	-0.008**	0.008**	0.009***
	-1.38	0.44	-2.50	-2.24	2.09	2.91
MKT	1.011***	0.936***	0.897***	0.798***	0.114	0.137**
	5.28	8.54	5.50	6.45	1.34	2.07
SMB		-0.815**		0.189		-1.004**
		-2.23		0.67		-2.38
HML		-0.355		-0.237		-0.118
		-1.61		-0.70		-0.61
UMD		-0.309**		-0.388**		0.079
		-2.54		-2.46		0.99
Adj. R-squared	0.610	0.693	0.533	0.572	0.013	0.180
Observations	120	120	120	120	120	120

Panel B

20% Threshold	(1) H	(2) H	(3) L	(4) L	(5) H−L	(6) H−L
Alpha	-0.006	0.003	-0.014***	-0.010**	0.008	0.012**
	-1.20	0.48	-2.60	-2.28	1.57	2.43
MKT	1.082***	1.056***	0.925***	0.826***	0.156	0.230**
	3.31	5.97	5.19	6.29	0.67	2.19
SMB		-1.331**		0.211		-1.543**
		-1.97		0.67		-2.06
HML		-0.934**		-0.354		-0.580*
		-2.52		-0.95		-1.91
UMD		-0.389*		-0.441***		0.052
		-1.65		-3.25		0.40
Adj. R-squared	0.424	0.558	0.486	0.532	0.008	0.189
Observations	120	120	120	120	120	120

Panel C

10% Threshold	(1) H	(2) H	(3) L	(4) L	(5) H−L	(6) H−L
Alpha	-0.003	-0.001	-0.018**	-0.017***	0.016***	0.016***
	-0.89	-0.27	-2.37	-2.78	2.66	4.11
MKT	0.909***	0.800***	0.721***	0.537***	0.188*	0.263***
	9.98	9.96	4.62	5.93	1.82	3.35
SMB		-0.300*		0.591*		-0.891**
		-1.76		1.69		-2.32
HML		0.252		0.472		-0.220
		1.18		1.17		-0.64
UMD		-0.220*		-0.297***		0.077
		-1.90		-2.78		0.95
Adj. R-squared	0.581	0.600	0.270	0.304	0.021	0.068
Observations	108	108	108	108	108	108

on high wage gap stocks is near-zero for the portfolio thresholds of 30% (0.0%), 20% (0.2%), and 10% (0.3%). Low wage gap stocks, instead, yield negative and significant risk-adjusted returns, increasing in absolute value across the thresholds of 30% (-0.9%), 20% (-1.1%), and 10% (-1.3%). As a result, the long-short portfolios yield positive and significant risk-adjusted returns increasing from 0.9% for the 30% portfolios to 1.7% for the 10% portfolios. The results are virtually unchanged when reestimating these coefficients using April cash-only wage gaps (unreported).

Restricting the wage gap calculation to same-state establishments implies that we only consider compensation for 18% of workers in the sample. For robustness, then, we repeat the analysis using workers from all establishments to calculate wage gaps. The results are in Table 30, and remain unchanged. High wage gap stocks exhibit near-zero risk-adjusted returns for all thresholds including 30% (0.1%), 20% (0.3%), and 10% (0.4%). Low wage gap stocks, instead, yield negative and significant risk-adjusted returns, moderately increasing in absolute value from the 30% (-0.9%) to the 10% threshold (-1.3%). As a result, the long-short portfolios yield positive and significant risk-adjusted returns increasing from 1.0% for the 30% portfolios to 1.7% for the 10% portfolios. For robustness, we repeat the analysis using only wages from workers from establishments in states other than the headquarters', and obtain the same results (unreported).

We also look into the relation between wage gap and stock returns in the panel, by estimating Equation (41). The results are in Table 31. In columns (1) and (2), we run Fama-MacBeth regressions, with unadjusted and robust standard errors, respectively. In both cases, we find that the top 30% wage gap stocks earn 1.2% higher returns than the bottom 30%. In columns (3) and (4), we estimate simple ordinary least squares once without and once with year fixed effects regressions. The coefficients are still positive, but outside the rejection region. The results, then, are in line with the previous tests.

6. Conclusion

We explore the hypothesis that the difference in pay between managers and workers constitutes a signal on the quality of management, and ultimately of the company. We propose an asset pricing model with unsophisticated traders and short-sales constraints, in which the wage gap between the CEO and rank-and-file workers is set efficiently and increases with managerial skills. In equilibrium, we show that firms with lower wage gaps should trade at a premium. The overpricing should be even more pronounced in the presence of investors who dislike pay inequality within firms.

Using a unique data set on German firms' employee compensation, we find strong evidence for the model's predictions. Stocks with low pay inequality earn negative and highly significant risk-adjusted returns, both in CAPM regressions and in Carhart's four-factor model. The magnitudes are also economically large, ranging between 0.9% per month for the 30% threshold and 1.6% per month for the 10% threshold. Stocks with high pay inequality, instead, exhibit no mispricing. As a result, a portfolio with a long position in stocks with high wage gaps and a short position in stocks with low wage gaps earns positive and significant risk-adjusted returns.

The empirical evidence suggests that the wage gap arises from efficient contracting, thus ruling out

Table 29: CEO-workers wage gap: April portfolios

CAPM and Carhart's (1997) four-factor model regressions of equally-weighted returns on a portfolio of stocks with high inequality between CEO and workers' pay (columns 1 and 2), a portfolio of stocks for which such inequality is low (columns 3 and 4), and a portfolio with a long position in high-inequality stocks and a short position in low-inequality stocks (columns 5 and 6). Pay inequality is defined as the log-difference in earnings between CEO and workers. CEO compensation is calculated as the overall annual pay, including cash and stocks, while workers' pay is measured as the average annual wage paid in establishments that are located in the same state as the firm's headquarters, weighted by the number of employees in each establishment. To construct our test portfolios, we rank all stocks in pay inequality quantiles as measured in the previous year, and rebalance them in April each year. In any given year, we define high and low wage gap stocks, respectively, as the stocks that lie at the top and the bottom 30% (Panel A), 20% (Panel B), and 10% (Panel C) of the distribution. All the regressors are from Kenneth French's website and refer to Europe. Observations are monthly, and the sample period is from January 2002 to December 2011. t statistics are below the coefficients, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Panel A

30% Threshold	(1)	(2)	(3)	(4)	(5)	(6)
	H	H	L	L	H−L	H−L
Alpha	-0.005*	-0.000	-0.013**	-0.009**	0.008**	0.009**
	-1.76	-0.00	-2.50	-2.48	2.00	2.43
MKT	1.052***	0.992***	0.905***	0.822***	0.148*	0.170***
	5.49	9.37	5.75	6.40	1.87	3.17
SMB		-0.781**		0.234		-1.015**
		-2.14		0.79		-2.26
HML		-0.378*		-0.301		-0.078
		-1.82		-0.81		-0.30
UMD		-0.281**		-0.375***		0.095*
		-2.10		-2.79		1.79
Adj. R-squared	0.627	0.697	0.537	0.575	0.025	0.176
Observations	120	120	120	120	120	120

Panel B

20% Threshold	(1)	(2)	(3)	(4)	(5)	(6)
	H	H	L	L	H−L	H−L
Alpha	-0.007	0.002	-0.015***	-0.011***	0.008*	0.013***
	-1.40	0.30	-3.03	-2.86	1.80	2.67
MKT	1.125***	1.072***	0.896***	0.830***	0.229	0.242**
	3.49	6.03	4.83	5.84	1.03	2.24
SMB		-1.467**		0.347		-1.814***
		-2.20		1.23		-2.71
HML		-0.838**		-0.431		-0.407
		-2.23		-1.21		-1.54
UMD		-0.415**		-0.394**		-0.021
		-1.97		-2.42		-0.21
Adj. R-squared	0.452	0.600	0.466	0.506	0.026	0.268
Observations	120	120	120	120	120	120

Panel C

10% Threshold	(1)	(2)	(3)	(4)	(5)	(6)
	H	H	L	L	H−L	H−L
Alpha	-0.006	0.003	-0.012	-0.013**	0.007	0.017***
	-1.00	0.69	-1.48	-2.05	0.58	3.59
MKT	1.129***	1.021***	0.757***	0.696***	0.371	0.325
	3.30	4.72	4.18	4.75	0.86	1.11
SMB		-1.580**		0.953***		-2.533***
		-2.50		3.31		-3.31
HML		-0.714		0.214		-0.928*
		-1.63		0.83		-1.95
UMD		-0.523**		-0.132**		-0.391*
		-2.24		-2.00		-1.79
Adj. R-squared	0.423	0.586	0.309	0.353	0.036	0.322
Observations	117	117	117	117	117	117

Table 30: CEO-workers wage gap: All establishments

CAPM and Carhart's (1997) four-factor model regressions of equally-weighted returns on a portfolio of stocks with high inequality between CEO and workers' pay (columns 1 and 2), a portfolio of stocks for which such inequality is low (columns 3 and 4), and a portfolio with a long position in high-inequality stocks and a short position in low-inequality stocks (columns 5 and 6). Pay inequality is defined as the log-difference in earnings between CEO and workers. CEO compensation is calculated as the overall annual pay, including cash and stocks, while workers' pay is measured as the average annual wage paid in all establishments, weighted by the number of employees in each establishment. To construct our test portfolios, we rank all stocks in pay inequality quantiles as measured in the previous year, and rebalance them at the beginning of each year. In any given year, we define high and low wage gap stocks, respectively, as the stocks that lie at the top and the bottom 30% (Panel A), 20% (Panel B), and 10% (Panel C) of the distribution. All the regressors are from Kenneth French's website and refer to Europe. Observations are monthly, and the sample period is from January 2002 to December 2011. t statistics are below the coefficients, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Panel A

30% Threshold	(1) H	(2) H	(3) L	(4) L	(5) H−L	(6) H−L
Alpha	-0.004	0.001	-0.014***	-0.009***	0.010***	0.010***
	-1.44	0.31	-2.71	-2.73	2.70	3.40
MKT	1.019***	0.949***	0.931***	0.845***	0.088	0.104*
	5.21	8.26	5.81	6.86	0.85	1.68
SMB		-0.891***		0.180		-1.071***
		-2.70		0.60		-3.09
HML		-0.307		-0.339		0.032
		-1.42		-0.84		0.12
UMD		-0.269**		-0.398**		0.129*
		-2.08		-2.56		1.88
Adj. R-squared	0.618	0.700	0.544	0.586	0.004	0.182
Observations	120	120	120	120	120	120

Panel B

20% Threshold	(1) H	(2) H	(3) L	(4) L	(5) H−L	(6) H−L
Alpha	-0.006	0.003	-0.017***	-0.013***	0.011**	0.016***
	-1.16	0.65	-3.37	-3.17	2.19	3.16
MKT	1.111***	1.014***	0.963***	0.867***	0.148	0.147
	3.48	5.67	5.37	6.36	0.61	1.19
SMB		-1.626***		0.373		-1.999***
		-2.69		1.18		-3.20
HML		-0.704*		-0.307		-0.397
		-1.90		-0.78		-1.54
UMD		-0.466**		-0.422***		-0.043
		-2.30		-2.81		-0.41
Adj. R-squared	0.435	0.605	0.498	0.539	0.005	0.276
Observations	120	120	120	120	120	120

Panel C

10% Threshold	(1) H	(2) H	(3) L	(4) L	(5) H−L	(6) H−L
Alpha	0.002	0.004	-0.016**	-0.013*	0.018***	0.017***
	1.18	1.54	-2.37	-1.91	2.81	3.01
MKT	0.739***	0.669***	0.980***	0.880***	-0.241	-0.211
	7.77	5.65	4.40	5.11	-0.98	-0.94
SMB		-0.827***		0.843**		-1.670***
		-5.56		2.15		-4.10
HML		0.249		-0.322		0.571
		1.00		-0.63		1.56
UMD		-0.109		-0.493		0.385
		-1.06		-1.45		1.19
Adj. R-squared	0.547	0.619	0.354	0.403	0.021	0.179
Observations	108	108	108	108	108	108

Table 31: CEO-workers wage gap: Panel regressions

Panel regressions from Brennan et al. (1998) of returns on German stocks on a dummy variable that takes on value one if firm i's wage gap is among the top 30% in the previous year, and a vector of firm characteristics, which includes: the log of the book-to-market ratio (calculated each July and held constant through the following June), the ratio of dividends in the previous fiscal year to market value at calendar year-end (calculated each July and held constant through the following June), the log of cumulative returns over months $t-3$ through $t-2$, months $t-6$ through $t-4$, and months $t-12$ through $t-7$, size (defined as the log of market capitalization at the end of month $t-2$), the log of the dollar volume of trading in the stock in month $t-2$, and the log of the stock price at the end of month $t-2$. We estimate Fama and MacBeth regressions in columns (1) and (2), ordinary least square regressions in column (3), and regressions with year fixed effects and robust standard errors in column (4). In all specifications, we exclude the middle 40% wage gap stocks. CEO compensation is measured as the CEO pay, including cash and stocks, while workers' pay is measured as the annual wage paid in establishments that are located in the same state as the firm's headquarters, calculated as a weighted average across establishments, where the weights are represented by the number of employees in each establishment. Data on workers' wages are from the "Establishment History Panel" database. Company-level accounting and stock market data from Worldscope and Datastream, and CEO compensation data from the companies' annual reports. Observations are monthly, and the sample period is from January 2002 to December 2011. t statistics are below the coefficients, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(1)	(2)	(3)	(4)
	Ret	Ret	Ret	Ret
Wage gap (d)	0.012*	0.012*	0.003	0.005
	1.74	1.80	0.41	1.19
Book-to-market	-0.025	-0.025	0.043	0.007
	-1.04	-1.10	1.37	0.25
Dividend yield	-0.004	-0.004	-0.008**	-0.003
	-0.94	-0.90	-2.23	-1.46
CumRet (2,3)	0.011	0.011	0.027*	-0.031
	0.48	0.49	1.87	-0.69
CumRet (4,6)	0.032**	0.032**	0.056***	0.031
	1.99	1.99	4.65	1.68
CumRet (7,12)	0.033***	0.033***	0.004	0.015
	3.43	3.33	0.46	0.75
Size (-2)	0.001	0.001	-0.001	0.000
	0.21	0.25	-0.35	-0.07
Trading volume (-2)	-0.004	-0.004	0.001	-0.001
	-1.02	-1.25	0.62	-0.25
Stock price (-2)	-0.007	-0.007*	-0.004	0.000
	-1.56	-1.81	-0.97	-0.01
Constant	0.006	0.006	-0.052	-0.027
	0.16	0.17	-1.30	-1.05
R-squared	0.598	0.598	0.015	0.006
Observations	2,331	2,331	2,331	2,450
Fama-MacBeth	Y	Y	Y	Y
Robust st. errors	N	Y	N	Y
Year fixed effects	N	N	N	Y

two alternative stories. First, the wage gap might be uninformative, and the mispricing might be entirely driven by inequality-averse investors. However, we show that there is a positive relation between the wage gap and firm performance. Second, a high wage gap might be the result of rent extraction. This implies a reversal of the types, as high wage gap firms exhibit agency issues while low wage gap firms are efficient. As a result, high wage gap stocks should be overpriced, which is at odds with our findings.

The present work contributes to the debate over pay inequality within firms, which has received significant attention from academics, regulators, and the media (see, e.g., Wade, O'Reilly, and Pollock (2006)). Recent research shows that pay inequality does not hinder firm performance (Cronqvist, Heyman, Nilsson, Svaleryd, and Vlachos, 2009), and actually improves it (Faleye, Reis, and Venkateswaran, 2013; Mueller, Ouimet, and Simintzi, 2017). However, this information is not fully impounded in the stock price. In this paper, we show that a story on limits of arbitrage does not fully explain the wage gap.

In particular, the findings suggest that at least some sophisticated market participants are able to infer information on salaries, and correctly incorporate it into their evaluations. Given the ample coverage that pay inequality receives in the media, however, this information may also be available to some unsophisticated investors, who could use it to discriminate against firms with high wage gaps and reward firms with low wage gaps. We refer to this as an inequality-aversion bias. The empirical evidence is consistent with this conjecture that investors exhibit inequality-aversion.

To the best of our knowledge, this is the first paper to show that investors exhibit inequality aversion. This finding contributes to the recent literature on values in asset pricing, which shows that investors also consider nonmonetary values in their trading strategies (see, e.g., Hong and Kostovetsky (2012)). The mechanism through which investors incorporate intangible information in the stock price also adds new insights to the literature on market learning (see, e.g., Edmans (2011)).

This paper can pave the way to future studies. It will be instructive to conduct this type of research on U.S. data in a few years' time, in light of the SEC new rule requiring firms to disclose the pay ratio between CEO and employees as of January 1, 2017. This regulatory change will also allow researchers to work with a much larger set of stocks.

Appendix A

Proof of Proposition 1

A.1. Firms' first-order conditions

If contracting is efficient, the objective function can be expressed as:

$$\max_{e,L} \pi(e, L) = \theta(e)K^{1-\alpha}L^{\alpha} - w(e)L - W(e), \tag{42}$$

The first-order condition with respect to e is:

$$\frac{d\pi(e,L)}{de} = \frac{\theta\left(e\right)}{de}K^{1-\alpha}L^{\alpha} - \frac{w\left(e\right)}{de}L - \frac{dW\left(e\right)}{de} = 0 \tag{43}$$

Using the following elasticity definitions:

$$\epsilon_{\theta} \equiv \frac{\partial\theta(e)}{\partial e}\frac{e}{\theta(e)} > 0, \tag{44}$$

$$\epsilon_{W} \equiv \frac{\partial W(e)}{\partial e}\frac{e}{W(e)} > 0, \tag{45}$$

$$\epsilon_{w} \equiv \frac{\partial w(e)}{\partial e}\frac{e}{w(e)} < 0, \tag{46}$$

then optimal managerial compensation is:

$$W\left(e^{*}\right) = \frac{\epsilon_{\theta}y\left(e^{*},L\right) - \epsilon_{w}w\left(e^{*}\right)L}{\epsilon_{W}} \equiv \frac{\epsilon_{\theta}y\left(e^{*},L\right) + |\epsilon_{w}|w\left(e^{*}\right)L}{\epsilon_{W}}, \tag{47}$$

where ϵ_W represents the price of managerial effort. Note that the optimal compensation for the manager has both a variable and a fixed component, calculated as a fraction of the firm's cash flow and labor costs, respectively.
The first-order condition with respect to labor is:

$$\frac{\partial\pi(e,L)}{\partial L} = \alpha L^{\alpha-1}\theta(e)K^{1-\alpha} - w(e) = 0, \tag{48}$$

which yields

$$L^{*} = \left(\frac{\alpha\theta(e^{*})}{w(e^{*})}\right)^{\frac{1}{1-\alpha}}K. \tag{49}$$

Optimal managerial compensation can then be expressed as

$$\begin{aligned} W\left(e^{*}\right) &= \frac{\epsilon_{\theta}}{\epsilon_{W}}\underbrace{\theta\left(e^{*}\right)K^{1-\alpha}\left(\frac{\alpha\theta\left(e^{*}\right)}{w\left(e^{*}\right)}\right)^{\frac{\alpha}{1-\alpha}}K^{\alpha}}_{\equiv y(e^{*},L^{*})} + \frac{|\epsilon_{w}|}{\epsilon_{W}}\underbrace{w\left(e^{*}\right)\left(\frac{\alpha\theta\left(e^{*}\right)}{w\left(e^{*}\right)}\right)^{\frac{1}{1-\alpha}}K}_{\equiv w(e^{*})L^{*}} = \\ &= K\left(\theta\left(e^{*}\right)\right)^{\frac{1}{1-\alpha}}\left(\frac{\alpha}{w\left(e^{*}\right)}\right)^{\frac{\alpha}{1-\alpha}}\frac{\epsilon_{\theta}+|\epsilon_{w}|\alpha}{\epsilon_{W}}, \end{aligned} \tag{50}$$

which is an increasing function of e, and therefore represents a signal for managerial effort. In particular, the firm chooses the optimal level of effort e^* by trading off the cost of managerial effort (ϵ_W) with its benefits ($\epsilon_\theta, |\epsilon_w|$).

Using the two first-order conditions, the firm's profits can be expressed as:

$$\pi(e^*,L^*) = \underbrace{\theta(e^*)K^{1-\alpha}\left(\frac{\alpha\theta(e^*)}{w(e^*)}\right)^{\frac{\alpha}{1-\alpha}}K^{\alpha}}_{\equiv y(e^*,L^*)} - \underbrace{w(e^*)\left(\frac{\alpha\theta(e^*)}{w(e^*)}\right)^{\frac{1}{1-\alpha}}K}_{\equiv w(e^*)L^*} + \underbrace{-K\left(\theta\left(e^*\right)\right)^{\frac{1}{1-\alpha}}\left(\frac{\alpha}{w\left(e^*\right)}\right)^{\frac{\alpha}{1-\alpha}}\frac{\epsilon_\theta+|\epsilon_w|\alpha}{\epsilon_W}}_{\equiv W(e^*)} = \tag{51}$$

$$= K\left(\theta\left(e^*\right)\right)^{\frac{1}{1-\alpha}}\left(\frac{\alpha}{w\left(e^*\right)}\right)^{\frac{\alpha}{1-\alpha}}\left(1-\alpha-\frac{\epsilon_\theta+|\epsilon_w|\alpha}{\epsilon_W}\right) \equiv \phi\left(e^*\right)K,$$

where $\phi(e^*)$ represents the profits per euro invested. Note that $\phi(e^*) \geq 1$ for the project to be started. From Equations (50) and (51), it is easy to see that high managerial effort leads to higher total compensation, a higher wage gap with respect to rank-and-file workers, and higher profitability.

A.2. Investors' first-order conditions

The first-order condition yields the following demand function for investor i in stock j:

$$n_{ij}^* = \gamma(E_i(v_j) - p_j), \tag{52}$$

Investors' evaluations are as follows. Arbitrageurs infer e^* from the wage gap, and correctly estimate:

$$E_A(\pi(e,L)) = \phi(e^*)K. \tag{53}$$

Noise traders neglect wage gaps and therefore have "flat" evaluations, using ϕ as an unconditional estimate instead:

$$E_U(\pi(e,L)) = \phi K. \tag{54}$$

Inequality-averse traders do observe wage gaps but exhibit a cognitive bias, as they perceive stocks with high (low) wage gaps less (more) favorably than noise traders. This translates into a bias towards firms with extreme profitability:

$$E_I(\pi(e,L)) = \phi(e^*)K + b(\phi - \phi(e^*))K, \tag{55}$$

where $b > 0$ is a parameter that captures the degree of inequality aversion, and the (unconditional) average wage gap ϕ is used as a reference point.

A.3. Equilibrium prices

The market clearing condition for a generic stock is:

$$
\begin{aligned}
q = \gamma(\phi(e^*)K - p(e^*))(1 - \lambda_U - \lambda_I) + \gamma(\phi K - p(e^*))\lambda_U + \\
+\gamma(\phi(e^*)K + b(\phi - \phi(e^*))K - p(e^*))\lambda_I.
\end{aligned}
\tag{56}
$$

Solving out for the equilibrium price:

$$
p(e^*) = K\phi(e^*) + (\phi - \phi(e^*))K\lambda_U + b(\phi - \phi(e^*))K\lambda_I - \frac{q}{\gamma}. \tag{57}
$$

It is easy to see that when $\phi < \phi(e^*)$, short-sales constraints are binding and the equilibrium price reduces to Equation (31).

Chapter V

Conclusion

Executive compensation is a tool to align the interest of shareholders and managers. High CEO compensation also arouses great attention from different groups of people in the society. Although previous studies provide abundant results concerning the optimal design and potential effects on executive compensation, there are still some phenomenon remains unsolved. In this thesis, I provide models and empirical analysis to explain the questions that are not answered by previous literature.

The first question is why stock option is a component of the CEO compensation packages. Previous literature shows that if the CEO is risk-averse, firms should provide low fixed salary and no options to CEOs (Dittmann and Maug, 2007). I provide an answer to this question by incorporating the probability weighting into the model. I approximate the probability weighting function with parameters that shift the shape of the normal distribution. I show that the optimal general contract contains options. I also calibrate the model with a wide range of parameters using the observed U.S. CEO compensation contracts. I find that the model with the probability weighting can explain the shape of the observed contracts better than the normal CRRA model. My finding gives an explanation to the option component in CEO's compensation packages.

The second question is what is the effect of CEO compensation on rank-and-file employee's pay. I establish a principal-agent model with one principal and two agents: the CEO and the employee who is behindness averse. I find that the wage of the employee is increasing in the CEO pay. This relationship is proved, both across firms and across time, by empirical tests on data set for German firms. I also find that the workers receive a significant increase in wage when CEO compensation is made public for the first time. The results suggest that the envy of workers associated with high CEO compensation brings extra costs to the firms.

The third question is whether the pay inequality between CEO and workers are priced on the stock markets. I set up an asset pricing model with noise traders and short-sales constraints, in which the optimal wage gap between the CEO and rank-and-file workers increases with managerial skills. In equilibrium, I show that firms with low wage gaps should be overpriced, and the effect should be even stronger in the presence of inequality-averse investors. The empirical tests show that a long-short portfolio of stocks with high and low wage gaps yields positive and robust risk-adjusted returns. This finding suggests that investors do trade on the pay inequality, and show that the mispricing comes from the overvaluation of low wage gap stocks.

To conclude, the CEO compensation not only acts as a tool to incentivize the CEO to maximize the shareholders' profits, but it also has effects on workers' wages and investors' decisions. All of these factors finally affect the value of the firm. The findings in this thesis can be a guidance to shareholders when they design the CEO compensation packages.

References

Aboody, David, and Baruch Lev, 1998, The value relevance of intangibles: The case of software capitalization, *Journal of Accounting Research* 36, 161–191.

Addison, John T., Alex Bryson, Paulino Teixeira, André Pahnke, and Lutz Bellmann, 2010, The state of collective bargaining and worker representation in Germany: The erosion continues, Discussion Paper 5030 Bonn.

Akerlof, George A, and Janet L Yellen, 1990, The fair wage-effort hypothesis and unemployment, *The Quarterly Journal of Economics* pp. 255–283.

Baker, Malcolm, and Jeffrey Wurgler, 2006, Investor sentiment and the cross-section of stock returns, *The Journal of Finance* 61, 1645–1680.

——— , 2007, Investor sentiment in the stock market, *Journal of Economic Perspectives* 21, 129–152.

Ben-David, Itzhak, John R Graham, and Campbell R Harvey, 2013, Managerial miscalibration, *The Quarterly Journal of Economics* 128, 1547–1584.

Bertrand, Marianne, and Sendhil Mullainathan, 2000, Agents with and without principals, *American Economic Review* 90, 203–208.

——— , 2001, Are CEOs rewarded for luck? The ones without principals are, *The Quarterly Journal of Economics* 116, 901–932.

Bhattacharya, Utpal, and Peter Groznik, 2008, Melting pot or salad bowl: Some evidence from US investments abroad, *Journal of Financial Markets* 11, 228–258.

Bizjak, John M, Swaminathan L Kalpathy, Zhichuan Frank Li, and Brian Young, 2017, The role of peer firm selection in explicit relative performance awards, .

Bloom, Matt, and John G. Michel, 2002, The relationships among organizational context, pay dispersion, and managerial turnover, *Academy of Management Journal* 45, 33–42.

Bolton, Gary E, and Axel Ockenfels, 2000, Erc: A theory of equity, reciprocity, and competition, *American Economic Review* pp. 166–193.

Brennan, Michael J, Tarun Chordia, and Avanidhar Subrahmanyam, 1998, Alternative factor specifications, security characteristics, and the cross-section of expected stock returns1, *Journal of Financial Economics* 49, 345–373.

Brown, Charles, and James Medoff, 1989, The employer size-wage effect, *Journal of Political Economy* 97, 1027–1059.

Card, David, Alexandre Mas, Enrico Moretti, and Emmanuel Saez, 2012, Inequality at work: The effect of peer salaries on job satisfaction, *American Economic Review* 102, 2981–3003.

Carhart, Mark M, 1997, On persistence in mutual fund performance, *The Journal of Finance* 52, 57–82.

Carpenter, Jennifer N, 1998, The exercise and valuation of executive stock options, *Journal of Financial Economics* 48, 127–158.

Chan, Louis KC, Josef Lakonishok, and Theodore Sougiannis, 2001, The stock market valuation of research and development expenditures, *The Journal of Finance* 56, 2431–2456.

Chen, Joseph, Harrison Hong, and Jeremy C Stein, 2002, Breadth of ownership and stock returns, *Journal of Financial Economics* 66, 171–205.

Clark, Andrew E, and Andrew J Oswald, 1996, Satisfaction and comparison income, *Journal of Public Economics* 61, 359–381.

Cronqvist, Henrik, Fredrik Heyman, Mattias Nilsson, Helena Svaleryd, and Jonas Vlachos, 2009, Do entrenched managers pay their workers more?, *The Journal of Finance* 64, 309–339.

Deng, Zhen, Baruch Lev, and Francis Narin, 1999, Science and technology as predictors of stock performance, *Financial Analysts Journal* 55, 20–32.

DeVault, Luke, Richard Sias, and Laura Starks, 2014, Who are the sentiment traders? Evidence from the cross-section of stock returns and demand, .

Dittmann, Ingolf, and Ernst Maug, 2007, Lower salaries and no options? On the optimal structure of executive pay, *The Journal of Finance* 62, 303–343.

———, and Oliver Spalt, 2010, Sticks or carrots? Optimal CEO compensation when managers are loss averse, *The Journal of Finance* 65, 2015–2050.

Dittmann, Ingolf, Ko-Chia Yu, and Dan Zhang, 2017, How important are risk-taking incentives in executive compensation?, *Review of Finance* 21, 1805–1846.

Dur, Robert, and Amihai Glazer, 2008, Optimal contracts when a worker envies his boss, *Journal of Law, Economics, and Organization* 24, 120–137.

Edmans, Alex, 2011, Does the stock market fully value intangibles? Employee satisfaction and equity prices, *Journal of Financial Economics* 101, 621–640.

———, Xavier Gabaix, and Dirk Jenter, 2017, *Executive compensation: A survey of theory and evidence* . pp. 383–539 (in B. Hermalin and M. Weisbach, eds.: Handbook of the Economics of Corporate Governance 1, North Holland, Amsterdam).

Ellguth, Peter, Hans-Dieter Gerner, and Jens Stegmaier, 2012, Wage bargaining in Germany: The role of works councils and opening clauses, IAB-Discussion Paper 5/2012 Nürnberg.

Faleye, Olubunmi, Ebru Reis, and Anand Venkateswaran, 2013, The determinants and effects of CEO-employee pay ratios, *Journal of Banking & Finance* 37, 3258–3272.

Fehr, Ernst, and Klaus M Schmidt, 1999, A theory of fairness, competition, and cooperation, *Quarterly Journal of Economics* pp. 817–868.

Gabaix, Xavier, and Augustin Landier, 2008, Why has CEO pay increased so much?, *The Quarterly Journal of Economics* 123, 49–100.

Gillan, Stuart L, Jay C Hartzell, and Robert Parrino, 2009, Explicit versus implicit contracts: Evidence from CEO employment agreements, *The Journal of Finance* 64, 1629–1655.

Graef, Crystal, 1991, In search of excess: The overcompensation of american executives, .

Grinblatt, Mark, and Matti Keloharju, 2001, How distance, language, and culture influence stockholdings and trades, *The Journal of Finance* 56, 1053–1073.

Grund, Christian, and Dirk Sliwka, 2005, Envy and compassion in tournaments, *Journal of Economics & Management Strategy* 14, 187–207.

Hassel, Anke, 1999, The erosion of the German system of industrial relations, *British Journal of Industrial Relations* 37, 483–505.

——— , and Britta Rehder, 2001, Institutional change in the German wage bargaining system: The role of big companies, MPIfG Working Paper 01/9 Max Planck Institute for the Study of Societies.

Hölmstrom, Bengt, 1979, Moral hazard and observability, *The Bell Journal of Economics* pp. 74–91.

Hong, Harrison, and Marcin Kacperczyk, 2009, The price of sin: The effects of social norms on markets, *Journal of Financial Economics* 93, 15–36.

Hong, Harrison, and Leonard Kostovetsky, 2012, Red and blue investing: Values and finance, *Journal of Financial Economics* 103, 1–19.

Huddart, Steven, and Mark Lang, 1996, Employee stock option exercises an empirical analysis, *Journal of Accounting and Economics* 21, 5–43.

Jung, Sven, and Claus Schnabel, 2011, Paying more than necessary? The wage cushion in Germany, *LABOUR* 25, 182–197.

Kaustia, Markku, and Sami Torstila, 2011, Stock market aversion? Political preferences and stock market participation, *Journal of Financial Economics* 100, 98–112.

Lev, Baruch, and Theodore Sougiannis, 1996, The capitalization, amortization, and value-relevance of r&d, *Journal of Accounting and Economics* 21, 107–138.

Lin, Chen, Thomas Schmid, and Yang Sun, 2016, Conflict or collusion? How employees in the boardroom affect compensation, Working Paper.

Main, Brian GM, Charles A O'Reilly, and James Wade, 1995, The CEO, the board of directors and executive compensation: Economic and psychological perspectives, *Industrial and Corporate Change* 4, 293–332.

Malmendier, Ulrike, and Geoffrey Tate, 2005, Does overconfidence affect corporate investment? CEO overconfidence measures revisited, *European Financial Management* 11, 649–659.

Morse, Adair, and Sophie Shive, 2011, Patriotism in your portfolio, *Journal of Financial Markets* 14, 411–440.

Mueller, Holger M, Paige P Ouimet, and Elena Simintzi, 2017, Within-firm pay inequality, *Review of Financial Studies*.

Neilson, William S, and Jill Stowe, 2010, Piece-rate contracts for other-regarding workers, *Economic Inquiry* 48, 575–586.

Newey, Whitney K, and Kenneth D West, 1987, Hypothesis testing with efficient method of moments estimation, *International Economic Review* pp. 777–787.

Otto, Clemens A, 2014, CEO optimism and incentive compensation, *Journal of Financial Economics* 114, 366–404.

Pfeffer, Jeffrey, and Nancy Langton, 1993, The effect of wage dispersion on satisfaction, productivity, and working collaboratively: Evidence from college and university faculty, *Administrative Science Quarterly* pp. 382–407.

Quiggin, John, 1982, A theory of anticipated utility, *Journal of Economic Behavior & Organization* 3, 323–343.

Rouen, Ethan, 2017, Rethinking measurement of pay disparity and its relation to firm performance, .

Schmitt, David R, and Gerald Marwell, 1972, Withdrawal and reward reallocation as responses to inequity, *Journal of Experimental Social Psychology* 8, 207–221.

Schnabel, Claus, and Joachim Wagner, 2007, The persistent decline in unionization in Western and Eastern Germany, 1980-2004: What can we learn from a decomposition analysis?, *Industrielle Beziehungen / The German Journal of Industrial Relations* 14, 118–132.

Spalt, Oliver G, 2013, Probability weighting and employee stock options, *Journal of Financial and Quantitative Analysis* 48, 1085–1118.

Stambaugh, Robert F, Jianfeng Yu, and Yu Yuan, 2012, The short of it: Investor sentiment and anomalies, *Journal of Financial Economics* 104, 288–302.

Tervio, Marko, 2008, The difference that CEOs make: An assignment model approach, *American Economic Review* 98, 642–68.

Tversky, Amos, and Daniel Kahneman, 1992, Advances in prospect theory: Cumulative representation of uncertainty, *Journal of Risk and Uncertainty* 5, 297–323.

Wade, James B, Charles A O'Reilly, and Timothy G Pollock, 2006, Overpaid CEOs and underpaid managers: Fairness and executive compensation, *Organization Science* 17, 527–544.

Yu, Jianfeng, and Yu Yuan, 2011, Investor sentiment and the mean–variance relation, *Journal of Financial Economics* 100, 367–381.

The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus University Rotterdam, University of Amsterdam and VU University Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. The following books recently appeared in the Tinbergen Institute Research Series:

676 F. GOMEZ MARTINEZ, Essays in Experimental Industrial Organization: How Information and Communication affect Market Outcomes
677 J.A. ATTEY, Causes and Macroeconomic Consequences of Time Variations in Wage Indexation
678 T. BOOT, Macroeconomic Forecasting under Regime Switching, Structural Breaks and High-dimensional Data
679 I. TIKOUDIS, Urban Second-best Road Pricing: Spatial General Equilibrium Perspectives
680 F.A. FELSŐ, Empirical Studies of Consumer and Government Purchase Decisions
681 Y. GAO, Stability and Adaptivity: Preferences over Time and under Risk
682 M.J. ZAMOJSKI, Panta Rhei, Measurement and Discovery of Change in Financial Markets
683 P.R. DENDERSKI, Essays on Information and Heterogeneity in Macroeconomics
684 U. TURMUNKH, Ambiguity in Social Dilemmas
685 U. KESKIN, Essays on Decision Making: Intertemporal Choice and Uncertainty
686 M. LAMMERS, Financial Incentives and Job Choice
687 Z. ZHANG, Topics in Forecasting Macroeconomic Time Series
688 X. XIAO, Options and Higher Order Risk Premiums
689 D.C. SMERDON, 'Everybody's doing it': Essays on Trust, Norms and Integration
690 S. SINGH, Three Essays on the Insurance of Income Risk and Monetary Policy
691 E. SILDE, The Econometrics of Financial Comovement
692 G. DE OLIVEIRA, Coercion and Integration
693 S. CHAN, Wake Me up before you CoCo: Implications of Contingent Convertible Capital for Financial Regulation
694 P. GAL, Essays on the role of frictions for firms, sectors and the macroeconomy
695 Z. FAN, Essays on International Portfolio Choice and Asset Pricing under Financial Contagion
696 H. ZHANG, Dealing with Health and Health Care System Challenges in China: Assessing Health Determinants and Health Care Reforms
697 M. VAN LENT, Essays on Intrinsic Motivation of Students and Workers
698 R.W. POLDERMANS, Accuracy of Method of Moments Based Inference
699 J.E. LUSTENHOUWER, Monetary and Fiscal Policy under Bounded Rationality and Heterogeneous Expectations
700 W. HUANG, Trading and Clearing in Modern Times
701 N. DE GROOT, Evaluating Labor Market Policy in the Netherlands
702 R.E.F. VAN MAURIK, The Economics of Pension Reforms
703 I. AYDOGAN, Decisions from Experience and from Description: Beliefs and Probability Weighting
704 T.B. CHILD, Political Economy of Development, Conflict, and Business Networks
705 O. HERLEM, Three Stories on Influence

706 J.D. ZHENG, Social Identity and Social Preferences: An Empirical Exploration
707 B.A. LOERAKKER, On the Role of Bonding, Emotional Leadership, and Partner Choice in Games of Cooperation and Conflict
708 L. ZIEGLER, Social Networks, Marital Sorting and Job Matching. Three Essays in Labor Economics
709 M.O. HOYER, Social Preferences and Emotions in Repeated Interactions
710 N. GHEBRIHIWET, Multinational Firms, Technology Transfer, and FDI Policy
711 H.FANG, Multivariate Density Forecast Evaluation and Nonparametric Granger Causality Testing
712 Y. KANTOR, Urban Form and the Labor Market
713 R.M. TEULINGS, Untangling Gravity
714 K.J.VAN WILGENBURG, Beliefs, Preferences and Health Insurance Behavior
715 L. SWART, Less Now or More Later? Essays on the Measurement of Time Preferences in Economic Experiments
716 D. NIBBERING, The Gains from Dimensionality
717 V. HOORNWEG, A Tradeoff in Econometrics
718 S. KUCINSKAS, Essays in Financial Economics
719 O. FURTUNA, Fiscal Austerity and Risk Sharing in Advanced Economies
720 E. JAKUCIONYTE, The Macroeconomic Consequences of Carry Trade Gone Wrong and Borrower Protection
721 M. LI, Essays on Time Series Models with Unobserved Components and Their Applications
722 N. CIURILĂ, Risk Sharing Properties and Labor Supply Disincentives of Pay-As-You-Go Pension Systems
723 N.M. BOSCH, Empirical Studies on Tax Incentives and Labour Market Behaviour
724 S.D. JAGAU, Listen to the Sirens: Understanding Psychological Mechanisms with Theory and Experimental Tests
725 S. ALBRECHT, Empirical Studies in Labour and Migration Economics