

Market Power in Hospital Markets



Selection in Health Insurance Markets

Ramsis Croes

Market Power in Hospital Markets and Selection in Health Insurance Markets

Ramsis Reynold Croes

ISBN: 978-94-6361-253-1

Lay-out and printing by Optima Grafische Communicatie

Market Power in Hospital Markets and Selection in Health Insurance Markets

Marktmacht in ziekenhuismarkten en selectie in zorgverzekeringsmarkten

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 16 May 2019 at 13:30 hrs

by

Ramsis Reynold Croes
born in Aruba

Erasmus University Rotterdam



Doctoral Committee:

Promoters:

Prof.dr. F.T. Schut

Prof.dr. M. Varkevisser

Other members:

Prof.dr. M. Lindeboom

Prof.dr. J. Boone

Prof.dr. W.P.M.M. van de Ven

Table of contents

Chapter 1	Introduction	7
Chapter 2	Price effects of a hospital merger: heterogeneity across health insurers, hospital products and hospital locations	17
Chapter 3	Back to the Future: Predictive Power of the Option Demand Method in the Dutch Hospital Industry	65
Chapter 4	Competition and quality indicators in the health care sector: empirical evidence from the Dutch hospital sector	97
Chapter 5	Evidence of selection in a mandatory health insurance market with risk adjustment	125
Chapter 6	Is adverse selection effectively mitigated by consumer inertia? Empirical evidence from the Dutch health insurance market	151
Chapter 7	Conclusion	175
	Samenvatting	183
	Portfolio & CV	191
	Acknowledgements	197

Chapter 1

Introduction

INTRODUCTION

In many countries, market mechanisms are used to generate incentives for efficiency in health care markets. However, markets may suffer from market failures, resulting in an inefficient use of resources. Economists can help to design and tweak markets such that the public goals that have been set are reached. It is thus important to understand these market failures and to advance policy proposals that can help to avoid these failures.

In most OECD countries, markets play an important role in the delivery of health care. Additionally, in several OECD countries markets play an important role in the financing of health care, by means of health insurance markets, as well. This is for example the case in the United States, the Netherlands, Switzerland and Israel (OECD 2018). Such health care markets consist of a complex interaction between consumers, insurers and providers. Generally, in these health systems consumers buy health plans from competing health insurers. These health plans give consumers access to a network of providers when they seek treatment. An insurer and a provider negotiate over the inclusion of the provider into the insurer's network. These negotiations result into contracts that include for example the reimbursement rates. Hospitals compete with each other to be included into an insurer's network based on their contracts. Health insurers compete with each other based on their health plan premiums and the breadth and quality of their provider networks (Gaynor et al., 2015).

A potentially important source of market failure in health care is insufficient competition between health care providers. In that case, health insurers are not able to bargain effectively with providers. This results in high reimbursement rates and, most likely, high premiums. Another potential source of market failure is that health insurance markets can be affected by selection problems and information frictions. There is selection in the health insurance market when insurers cannot price all risk characteristics and riskier consumers choose more comprehensive health plans (Handel, 2013). Adverse selection can induce insurers to offer health plans with the goal to encourage self-selection by healthy individuals. Furthermore, information frictions may hamper optimal consumer choice in health insurance markets. A mismatch between people's preferences and their health insurance choices is then created (Handel & Kolstad, 2015). This thesis examines these market failures in the context of the Dutch health care system. It focuses on (i) market power in the provider market and (ii) selection and inertia in the insurance market.

1. DUTCH HEALTH CARE SYSTEM

In the Dutch health care system, market mechanisms were stepwise introduced with a major market-oriented reform in 2006 (Helderman, Schut, Van der Grinten, & Van de Ven, 2005). This reform was largely based on the Dekker plan (Dekker, 1987), which was drafted almost 20 years earlier. With the reform, 'regulated competition' was introduced in the Dutch health system: i.e. a regulated market with competing providers and competing health insurers.

The current Dutch health care system is characterized by mandatory health insurance with open enrollment. Since 1 January 2006 each person who lives or works in the Netherlands is obliged to buy health insurance. Private health insurer companies offer health plans based on a standardized benefits package, which is defined by the government. Health insurers have to accept all applicants for each health plan that they offer at a community-rated premium. Each year, individuals can switch health plans during the annual open enrollment period (December-January). Insurers compete with each other mainly on premium and service. The government expects that the insurers become prudent buyers of care on behalf of their insured populations (Van de Ven & Schut, 2008). Since health insurers cannot refuse applicants or charge a higher premium for individuals that are expected to have high health care costs, they have an incentive to select individuals that are expected to have low costs. To mitigate this selection incentive, a risk-adjustment system is in place. This risk-adjustment system aims to eliminate as much as possible the predictable profits (losses) on low-risk profile (high-risk profile) individuals emanating from community rating.

Since 2005, health care providers gradually received more freedom to determine prices as well as the quantity and quality of production. For example, in the hospital market, which is largest market in terms of total expenditure, the percentage of revenue for which providers can freely set their prices increased from on average 10% in 2005 to about 70% in 2012. The role of the insurers in this regulated market is twofold. First, they offer health plans to citizens. Second, they negotiate contracts with health care providers concerning the delivery of health care. An important bargaining instrument for health insurers is selective contracting. When offering health plans, insurers can limit the set of contracted providers for which the costs are fully reimbursed. Hence, they are allowed to contract selectively. This threat empowers insurers to negotiate lower prices or higher quality.

2. HOSPITAL MARKET FAILURE

The first potential market failure discussed in this thesis is market power in the provider market. We focus on hospital markets, which accounts for the largest share (i.e. more than one third) of total healthcare expenditure in the Netherlands¹. Since the 1980s, hospital sectors in many OECD countries have become increasingly concentrated as a result of mergers (Gaynor, Ho & Town, 2015). In hospital markets, where competition should create incentives for efficiency (Gaynor & Town, 2011), an increase in market concentration means that there may be less competitive pressure in the market. This may result in higher prices and/or lower quality.

An increasing number of empirical studies have investigated the price effects of hospital mergers. Most of these studies have shown that although mergers may bring about meaningful reductions in marginal costs, mergers between rival hospitals are likely to raise prices in concentrated hospital markets (Gaynor & Town, 2011; Gaynor et al., 2015).

In many developed countries, an important policy instrument to prevent dominant positions (market power) is merger control. Competition authorities are entrusted with this instrument to prohibit anticompetitive mergers in order to keep markets competitive. However, the traditional approaches used to prospectively review mergers are especially problematic in hospital markets. Generally, these methods first define the relevant market for the industry being studied and then use market shares to infer how the merger could affect competition in that market (Shapiro, 2010). In order to delineate the relevant market, they typically rely on disputed methodologies and the conclusions drawn from the resulting analysis typically depends heavily on how that market is defined. Given the critical role for merger control in keeping hospital markets competitive, it is important to develop reliable methods that competition authorities can implement to effectively review mergers. A recent development is the use of merger simulation models to assess the competitive effects of hospital mergers.

My first three research questions are about market concentration in the Dutch hospital industry and its (potential) implications for market outcomes. Due to a large number of hospital mergers, hospital market concentration in the Netherlands has substantially increased. From 2004 to (August) 2018, 30 hospital mergers were reviewed by the Dutch Competition Authority. From these 30 reviewed mergers, one merger was forbidden by the Competition Authority, one was accepted with remedies and three approved mergers were in the end not realized by the merging parties. Currently (August 2018) there are 68 hospital organizations in the Netherlands. Due to the above described

1 In 2018 total hospital expenditure is projected at 24.9 billion euro, which is 36% of total projected expenditure on curative and long-term health care (69.2 billion euro). These figures are derived from the 2019 Budget Memorandum, available at <http://www.rijksbegroting.nl/rijksbegroting2019.html>.

merger activities, 21 of the current 68 hospital organizations consists of hospitals that have merged between 2004 and (August) 2018². This indicates that the concentration in the Dutch hospital market has substantially increased during the last 14 years. In this thesis we will: (i) investigate the effect of a hospital merger on prices, (ii) evaluate the use of a merger simulation model to assess the price-effects of a hospital merger, and (iii) examine the relationship between hospital market concentration and quality of care.

To date, research on the causal effect of mergers on prices in the Dutch hospital market is limited. By examining the effects of one specific hospital merger case on various hospital products in depth, we want to determine how mergers may affect prices in the Dutch hospital market. In most hospital merger studies, the unit of observation is the merged hospital. That is, hospital products are aggregated and prices are averaged across different payers. In practice, however, price effects of a hospital merger may vary between different hospital locations, different products and different insurers. Therefore, my first research question is (discussed in Chapter 2): ***For a merger between two neighboring hospitals in the Netherlands, do the prices increase after a merger and do we observe differential price changes between different hospital locations, different products and different insurers?***

As discussed above, recently, economists have developed new methods for reviewing mergers: merger simulation models. Specifically for hospital merger simulation, the Option Demand Method has been developed. This method uses a Nash-bargaining framework to model the hospital-insurer interactions, where the added value that a hospital brings to a health plan network is measured through a logit demand model. However, there is still not much empirical evidence that such simulation models are able to produce reliable predictions. An approach to evaluate the predictive power of a merger simulation model is to compare the predicted price effects from the simulation model with the actual price effect of a consummated hospital merger. Hence, my second research question is (discussed in Chapter 3): ***What is the predictive power of the Option Demand Method for mergers in the Dutch hospital market?*** We examine this question by contrasting the prediction of the simulation model with the post-merger price effects examined in Chapter 2.

Next to prices, quality of care may also be affected by (changes in) the level of competition. Gaynor and Town (2011) argue that in hospital markets without regulated prices the impact of competition on quality is ambiguous. This is supported by empirical evidence from the US and UK. To date empirical evidence about the relationship between competition and quality in the Netherlands is lacking. My third research question is

2 Author's own calculation based on data on hospital organizations from Volksgezondheidenzorg.info and a merger list from the Authority for Consumers and Markets (ACM). For a detailed historical description of merger activities in the Dutch hospital market, see chapter 2 in Roos (2018).

(discussed in Chapter 4): ***What is the relationship between competition and quality indicators in the Dutch non-price-regulated hospital market?*** We examine this question by analyzing three diagnosis groups (cataract, adenoid and tonsils, bladder tumor) delivered by Dutch hospitals in the period 2008-2011. For these three diagnoses the Health Inspectorate (current name: Dutch Health and Youth Care Inspectorate) published for each hospital outcome, process and structure indicators. Using these indicators, we evaluate the relationship between hospitals' quality scores and market concentration.

3. MARKET FAILURES IN HEALTH INSURANCE

Health insurance markets play an important role in the financing of health care in various countries, including the United States, Netherlands, Switzerland and Israel. In these countries, consumers buy health plans from insurers that give them access to health care in case they fall ill. Insurers bargain with providers to create provider networks for their enrollees. It is thus important to have an efficient health insurance market that give health insurers appropriate incentives to buy health care efficiently and to offer health plans that matches the preferences of the enrollees. As discussed above, the presence of adverse selection and inertia are impediments to an efficient health insurance market.

For many people, choosing the right health plan is a complex task. First, people have to search for relevant information and have to understand all the dimensions of the health plans (such as provider networks, reimbursement rates and deductible schemes). In addition, people have to predict their health care demand and expenses for the next year, since at the moment of choosing their health plan for the coming year, they do not exactly know what their health care needs will be. There is a growing number of studies finding that most consumers in health insurance markets are poorly informed and often do not actively and carefully evaluate their health plan choices (Ericson & Sydnor, 2017). Optimal consumer choice in health insurance markets is hampered by both frictions (inertia, search and switching cost) and a lack of knowledge (health insurance illiteracy); see for example Abaluck & Gruber, 2011, Bhargava, Loewenstein, & Sydnor, 2015, Handel, 2013, and Handel & Kolstad, 2015. A mismatch between the theoretical optimal choices and the actual choices of the individuals can lower consumer welfare.

On the other hand, well-informed consumer choices can lead to adverse selection when high-risk consumers choose more comprehensive health plans and/or low-risk consumers choose less comprehensive health plans. Rothschild and Stiglitz (1976) show that adverse selection can result in the underinsurance of low-risk individuals and even in a market without a stable equilibrium.

In this thesis, we identify and examine inertia and selection in the context of voluntary deductibles in the Dutch health insurance market. This analysis starts with the empirical

identification of the exact magnitude of selection, which is challenging, because of the interaction of selection with moral hazard (Cohen & Siegelman, 2010). The mere fact that people who opt for a voluntary deductible have lower health care expenses does not, in itself, prove the presence of adverse selection, because the deductible may also induce more cost conscious behavior (i.e. less moral hazard). Importantly, in the Dutch health insurance market, which contains an comprehensive risk-adjustment system, it is a priori not clear if contracts with voluntary deductibles are profitable for insurers given the prevailing risk-adjustment system, since the goal of the system is to eliminate the predictable gains (losses) on low-risk profile (high-risk profile) individuals. My fourth research question is thus (discussed in Chapter 5): ***How large is the selection effect for individuals who chose voluntary deductibles in a health insurance market with risk-adjustment?*** By examining people's deductible choice taking account of their prior health care expenses, we are able to identify the selection effect separately from moral hazard effect. In addition, we examine whether people opting for voluntary deductibles are over- or undercompensated by the risk adjustment scheme.

As shown by Handel (2013) and Handel et al. (2015), adverse selection may be counteracted by consumer inertia. Hence, in case of consumer inertia the welfare loss due to suboptimal health plan choice may be compensated by a welfare gain due to less adverse selection. This interaction between adverse selection and consumer inertia is the subject of my fifth and final research question (discussed in Chapter 6): ***Is adverse selection effectively mitigated by consumer inertia in the Dutch health insurance market?*** By examining how previous as well as future health costs affect the choices individuals make regarding their voluntary deductible, we try to determine the extent of inertia and adverse selection regarding the uptake of voluntary deductible contracts. The Dutch market for mandatory basic health insurance has an interesting feature: i.e. people can annually adjust the voluntary deductible without changing health plans. This feature can be exploited for an empirical identification of the relationship between prior health care expenses and deductible choice, which can provide an indication of the extent of inertia and adverse selection in deductible choice.

REFERENCES

- Abaluck, J., & Gruber, J. (2011). Choice inconsistencies among the elderly: evidence from plan choice in the medicare part d program. *American Economic Review*, 101 (4), 1180–1210.
- Bhargava, S., Loewenstein, G., & Sydnor, J. (2015). Do individuals make sensible health insurance decisions? evidence from a menu with dominated options. National Bureau of Economic Research.
- Capps, C., Dranove, D., & Satterthwaite, M. (2003). Competition and market power in option demand markets. *RAND Journal of Economics*, 737–763
- Cohen, A., & Siegelman, P. (2010). Testing for adverse selection in insurance markets. *Journal of Risk and Insurance*, 77 (1), 39–84.
- Dekker, C. (1987). *Bereidheid tot verandering (willingness to change)*. The Hague: Distributiecentrum Overheidspublicaties.
- Ericson, K. M., & Sydnor, J. (2017). The questionable value of having a choice of levels of health insurance coverage. *Journal of Economic Perspectives*, 31 (4), 51–72.
- Gaynor, M., Ho, K., & Town, R. J. (2015). The industrial organization of health-care markets. *Journal of Economic Literature*, 53 (2), 235–84.
- Gaynor, M., & Town, R. (2011). Competition in health care markets. In T. G. M. Mark V. Pauly & P. P. Barros (Eds.), *Handbook of health economics* (Vol. 2, p. 499 - 637). Elsevier.
- Gowrisankaran, G., Nevo, A., & Town, R. (2015). Mergers when prices are negotiated: Evidence from the hospital industry. *American Economic Review*, 105 (1), 172–203.
- Handel, B. (2013). Adverse selection and inertia in health insurance markets: When nudging hurts. *American Economic Review*, 103 (7), 2643–82.
- Handel, B. R., & Kolstad, J. T. (2015). Health insurance for “humans”: Information frictions, plan choice, and consumer welfare. *American Economic Review*, 105 (8), 2449–2500.
- Handel, B. R., Kolstad, J. T., & Spinnewijn, J. (2015). Information frictions and adverse selection: Policy interventions in health insurance markets (Tech. Rep.). National Bureau of Economic Research.
- Helderman, J.-K., Schut, F., van der Grinten, T., & van de Ven, W. (2005). Market-oriented health care reforms and policy learning in the netherlands. *Journal of Health Politics, Policy and Law*, 30 (1-2), 189–210.
- OECD (2018). *Health Statistics - Social Protection*, OECD Publishing, Paris. https://stats.oecd.org/index.aspx?DataSetCode=HEALTH_STAT#. Retrieved on 14 Sep 2018 13:49 UTC (GMT).
- Rothschild, M., & Stiglitz, J. (1976). Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *Quarterly Journal of Economics*, 90(4), 629–649.
- Roos, A.F. (2018). *Mergers and Competition in the Dutch Healthcare Sector* (Doctoral dissertation). Retrieved from <https://repub.eur.nl/pub/108851>.
- Shapiro, C. (2010). The 2010 horizontal merger guidelines: From hedgehog to fox in forty years. *Antitrust Law Journal*, 77 (1), 49–107.
- van de Ven, W., & Schut, F. (2008). Universal mandatory health insurance in the netherlands: a model for the United States? *Health Affairs*, 27 (3), 771–781.

Chapter 2

Price effects of a hospital merger: heterogeneity across health insurers, hospital products and hospital locations

With Anne-Fleur Roos, Victoria Shestalova, Marco Varkevisser and Frederik Schut

A version of this chapter has been submitted to a journal and is currently under review.

ACKNOWLEDGEMENTS

This project was partly funded by the Dutch Healthcare Authority. The authors would like to thank the Dutch Healthcare Authority for providing access to the data. We are also grateful to the following for their valuable comments on the article: two anonymous reviewers; C.-t. A. Ma, Ph.D. (Boston University); P.P. Barros, Ph.D. (Universidade Nova de Lisboa); R. Halbersma, Ph.D. (Dutch Gaming Authority); M. Mikkers, Ph.D. (Dutch Healthcare Authority); the participants in the CINCH academy: Essen Summer School in Health Economics (June 3-9, 2013 in Essen, Germany – in particular, A. Schmid, Ph.D. and A. Chorniy, Ph.D.), the participants in the Health Insurance Meeting (May 15, 2013 in Rotterdam, the Netherlands – in particular, R. van Kleef, Ph.D. and R. Douven, Ph.D.) and the participants in the International Invitational Conference on Competition and Antitrust in Hospital Markets (September 22-23, 2014 in Bayreuth, Germany).

ABSTRACT

In most studies on hospital merger effects, the unit of observation is the merged hospital, whereas the observed price is the weighted average across hospital products and across payers. However, little is known about whether price effects vary between hospital locations, products and payers. We expand existing bargaining models to allow for heterogeneous price effects and use a difference-in-differences model in which price changes at the merging hospitals are compared to price changes at comparison hospitals. We find evidence of heterogeneous price effects across health insurers, hospital products and hospital locations. These findings have implications for ex ante merger scrutiny.

1. INTRODUCTION

An increasing number of empirical studies have been conducted concerning the price effects of hospital mergers. In general, the aim of these studies is to test the effectiveness of antitrust policy. In competitive markets, the aim of preventive merger control is to prohibit anticompetitive consolidation. To determine whether a merger between two or more firms will result in anticompetitive price increases and/or quality decreases, antitrust authorities need to carry out a prospective review of the merger. However, merger reviews in the healthcare sector encounter specific difficulties because there are unique factors that render the most commonly used tests for measuring geographic markets less reliable in healthcare than in other sectors (Elzinga & Swisher, 2011). Retrospective studies are aimed at providing a better understanding of the effects of mergers, which, in turn may improve future antitrust policy.

The majority of the studies on retrospective merger analyses indicate a positive correlation between hospital mergers and prices (see e.g. Gaynor & Town, 2012; Vogt & Town, 2006; Gaynor & Vogt, 2000 for reviews). In most of these studies, the unit of observation is the merged hospital, whereas the observed price is the weighted average across different hospital products and across different payers. However, little is known about whether price effects vary between different hospital locations, different products and different payers. Because merged hospitals often continue to operate at different locations, produce multiple products and negotiate prices with a range of payers, an interesting question is whether these differences matter. If it turns out that they do matter, this may have important implications for *ex ante* merger scrutiny by antitrust authorities.

This article considers the question of whether the price effects of a hospital merger vary between locations, products and third-party payers (i.e. health insurers). By means of a hospital-insurer bargaining model, we show that the price effects of a hospital merger may vary and that the differences between locations, products and insurers may influence the outcome of hospital-insurer price setting differently. We show that the price effects differ between locations, products and insurers depending on: (i) the degree of substitution between the merging hospitals for different products, (ii) the relative bargaining ability of hospitals and insurers and (iii) the pre-merger price-cost margins. We then use a unique national dataset on hospital-insurer negotiated contract prices for each hospital product in the Netherlands to investigate whether the price effects of a merger between a general acute care hospital (henceforth hospital M1) and a neighboring general acute care hospital that also provides tertiary hospital care (henceforth hospital M2) vary between different hospital locations, different products

and different insurers. The merger that we study was consummated in the Netherlands in year t^3 .

Our article relates to two literatures. First, we build on the literature that structurally estimates multilateral bargaining models of healthcare competition. In general, these models contribute to our understanding of price setting mechanisms in the healthcare industry. This is relevant because standard oligopoly models are not applicable to the hospital industry (Gaynor et al., 2015). Because the current Dutch healthcare system bears evident similarities with the US healthcare system, we are able to build on the models that were developed for the US health market by Gaynor and Town (2012) (hereafter: GT) and Gowrisankaran et al. (2015) (hereafter: GNT). Following these models, we describe a bargaining model in which hospital-product prices are bilaterally negotiated between insurers and hospitals. We show how hospital-insurer negotiations translate into product prices, and by adapting the GT and GNT models for hospital mergers we show that the price effect of a merger between two hospitals may be heterogeneous depending on the degree of substitution between hospitals, the relative bargaining ability of hospitals and insurers and the pre-merger price-cost margins of different products at both merging hospitals. The most important contribution of this article to hospital-insurer bargaining models is that we, unlike GT and GNT, endogenize the product price ratio. That is, the models by GT and the GNT both assume that hospitals and health insurers bargain over a single base price per hospital, holding product-price ratios of each hospital fixed. This means that in both benchmark models each hospital entering a network always provides all treatments. Our model, in contrast, allows for the situation in which a hospital may be contracted only for a subset of treatments. In section 2, we explain that this assumption better matches current practice where contracts between hospitals and insurers can be concluded for a subset of treatments.

Second, we build on the literature on retrospective analyses of hospital mergers. Since the 1980s, hospital sectors in many OECD countries have become increasingly concentrated as a result of mergers (Gaynor and Town, 2012). Merger activity has fueled a public and scientific debate about the consequences of mergers and the desirability of further concentration of healthcare sectors. An increasing number of empirical studies have been conducted concerning the price effects of hospital mergers. Most of these studies have shown that although mergers may bring about meaningful reductions in marginal costs and therefore improve welfare overall, mergers between rival hospitals are likely to raise the price of inpatient care in concentrated markets (Gaynor and Town, 2012). We

3 For reasons of confidentiality, we only report those results that are of direct interest to this article. We anonymize the names of the merged hospitals, rival hospitals and insurers. For the same reason, the merger year is reported as t (which was between 2005 and 2012), with the year preceding the merger as $t-1$ and the year following the merger as $t+1$.

build on these studies, but disaggregate the merger price effect and show that the price effects of a merger between two hospitals may differ between locations, providers and products. With that, we contribute to a better understanding of the effects of mergers, which, in turn may also improve future antitrust policy.

This article is structured as follows. We start with the bargaining model. We then discuss the applicability of this model to the Dutch hospital market (section 3) and describe the merger that we study (section 4). The next sections concern the empirical model (section 5) and the data (section 6). In section 7 we present the results and section 8 discusses the policy implications. Finally, our main findings are summarized in section 9.

2. THE MODEL

To explain the possibility of heterogeneous price effects of hospital mergers we consider a game-theoretical model of hospital-insurer bargaining, following the lines suggested by Gaynor and Town (2012) (GT) and Gowrisankaran et al. (2015) (GNT). These papers build on earlier literature analyzing hospital-insurer bargaining, notably Gal-Or (1997); Town and Vistnes (2001); Capps et al. (2003) and Gaynor and Vogt (2003).

To keep our model as simple as possible, we adopt a two-stage set-up following the base model of GNT. In the first stage of this model, health insurers⁴ bargain and contract with hospitals on behalf of their insured and in the second stage, each consumer receives a health draw and seeks treatment at the hospital that maximizes his utility. Because the consumer commits to a restricted network of hospitals when he buys health insurance, he has the option of visiting any of the contracted hospitals when he is in need of specific care.

Like in the models by GT and GNT, we simplify some elements of the bargaining game: we condition on the network of the insurer⁵ and do not allow consumers to switch insurers in response to a network change. Following GT and GNT, the bargaining solution used in this article is based on the framework that was developed by Horn and Wolinsky (1988). While not imposing a complete non-cooperative structure, this framework nests a non-cooperative Nash equilibrium within a cooperative game theoretical concept of a Nash bargaining solution.

4 GNT use the term managed care organization or MCO if they refer to insurers that use provider networks and negotiate prices with providers. We refer to similar organizations, but use the term 'health insurer' instead as this is the more commonly used term in the Netherlands.

5 There is some work on network formation games, with Ho (2009) being the most notable. Ho (2009) estimates the parameters of managed care organization's (MCO) choices of provider network focusing on the role of different networks on downstream MCO competition (Gowrisankaran et al., 2015). Like GT and GNT, we treat the insurers' network structure as given.

To be able to explain heterogeneous price effects over products, we need to allow for flexibility in the price ratios between different products of the same hospital. Both the GT and the GNT models consider heterogeneous insurers, hospital locations and hospital products. However, they fix all the product-price ratios at the level of the respective disease-weight ratios. In their models, the hospitals are constrained to negotiate a single base price per hospital location and the prices for different products are computed as a product of the base price and the disease weight⁶. Our model deviates from this assumption by freeing the product-price ratios. While in both benchmark models each hospital that enters a network always provides all treatments, our model allows for the situation in which a hospital may be contracted only for a subset of treatments. This also better matches practice where contracts between hospitals and insurers can be concluded for a subset of treatments. For example, in the US, we observe cases in which hospitals shifted resources and activities to central profitable services, while reducing or eliminating some loss making services (i.e. the so-called specialty service lines) (Benson et al., 2006). This is in line with the anticipated strategy change towards integrated care delivery systems (Porter, 2009). Furthermore, there is an increase in the use of bundled payments, global payments or alternative quality contracts by health insurers (e.g. Chernew et al. 2011; Delbanco, 2014; Song et al. 2014). In these settings, a single payment covers the services that providers deliver to treat a given condition or provide a given treatment. Hence, in these cases, a price has to be determined for each bundle. Also in the Netherlands, which data we use when estimating the model parameters, hospitals may be contracted only for a subset of services. Interviews with health insurers and hospital representatives who were involved in contractual negotiations during our study period indicated that especially for high-revenue products insurers and hospitals bargain separate prices. In the Netherlands, it is usually the insurers that initiate selective contracting of procedures. Dutch health insurers have imposed rules on contracting certain types of operations. For example, one insurer selectively contracts providers of breast cancer surgeries (CZ, 2015), whereas another selectively contracts 15 hospital products (VGZ, 2014). As a result of selective contracting or hospitals' choices, in practice, the full hospital or a subset of procedures in a hospital may be contracted.

Model set-up

Following GT and GNT, we analyze hospital-insurer bargaining in a model with multiple hospitals and health insurers. For ease of comparison, we follow the model notation by GNT. In this model, there is a set of hospitals that is indexed by $j = 1, \dots, J$; and a set of

6 Each year, the Center for Medicare Services publishes DRG weights. The DRG weights measure the mean resource usage by diagnosis. In the model, they reflect the resource intensity of treatment. Using the DRG weights with a base price does not allow for heterogeneous price effects of mergers.

health insurance companies indexed by $m = 1, \dots, M$. Each consumer buys insurance at a particular health insurer and hence the set of enrollees for a particular health insurer is indexed by $i = 1, \dots, I$. With probability f_{id} enrollees may be stricken by illness $d \in \{0, 1, \dots, D\}$, where $d = 0$ means no illness.

In our model, we associate each illness with a hospital product⁷. Let D_j denote the list of all products of hospital j . We assume that the range of products may differ between hospitals. The set of all hospitals (each of which delivers a certain range of products) is subdivided over $S \leq J$ systems. Here J denotes the number of hospitals, and S denotes the number of hospital systems. M_s will denote the respective set of all systems. Each system $s \in M_s$ is associated with a subset in the hospital-product space of all treatment options (jd) that can be provided by this system, where index j refers to hospitals and index d to products. L_s denotes the list of treatment options (jd) with which hospital j of system s enters the hospital-insurer bargaining game. For the sake of simplicity, we consider the situation in which each system is initially represented by one hospital (i.e. $S = J$).

For any consumer i , we denote his health insurer by $m(i)$. Following the base model version of GNT, we assume that $m(i)$ is chosen via long-run employer/health insurer contracts and hence, we assume that $m(i)$ is fixed. This implies that we do not allow consumers to switch insurers in response to a network change⁸. We also treat the network of each health insurer as given. That is, we assume that each health insurer enters the negotiations with some set of hospital systems and bargains with each of these systems over the prices of products. The network of insurer m denoted by N_m defines all hospital-product pairs available to the enrollees of insurer m . By introducing the notation N_{md} for the subset of hospitals that provide product d in network N_m , we obtain the expression: $N_m = \cup_{d \in \{1, \dots, D\}} \{j \in N_{md} \mid (jd) \in N_m\}$.

Value functions of a health insurer and a hospital system

When falling ill with illness d , the patient seeks treatment at a hospital that gives him the highest utility level. The utility function from the treatment of illness d at hospitals j is given by

⁷ Please note that d can also be a cluster of products.

⁸ GNT also present a modification of their base model to include the possibility that an enrollee may choose between different health insurers. In their posted premium model extension, the framework is as follows: (i) the health insurers set their network, (ii) the health insurers post their premiums simultaneously and (iii) the enrollees choose their health insurers. The bargaining process of the posted premium model is similar to the base model, except that the threat points are different. Since the results of the base model broadly align with the extended posted premium model, we follow the relatively simpler base model.

$$u_{ijd} = \beta \mathbf{x}_{ijd} + e_{ij} \quad (1)$$

where \mathbf{x}_{ijd} is a vector of hospital and patient characteristics such as travel time, hospital quality, or other characteristics, β is the associated vector of parameters and e_{ij} is an i.i.d. error term that is distributed type 1 extreme value. We assume that getting treated at a hospital does not require an out-of-pocket payment from the patient (see below). The patient with illness d may visit any of the contracted hospitals that provide this treatment in the insurer's network or an outside option. Following GNT, we assume that the outside option is treatment at a hospital located outside the market. The outside option is denoted by $j = 0$, so that the associated characteristics are normalized: $\mathbf{x}_{i0d} = 0$.

Health insurer m provides its enrollees a set of treatment options at hospitals in its network N_m , where each option $(jd) \in N_m$ listed in the insurance policy allows patients access to hospital j for treatment of disease d . Therefore, the utility function of enrollees introduced above results in the following expression for the probability that patient i with disease d chooses hospital j :

$$s_{ijd}(N_{m(i)}) = \frac{\delta_{ijd}}{\sum_{k \in \{0, N_{m(i),d}\}} \delta_{ikd}} \quad (2)$$

where $\delta_{ijd} = \beta \mathbf{x}_{ijd}$, $j \in \{0, N_{m(i),d}\}$. The notation $N_{m(i),d}$ denotes the subset of treatment options available to individual i enrolled at insurer m for treatment of illness d . Since the right hand side of equation (2) does not depend on prices and only includes product vd , $s_{ijd}(N_{m(i)}) = s_{ijd}(N_{m(i),d})$.

It is important to note that GT and GNT differ in their position towards copayments. GT assumes that enrollees pay a premium to their insurer, which gets them access to the provider network without any additional payments, whereas GNT considers an extension in which they also model out-of-pocket payments (i.e. the negotiated base price multiplied by the coinsurance rate and the resource intensity of the illness). The GT model without copayments is in this respect similar to the GNT model with zero coinsurance rates. Because our empirical analysis focuses on the Netherlands and in the Netherlands, coinsurance as defined by GNT in the hospital sector is nonexistent⁹, we follow the approach of GT or, put differently, the approach of GNT with zero coinsurance rates. For our model this means that the utility from treatment does not depend on hospital prices and hence the resulting choice probabilities are also independent of product prices.

⁹ Also copayments are very limited. There is a yearly mandatory deductible that the patient pays when he starts using healthcare. However, the deductible is limited to a fixed amount. Since most hospital prices are higher than this amount, each patient receiving treatment at any hospital would generally pay the same deductible. Hence, deductibles are expected to hardly affect patient hospital choice.

The *ex ante* expected utility to patient i from network $N_{m(i)}$ is then:

$$w_i(N_{m(i)}) = \sum_{d=1}^D f_{id} \ln(\sum_{j \in \{0, N_{m(i),d}\}} \exp(\delta_{ijd})) \quad (3)$$

Aggregating over the enrollees of insurer m , we obtain:

$$W_m(N_m) = \sum_{i=1}^I 1\{m(i) = m\} w_i(N_m)$$

Denoting the prices that insurer m pays to hospital j for treatment d by $p_{mj d}$, we obtain the insurer's total cost as follows:

$$TC_m(N_m, \mathbf{p}_m) = \sum_{i=1}^I \sum_{d=1}^D 1\{m(i) = m\} f_{id} \sum_{j \in \{0, N_{m(i),d}\}} p_{mj d} s_{ij d}(N_m) \quad (4)$$

Following GNT, we assume that the health insurer is seeking to maximize the sum of the enrollee surplus (equal to $w_i - \text{premium}_m$ for each consumer) and the insurer's profit (equal to $\text{premium}_m - \text{expected cost}_m(i)$ for each consumer) over all enrollees. Under this assumption, the value function of the health insurer is the difference between the *ex ante* expected utility of all the enrollees and the total payment to the hospitals treating these enrollees:

$$V_m(N_m, \mathbf{p}_m) = W_m(N_m) - TC_m(N_m, \mathbf{p}_m) \quad (5)$$

Note that in GNT the health insurer acts as an agent for the employer and, thus, cares equally about both enrollee welfare and insurer profit¹⁰. With that, it is assumed that the incentives of health insurers and enrollees are perfectly aligned which implies that both terms in equation (5) will have equal weights¹¹.

Substituting into this expression equations (3) and (4), and rearranging the terms, we derive the same expression in terms of prices and choice probabilities. Since both expected utility and the payment to the hospital are separable in products d , the total value function of a health insurer has an additive structure over the products. This can be seen as follows:

10 This is also a reasonable assumption in the Netherlands, where the provision of basic insurance is subject to strict rules so that Dutch health insurers too not only care about profit maximization, but also enrollee welfare. Having originated from social insurance funds, some insurers even explicitly state that they continue to carry out social mission. See section 3 for more details.

11 If we assume stronger power on the enrollee or the health insurer side, we would have to impose a higher weight to the respective term (as discussed in Gowrisankaran et al., 2015 and Gaynor et al., 2015).

$$\begin{aligned}
V_m(N_m, \mathbf{p}_m) &= W_m(N_m) - TC_m(N_m, \mathbf{p}_m) \\
&= \sum_i 1\{m(i) \\
&= m\} \sum_d f_{id} \left(\ln \left[\sum_{j \in \{0, N_{md}\}} \exp(\delta_{ijd}) \right] \right) - \sum_{j \in \{0, N_{md}\}} p_{mjd} s_{ijd}(N_m) \\
&= \sum_d \sum_i 1\{m(i) = m\} f_{id} \left(\ln \left[\sum_{j \in \{0, N_{md}\}} \exp(\delta_{ijd}) \right] \right) \\
&\quad - \sum_{j \in \{0, N_{md}\}} p_{mjd} s_{ijd}(N_m) \\
&= \sum_d W_{md}(N_{md}) - TC_{md}(N_{md}, \mathbf{p}_{md}) = \sum_d V_{md}(N_{md}, \mathbf{p}_{md})
\end{aligned}$$

where \mathbf{p}_m is the price vector of all product prices negotiated by insurer m , \mathbf{p}_{md} denotes the subvector of product d 's prices, N_{md} is the subset of options for product d , $W_{md}(N_{md}) = \sum_i 1\{m(i) = m\} f_{id} \ln \left[\sum_{j \in \{0, N_{md}\}} \exp(\delta_{ijd}) \right]$ and $TC_{md}(N_{md}, \mathbf{p}_{md}) = \sum_d \sum_i 1\{m(i) = m\} f_{id} \sum_{j \in \{0, N_{md}\}} p_{mjd} s_{ijd}(N_m)$. Since the choice probabilities do not depend on product prices, the enrollee surplus from each product also does not depend on prices of other products.

Following GT and GNT, we assume profit maximizing hospitals, which is typical in the health economics literature, especially because numerous studies found that the behavior of for-profit and not-for-profit hospitals is similar¹². The marginal cost of providing product d in hospital j for health insurer m can then be denoted by mc_{mjd} :

$$mc_{mjd} = \gamma v_{mjd} + \epsilon_{mjd} \quad (6)$$

where v_{mjd} denotes a fixed effect, γ is the associated parameter and ϵ_{mjd} is an error term. Because we assume that hospitals are maximizing their profits, we let each hospital system s maximize the total profits earned from the contracts with health insurers:

$$\pi(M_s, N_m, \mathbf{p}_m) = \sum_{m \in M_s} \sum_{(j,d) \in L_s} (p_{mjd} - mc_{mjd}) q_{mjd}(N_m) \quad (7)$$

where q_{mjd} denotes the production volumes of the hospitals under hospital-product system s and mc_{mjd} is the marginal cost of treatment d at hospital j for enrollees of insurer

¹² In this article, we assume that hospitals are profit maximizers, but Lakdawalla and Philipson (2006) and Gaynor et al. (2015) have shown that output maximization can be incorporated in the standard hospital utility function in addition to profit maximization by using perceived marginal costs instead of actual marginal costs.

m .¹³ Because of our assumption on the consumer utility function, the volume delivered by the hospital system only depends on the set of treatment options included in the network and not on the prices of these options. The production quantities of hospital j are then expressed by:

$$q_{mjd}(N_m) = \sum_i 1\{m(i) = m\} f_{id} s_{ijd}(N_m) \quad (8)$$

Bargaining problem

There are $M \times S$ potential contracts. However, in our model, each contract specifies the prices of treatment options that are contracted by the insurer and the hospital system, and not the base prices of the hospitals that enter the system, as in the models by GT and GNT. Following GT and GNT, we assume that bargaining occurs under complete information about the characteristics of enrollees and hospitals and we consider the Nash Bargaining solution price vector that results from the maximization of the product of the exponentiated value functions of both parties from agreement, conditional on all other prices. Based on the theoretical contributions by Binmore et al. (1986), Horn and Wolinsky (1988) and Collard-Wexler et al. (2014), it is assumed that the prices of each contract are negotiated conditional on the prices of all other contracts and that the agents do not change their strategies when they observe the outcome of the contracts that have already been concluded. That is, if one negotiating pair fails, the other pairs will continue the negotiation process conditional on their initial assumptions regarding the pricing outcomes of the other pairs ('passive beliefs'). The introduction of these assumptions corresponds with the models that were developed in the recent literature on hospital-insurer negotiations (in particular, GT and GNT). Here, we additionally assume that both insurers and hospitals appoint their negotiating teams per product. Therefore, bargaining on one product occurs separately from other products.

Under these assumptions, the objective of the Nash bargaining problem of health insurer m and system s is as follows:

$$\begin{aligned} & NB^{m,s}(\mathbf{p}_{m,s} | \mathbf{p}_{m,-s}) \\ &= \left(\sum_d \left[\sum_{\{jd\} \in L_s} q_{mjd}(N_m) (p_{mjd} - mc_{mjd}) \right] \right)^{b_{ism}} \\ &\times \left(\sum_d [V_m(N_m, \mathbf{p}_m) - V_m(N_m \setminus L_s, \mathbf{p}_m)] \right)^{b_{msd}} \end{aligned}$$

13 Marginal costs may differ between insurers, for example because of differences in administrative costs. If we assume, however, that marginal costs are the same over insurers, we could drop index m from the notation of marginal costs.

where $b_{s(m)}$ and $b_{m(s)}$ are the bargaining weights of system s and health insurer m respectively. The weights characterize the bargaining abilities of both negotiating parties. They are normalized to sum up to one. $\mathbf{p}_{m,s}$ and $\mathbf{p}_{m,-s}$ denote the insurer's prices of the treatment options at hospitals that participate in hospital system s and those that do not participate in the system, respectively.

The Nash equilibrium is a vector of prices that maximizes the Nash bargaining value specified above. Each price vector maximizes the value for the negotiating pair, conditional on the other prices:

$$p_{mjd}^* = \underset{p_{mjd}}{\operatorname{argmax}} NB^{m,s}(p_{mjd}, \mathbf{p}_{m,-(jd)}^* | \mathbf{p}_{m,-s}^*) \quad (9)$$

The new notation $\mathbf{p}_{m,-(jd)}^*$ denotes the equilibrium price vector consisting of all negotiated prices between insurer m and system s except for p_{mjd} .

Although each team negotiates separately, different negotiating teams of the same agent would generally take into account the effect of their decisions on patient flows for other products of the same agent. However, as according to equation (2) patient flows are fully determined by the network structure (i.e., the set of treatment options) and not by prices, the decisions of different product teams of the same agent will not be dependent on each other. This can be seen as follows. Consider that hospital j negotiates with insurer m over the price of product d , conditional on the other prices. We partition the set of all diseases into $\{D', d, D''\} = \{d_1, \dots, d_D\}$, where $\{D', d\}$ covers the subset of products with which hospital j enters the network of insurer m and D'' covers the rest. Because $m(i)$ is fixed, a hospital system that fails to reach agreement with a particular insurer regarding treatment option (jd) cannot capture any profit on this treatment option from the enrollees of this health insurer. Therefore, the disagreement outcome of the hospital system in negotiation over this treatment option will be zero. The payoff structure in bargaining between insurer m and hospital j over (jd) will then be:

$$\begin{aligned} j_{agree}^d &= \pi_{jd}(N_{md}, \mathbf{p}_{md}) + \pi_{j,D'}(N_{mD'}, \mathbf{p}_{mD'}) \\ j_{disagree}^d &= \pi_{jD''}(N_{mD''}, \mathbf{p}_{mD''}) \\ m_{agree}^d &= V_{md}(N_{md}, \mathbf{p}_{md}) + V_{mD'}(N_{mD'}, \mathbf{p}_{mD'}) + V_{mD''}(N_{mD''}, \mathbf{p}_{mD''}) \\ m_{disagree}^d &= V_{md}(N_{md}, \mathbf{p}_{md}) + V_{mD'}(N_{mD'}, \mathbf{p}_{mD'}) + V_{mD''}(N_{mD''}, \mathbf{p}_{mD''}) \end{aligned}$$

This payoff structure implies that the difference between the agreement and disagreement payoffs in negotiations on any product d only depends on the part related to that particular product. In particular, $j_{agree}^d - j_{disagree}^d = \pi_{jd}(N_{md}, \mathbf{p}_{md})$ and $m_{agree}^d - m_{disagree}^d = V_{md}(N_{md}, \mathbf{p}_{md}) - V_{md}(N_{md}, \mathbf{p}_{md})$. Hence, only these terms will be relevant for the derivation of the price p_{mjd} . Note that bargaining over this price only occurs if the sum of the payoffs

is positive: $j_{agree}^d - j_{disagree}^d + m_{agree}^d - m_{disagree}^d > 0$, therefore each 'link' (jd) included in the network must satisfy:

$$\pi_{jd}(N_{md}, \mathbf{p}_{md}) + V_{md}(N_{md}, \mathbf{p}_{md}) - V_{md}(N_{md} \setminus j, \mathbf{p}_{md}) = W_{md}(N_{md}) - W_{md}(N_{md} \setminus j) - mc_{mjd}$$

$$q_{mjd}(N_{md}) - \sum_{l \neq j, l \in \{0, N_{md}\}} p_{mjl}(N_{md})(q_{mld}(N_{md} \setminus j) - q_{mld}(N_{md})) > 0$$

Hence, hospital j will produce product d only if this condition is satisfied. The payoff structure outlined above leads to the following Nash bargaining problem with respect to p_{mjd} :

$$\max_{p_{mjd} | N_{md}, \mathbf{p}_{m,-j,d}} (j_{agree}^d - j_{disagree}^d)^{b_{s(m)}} (m_{agree}^d - m_{disagree}^d)^{b_{m(s)}}$$

where $\mathbf{p}_{m,-j,d}$ corresponds to the price vector of contract prices of hospitals other than j in the subset of treatments options N_{md} . The same type of Nash bargaining problem as described above is considered in GNT and GT, with the difference that their problem is formulated for a hospital's base price, keeping a product weight fixed in accordance to the disease weights of different diagnoses.

From the first order condition (FOC) of this problem, we derive the expression for product prices:

$$p_{mjd} = b_{s(m)} \frac{W_{md}(N_{md}) - W_{md}(N_{md} \setminus j)}{q_{mjd}} + b_{m(s)} mc_{mjd} + b_{s(m)} \sum_{k \neq j} [p_{mkd} d_{md}^{jk}]$$

where $d_{md}^{jk} = \frac{q_{mkd}(N_{md} \setminus j) - q_{mkd}(N_{md})}{q_{mjd}}$. The numerator of this ratio shows how many patients of insurer m with illness d will flow to hospital k if hospital j no longer treats this illness, and therefore d_{md}^{jk} defines the disease-specific diversion share of patients with illness d from hospital j to hospital k . A higher value of the diversion share suggests a higher degree of substitution between two hospitals in treating this illness.

The expression for p_{mjd} suggests that a product price of a hospital is increasing in the hospital's marginal costs of this product, the product prices of other hospitals, and net value that the inclusion of treatment option (jd) brings to the insurer's network. In addition to these factors, negotiated prices also depend on the bargaining abilities/weights of the hospital and the insurer. Differences in these parameters can explain the presence of price differences between health insurers, hospital locations and hospital products.

Merger analysis

The merger analysis considered in our article adopts a method proposed by GT. The method by GT allows us to derive the expressions of product price changes in a closed form, which simplifies the price comparison across products and players.

GT consider two alternative approaches to model a hospital merger of hospitals j and k . In the first approach, it is assumed that after the merger, these hospitals still negotiate prices per hospital, but take into account the impact of disagreement on the flow of patients to each other. In the second approach it is assumed that hospitals negotiate jointly and will charge the same price after the merger. Because our empirical application deals with the situation in which hospitals continue to charge different prices after they merged, we follow the first approach. Please note that because in our model the patient flows of different products are independent of each other, the problem can be split and analyzed separately for each product.

Drawing from GT, we analyze the situation in which two hospitals that enter the same network are merging and consider the bargaining problem for product d after their merger has taken place (assuming that the network covers treatment options of d at both hospitals). If each of the merged hospitals negotiates its own price of the product, but accounts for the effect on the other's patient flow, we obtain the following expressions for the agreement and disagreement payoffs in the bargaining problem of hospital j :

$$\begin{aligned}(j+k)^d_{agree} &= [p_{mjd} - mc_{mjd}] q_{mjd}(N_{md}) + [p_{mkd} - mc_{mkd}] q_{mkd}(N_{md}) \\(j+k)^d_{disagree} &= [p_{mkd} - mc_{mkd}] q_{mkd}(N_{md} \setminus j) \\m^d_{agree} &= W_{md}(N_{md}) - p_{mjd} q_{mjd}(N_{md}) - \sum_{l \neq j} p_{mld} q_{mld}(N_{md}) \\m^d_{disagree} &= W_{md}(N_{md} \setminus j) - \sum_{l \neq j} p_{mld} q_{mld}(N_{md} \setminus j)\end{aligned}$$

Writing down the Nash bargaining solution for this game and transforming the FOC of this problem, we derive the price of hospital j 's product d after the merger, $p^{(j+k)}_{mjd}$, as follows:

$$p^{(j+k)}_{mjd} = b_{s(m)} \frac{W_{md}(N_{md}) - W_{md}(N_{md} \setminus j)}{q_{mjd}} + b_{m(s)} mc_{mjd} + \frac{p_{mkd} d^{jk}_{md}}{q_{mjd}} + b_{s(m)} \sum_{l \neq j} [p_{mld} d^{jl}_{md}]$$

If we then take the difference between this price and the initial price level of hospital j , we obtain the expression for price change due to merger (given that the marginal costs are not affected by the merger):

$$p^{(j+k)}_{mjd} - p_{mjd} = b_{m(s)} (p_{mkd} - mc_{mkd}) d^{jk}_{md} \quad (10)$$

The same type of derivations can be done for hospital k , with indices k and j changing places.

Heterogeneous price effects of hospital mergers

There are a few important conclusions that can be drawn from equation (10) with respect to the price effect of a hospital merger. The first important finding is that product

d 's price change after the merger in each hospital is increasing in the diversion share between these hospitals. Since the diversion share reflects the degree of substitution between the hospitals, this result tells us that a merger will increase the product's price more if the hospitals that partner in the merger are close substitutes with respect to that product. Therefore, if substitution between hospitals is stronger for one product than for another product¹⁴, the price increase after the merger will be higher for the first product and hence hospital mergers may lead to heterogeneous price effects across different products and different locations.

The second most important conclusion that follows from our model is that, according to equation (10), the price change caused by merger is proportional to the difference between the price and the marginal cost of the other hospital (i.e. the merger partner). Therefore, these differences also contribute to explaining the heterogeneity of price changes after the merger for different products and locations. Merging with a hospital whose price of product d is higher, whereas the marginal cost are lower, would result in a greater price increase (other things being equal).

Finally, we observe, perhaps at first sight somewhat contra-intuitively, that a price increase caused by merger is proportional to the bargaining ability $b_{m(s)}$ of the insurer. Thus, a health insurer with greater bargaining ability against hospital system s is confronted with a higher price increase after the merger. This result suggests that, although a greater relative bargaining ability of the insurer in comparison to hospitals provides the insurer with more leverage against these hospitals, this leverage advantage is reduced after the merger of the hospitals.

3. THE DUTCH HOSPITAL MARKET

In this article, we estimate the price changes of a merger between two Dutch hospitals. From the viewpoint of the bilateral bargaining model, the current Dutch healthcare system bears important similarities with the US healthcare system. In recent decades, the Netherlands, like several other OECD countries, has embraced a market-oriented approach to healthcare. After decades of strict governmental supply-side regulation, the Dutch healthcare system is currently undergoing a transition towards regulated (or 'managed') competition (Van de Ven & Schut, 2009; 2008; Schut & Van de Ven, 2005). The main goal of the market-oriented healthcare reforms is to increase the efficiency of the

14 The substitution rates may differ across products, for example, because for some hospital products patients' willingness to travel might be higher, there is more intense competition with nearby hospitals over those products or the transparency of different product markets differs.

system and its responsiveness to patients' needs, whereas maintaining universal access to care (Schut & Van de Ven, 2005).

Of particular importance to this article are the introduction of the Health Insurance Act (HIA) in 2006 and the introduction of hospital-insurer bargaining in 2005. Under the HIA, all Dutch citizens are obliged to buy standardized individual basic health insurance from a private insurer. The standardized basic benefits package specified in the HIA is fairly comprehensive and includes hospital care, GP services, prescription drugs and maternity care. Having bought an insurance policy, the enrollee gets access to all hospitals of the contracted network without co-payments. As described in section 2, there is an annual deductible per adult individual, although most hospital product prices are higher than the fixed amount that is set by the deductible¹⁵ and hence the deductible does not play a role in patients' hospital choices. Dutch health insurers are furthermore required to offer all applicants standardized coverage at a community-rated premium, the insurers have to offer all basic health insurance policies to all applicants (i.e. a guaranteed issue requirement) and consumers are free to choose their health insurer during an annual enrolment period. Risk equalization across insurers takes place to ensure a level playing field for health insurers and to prevent risk selection. The insurers' market shares are relatively stable¹⁶.

In 2005, a product classification system for hospital and medical specialist care was introduced. Each activity and/or service provided by a hospital, including outpatient care, which is associated with a patient's demand for care, is referred to as a Diagnosis and Treatment Combination (DTC)¹⁷. Following the introduction of the DTC system, the scope for free negotiations of prices between hospitals and health insurance companies has gradually increased from 10% of hospital revenue in 2005, to 20% in 2008, to 34% in 2009 and to 70% in 2012. For the remaining part, hospital prices are still regulated. For products and services included in the free-pricing segment, each hospital typically renegotiates the terms of its contracts with health insurers on an annual basis. Dutch health insurers are allowed to engage in selective contracting with healthcare providers. As explained in section 2, there are several cases in which the insurer contracts only a subset of treatments in hospitals.

15 Just 11% of all patients received treatments that cost less than 165 euro in 2011. The prices of the products that we consider in our article all exceed the deductible during the study period.

16 For example, the switching rate between health insurance companies in the Netherlands was 6% in 2012.

17 The DTC system is based on the concept of Diagnosis-Related Groups but constitutes a newly developed classification system. The Dutch system originally contained 29,000 DTCs. In 2007, a project was initiated to decrease the number of DTCs to about 3,000. This was known as the 'DOT revision' and was implemented in January 2012.

The two-stage model that underlies the bargaining theory developed above reflects how Dutch health insurers and hospitals negotiate over the products in the free-pricing segment: consumers buy health insurance from health insurers and health insurers bargain and contract with hospitals on behalf of those that they insure. In the early years of the reform selective contracting was limitedly used, but over the years, the number of health insurers offering contracts with restricted provider networks has increased. Furthermore, the available evidence on the nature of hospital-insurer negotiations in the Netherlands suggests that until 2012, hospital-insurer bargaining focused on price, rather than on quality of volume of care (Ruwaard et al., 2014; Meijer et al., 2010; NZa, 2009). The introduction of the HIA has led to strong price competition between health insurers and health insurers have put increasing pressure on hospitals to charge lower prices (Schut & Van de Ven, 2011). It seems as if the threat of selective contracting, rather than its actual use, may already have had an impact on hospital-insurer bargaining.

4. THE MERGER

Dutch local and regional hospital markets are highly concentrated¹⁸ and mergers represent the largest change in the Dutch hospital industry nowadays as no hospitals have entered or exited the market since 2005. Between 2005 and 2012, 17 mergers involving 34 hospitals were cleared by the Authority for Consumers and Markets (ACM)¹⁹ (www.acm.nl), among which the merger that we study in this article. All mergers took place between neighboring hospitals.

The merger that we study was consummated in year t (which was between 2005 and 2012). The merger was notified to the ACM prior to taking place²⁰. Following the notification, the ACM carried out a general review of the proposed merger in which they made prospective inferences regarding the expected anticompetitive effects of the

18 In 2006, the average HHI of Dutch hospitals equaled 2.350 (Halbersma et al., 2010) and since then no hospitals entered or exited the hospital market. Only mergers have decreased the number of hospitals.

19 The Authority for Consumers and Markets is the Dutch antitrust agency. The legal predecessor of the Authority for Consumers and Markets, the Netherlands Competition Authority, has carried out the review of some of these mergers. For reasons of clarity, however, we ascribe the decisions made by the Netherlands Competition Authority to its legal successor, which has been in charge since April 1, 2013: the Authority for Consumers and Markets.

20 According to most antitrust laws, mergers must be reported to an antitrust authority prior to consummation (see 15 USC §18A for the US and the competition laws of the EU Member States or EC: 2004 for the European Union's rules on prior merger notification). The Dutch antitrust law is no exception (Mededingingswet, section 37.2).

merger on the market. In the Netherlands, a merger requires a license when there is reason to assume that 'a dominant position that appreciably restricts competition on the Dutch market or a part thereof could arise or be strengthened as a result of the said concentration' (Mededingingswet, Section 37.2). The merger that we study did not require a license and was cleared after the first general review. The decision to clear the merger evoked critical appraisal by health economists, however, who argued that the prospective merger analysis by the antitrust authority had been lacking and that it was likely that the merger had created a dominant position for the two hospitals involved (Varkevisser & Schut, 2008). Hence, this merger seems to be on the enforcement margin, making it an interesting case for further retrospective studies.

The locations

The merger involved a general acute care hospital (hospital M1) and a neighboring general acute care hospital that also provides tertiary hospital care (hospital M2). Hospital M1 is located in an isolated geographical area, whereas hospital M2 is located in a more densely populated region with several other hospitals nearby. The distance between hospitals M1 and M2 is about 50 kilometers²¹. According to the ACM, the merging hospitals were subject to competition from five other hospitals before the merger took place. Prior to the merger, hospital M2 was the largest competitor to hospital M1 and therefore posed a major constraint on hospital M1's prices, whereas hospital M2 had multiple competitors. After the merger, hospital M1 was expected to experience competitive pressure from only one rival hospital, whereas hospital M2 was expected to experience notable competitive pressure from five other hospitals²². The differences in competitive pressure in the markets of hospitals M1 and M2 may result in heterogeneous price effects of the merger (see section 2). To find out whether the merging hospitals exploited this opportunity, we disaggregated the merger effect for each of the two merging hospital locations.

The products

In this article, we estimated the impact of the merger in three separate product markets that jointly make up 47.5 percent of the merged hospital's turnover in the segment for which Dutch insurers and hospitals were allowed to freely negotiate prices at the time of the merger. We looked at hip replacements, knee replacements and cataract surgery. Most hospitals provide these services. In year t , 95% of all Dutch hospitals ($n=97$) and

21 1 kilometer is approximately 0.621 miles

22 None of these rivals provides tertiary hospital care.

Table 1. Descriptive statistics

	Hip replacements		Knee replacements		Cataract surgery	
	<i>t</i> - 1	<i>t</i> + 1	<i>t</i> - 1	<i>t</i> + 1	<i>t</i> - 1	<i>t</i> + 1
<i>Panel A. Hospital M1</i>						
Volume	174	175	223	293	387	361
Gender (% male)	0.28	0.38	0.34	0.43	0.38	0.35
Patients' average age	68	68	64	56	72	73
Patients' average SES score	0.05	-0.14	0.15	0	0.09	-0.06
<i>Panel B. Hospital M2</i>						
Volume	390	511	271	299	2144	2113
Gender (% male)	0.34	0.35	0.34	0.32	0.41	0.40
Patients' average age	68	70	69	69	72	73
Patients' average SES score	0.31	0.42	0.39	0.48	0.35	0.42
<i>Panel C. Rival 1</i>						
Volume	165	154	164	135	1026	1045
Gender (% male)	0.27	0.36	0.27	0.29	0.41	0.37
Patients' average age	70	71	71	69	75	75
Patients' average SES score	-0.22	-0.05	-0.06	-0.09	-0.09	-0.02
<i>Panel D. Rival 2</i>						
Volume	237	195	162	162	881	1088
Gender (% male)	0.32	0.34	0.32	0.38	0.43	0.41
Patients' average age	70	68	68	68	73	72
Patients' average SES score	0.15	0.12	0.15	0.28	0.22	0.26
<i>Panel E. Rival 3</i>						
Volume	136	114	146	118	650	972
Gender (% male)	0.34	0.28	0.40	0.29	0.38	0.42
Patients' average age	70	62	70	70	75	74
Patients' average SES score	-0.83	-0.88	-0.76	-0.69	-1.01	-0.96
<i>Panel F. Rival 4</i>						
Volume	169	155	101	151	855	763
Gender (% male)	0.34	0.26	0.38	0.35	0.43	0.44
Patients' average age	69	73	70	71	75	75
Patients' average SES score	0.24	0.46	0.09	0.36	0.17	0.4
<i>Panel G. Other hospitals</i>						
Volume	231 (14)	234 (15)	196 (12)	199 (12)	1590 (146)	1545 (137)
Gender (% male)	0.33	0.34	0.32	0.33	0.39	0.41
Patients' average age	69 (0.37)	69 (0.25)	69 (0.27)	69 (0.26)	73 (0.32)	73 (0.29)
Patients' average SES score	-0.04 (0.05)	-0.18 (0.08)	0 (0.05)	-0.11 (0.07)	0.01 (0.05)	-0.09 (0.07)

Notes: The standard errors are in parentheses. We excluded all hospitals that had more than 15% missing prices for either hip or knee replacements or cataract surgeries in the period *t*-2 to *t*+2. The fifth rival hospital was therefore excluded from this analysis. Panel G displays the descriptive statistics of the hospitals other than hospitals M1, M2 and the rival hospitals. Within panel G, 51 hospitals performed hip replacements, 56 hospitals performed knee replacements and 57 hospitals performed cataract surgeries. The rows on volume only report cases which have a valid gender, age and SES-score.

2.7% of all Dutch Independent Treatment Centers (ITCs)²³ ($n=73$) provided hip replacements, 95% (hospitals) and 7% (ITCs) provided knee replacements and 96% (hospitals) and 15% (ITCs) provided cataract surgery. These products were also provided by hospitals M1 and M2 and all five rivals in year t . At time of the merger, there were no ITCs in the regional market that offered any of the hospital products considered. Table 1 presents descriptive statistics on the patients for each product in hospitals M1 and M2 and four rivals²⁴ before and after the merger.

After merger, the hospitals had an opportunity to concentrate care in one of the two hospital locations. This does not seem to have occurred, however. Even though it follows from table 1 that hospital M2 provided many more hip replacements in year $t+1$ than in $t-1$, the provision of hip replacements in hospital M1 did not change significantly. The hospitals therefore do not seem to have concentrated care in hospital M2 after the merger. Rather, it seems that hospital M2 is, post-merger, better able to attract patients in need of hip replacements because the number of hip replacements performed in rival hospitals decreased slightly whereas the total number of patients in the market did not change significantly.

In hospital M1, the average age of patients undergoing knee replacements dropped between $t-1$ and $t+1$. Again, this does not seem to be an attempt to change patient flows in the merged hospitals, as the mean age of patients undergoing knee replacement surgery in hospital M2 did not change. However, according to hospital M1's website, the hospital has been testing out an innovative procedure for knee replacements since year t for which only patients under 60 years old are eligible. This is likely unrelated to the merger, but could potentially explain the decrease in the patients' average age observed in the data.

23 ITCs are comparable to freestanding Ambulatory Surgery Centers (ASCs) that operate in the US and UK healthcare markets (see e.g. Gaynor & Town, 2012; Carey et al., 2011). Independent Treatment Centers (ITCs) are typically much smaller than general hospitals and only compete on a narrow range of specialties. Their market share is relatively small, but their influence has increased because they usually offer elective care treatments, focus on the free-pricing segment and have rapidly grown in number and size (NZa, 2012; 2009). The joint market share of all ITCs increased from 1.5 percent (2005) to 4 percent (2007) of the free-pricing segment's total returns (NZa, 2009) and from 1 percent (2007) to 2.3 percent (2010) of total medical specialist care (NZa, 2012).

24 We excluded all hospitals that had more than 15% missing prices for either hip or knee replacements or cataract surgeries in the period $t-2$ to $t+2$. The fifth rival hospital was therefore excluded from the analysis. See section 5 for more information on the exclusion criteria.

The health insurers

At the time of the merger, at least five health insurers were active in the region²⁵. Four of these were independent health insurers, whereas the fifth was in fact a joint purchasing organization representing the majority of smaller health insurers. For reasons of clarity, we will henceforth treat this purchasing entity as a health insurer. All five health insurers are active on the national insurance market. According to table 1, the volume of patients has not changed significantly across hospitals, indicating that health insurers did not shift enrollees away from the merged hospitals to rival hospitals in $t+1$.

Table 2 shows the insurers' market share for each product and for each hospital in years $t-1$ and $t+1$. The market shares have not changed significantly over the years. Although insurer 1 has the largest market share per product per hospital (its market share ranges from 61% to 84%) it is not the largest health insurer nationally²⁶. Regional market shares reflect the continuing effect of the former regional legal monopoly positions of local health insurers (a policy that was abolished in 1992) (Halbersma et al., 2010).

Table 2. Health insurers' market share per product per hospital in $t-1$ and $t+1$

	Market share insurer 1		Market share insurer 2		Market share insurer 3		Market share insurer 4		Market share insurer 5	
	$t-1$	$t+1$	$t-1$	$t+1$	$t-1$	$t+1$	$t-1$	$t+1$	$t-1$	$t+1$
<i>Panel A. Hospital M1</i>										
Hip replacements	0.76	0.74	0.05	0.04	0.09	0.13	0.05	0.02	0.05	0.07
Knee replacements	0.69	0.61	0.05	0.06	0.16	0.20	0.06	0.08	0.05	0.06
Cataract surgery	0.84	0.77	0.01	0.03	0.09	0.09	0.04	0.05	0.02	0.06
<i>Panel B. Hospital M2</i>										
Hip replacements	0.62	0.62	0.08	0.06	0.19	0.17	0.04	0.06	0.07	0.08
Knee replacements	0.69	0.62	0.04	0.03	0.17	0.20	0.01	0.06	0.09	0.08
Cataract surgery	0.70	0.71	0.04	0.05	0.16	0.14	0.02	0.03	0.08	0.08

Notes: The health insurers' market shares are based on the number of cases per hospital-insurer-product combination.

25 In fact, there are six health insurers active in the region. However, for the sixth health insurer, we did not have valid prices in the post-merger year ($t+1$) for the merging hospitals M1 and M2. This health insurer was therefore not included in the difference-in-differences estimates or in any other analysis. The effect of excluding this health insurer for hospital M1 and hospital M2 is most likely negligible, however, because the health insurer only accounts for less than 2% of all hip, knee and cataract patients in hospitals M1 and M2.

26 For reasons of confidentiality, we cannot report the national market shares of the health insurers.

5. EMPIRICAL MODEL SPECIFICATION

We use data on hospital-insurer negotiated contract prices in the Netherlands for each of the three hospital products considered, to investigate whether the merger between hospitals M1 and M2 has led to price changes and if so, whether this effect varies between locations, payers and products. There are several ways to calculate price changes post-merger. The first method is to calculate the post-merger price change for each hospital product indexed on, for example, the average price change over all hospitals. However, these price changes would only give us a crude indication of the effect of the merger as it does not take account of changes in prices that would also have occurred if the merger had not taken place.

Although our model focuses on the price effects that follow from the interaction between health insurers and hospitals, large post-merger price increases for merged hospitals in comparison to prices among a control group could be consistent with at least four hypotheses according to the empirical literature (Haas-Wilson & Garmon, 2011; Adams & Noether, 2011): (i) the merger created or enhanced the hospital's power to raise its prices for general acute inpatient services; (ii) between the years $t-1$ and $t+1$ there was an increase in the product complexity of inpatient cases or an increase in the severity of patients' illness in the merging hospitals relative to non-merging hospitals; (iii) between the years $t-1$ and $t+1$, the quality of care associated with the products improved at the merging hospitals relative to non-merging hospitals, which increased value and (perhaps) cost and (iv) pre-merger prices at the merging hospitals were lower than the competitive equilibrium prices. In other words, the post-merger price increases at the merged hospital could be an adjustment towards equilibrium (Garmon & Haas-Wilson, 2011). We call this latter phenomenon 'catching up'. When interpreting our results in section 8, we will also reflect on these alternative explanations, arguing that the first explanation is the most likely in our case.

Because we wanted to control for price changes that would have occurred even if the merger had not taken place, we used a difference-in-differences (DID) model in which price changes at the merging hospitals are compared to price changes among a group of comparison hospitals which were unaffected by the merger (i.e. the control group). The identifying assumption of a difference-in-differences estimation is that trends (price trends) would be the same in both groups in the absence of the event (merger). This assumption is referred to as the 'common trend assumption'. We visually investigated whether the common trend assumption applies by using data on multiple periods.

To examine the effect of aggregating the merger price effect, we estimated difference-in-differences models at various aggregation levels. As a benchmark, we started with the most aggregated model. In other words, we first estimated the price effect for the merged hospital fully aggregated over hospital locations, products and insurers. We

then disaggregated this effect stepwise to ultimately arrive at the most differentiated model in which we fully differentiated the merger price effect across hospital locations, products and insurers. Table 3 provides a summary of the different models.

Table 3. Continuum of aggregated and disaggregated models

Models	Merger price effect
Baseline model	Fully aggregated over hospital locations, products and insurers
First disaggregated model	Aggregated over hospital products and insurers; disaggregated across locations
Second disaggregated model	Aggregated over hospital locations and insurers; disaggregated across products
Third disaggregated model	Aggregated over insurers; disaggregated across products and locations
Fourth disaggregated model	Aggregated over hospital locations and products; disaggregated across insurers
Fifth disaggregated model	Aggregated over hospital products; disaggregated across insurers and locations
Disaggregated model	Fully disaggregated across hospital locations, products and insurers

We first estimated the most aggregated model:

$$(11) \quad \ln p_{jt} = \alpha + \lambda \cdot POST_t + \delta \cdot POST_t \cdot MERGED_j + \vartheta_j + \varepsilon_{jt}$$

where p_{jt} was the weighted average hospital negotiated price. $POST_t$ is one in year $t+1$ (the post-merger year) and zero in year $t-1$ (the pre-merger year), $MERGED_j$ is one for the merger hospitals and zero for the control group hospitals, $\lambda \cdot POST_t$ denotes the change in the average price in year $t+1$ compared to year $t-1$, δ is the DID estimator (i.e. the average treatment effect on the treated; see Blundell & Costa Dias, 2009) and ϑ_j is a hospital fixed effect. To account for potential endogeneity of the merging policy, we matched a control group to the event group (i.e. hospitals M1 and M2). In this control group, we included all Dutch hospitals that provided the three products and excluded any other hospitals that also merged between years $t-2$ and $t+2$ and Independent Treatment Centers.

To estimate the most aggregated difference-in-differences model we aggregated the patient-level hospital data to an average price per hospital. It is important to note that in the Netherlands, negotiated prices differ between health insurers but not between patients with the same health insurer who are treated in the same hospital. Therefore, we can aggregate the data to hospital-insurer level without a loss of information. Furthermore, due to aggregation, we do not have to consider the correlation between prices within each hospital-insurer combination, which would otherwise lead to biased standard errors (see for example Thompson, 2011; Donald & Lang, 2007 and Bertrand et al., 2004). First, we calculated an average price per product for each hospital-insurer pair. Second, we aggregated these prices over the insurers to an average price for each hospital-product combination, whereby we weighted the prices with the insurer's spe-

cific volume shares in year $t-1$. Third, we aggregated over the products to an average price per hospital, whereby we weighted the hospital-product prices with the market-wide revenue shares for each product in $t-1$ ²⁷. We calculated an average price for the merged entity $M1 + M2$, by weighting the prices for hospitals $M1$ and $M2$ with their corresponding revenue shares in year $t-1$. We then removed the aggregations stepwise to show the effect of aggregating over products, locations and insurers until, finally, our results were disaggregated over all three sources of heterogeneity.

We investigated whether our results from the disaggregated model were robust to changes in the control groups by using six different control groups²⁸: (1) all Dutch hospitals that provide the product, excluding hospitals that also merged between years $t-2$ and $t+2$ and Independent Treatment Centers; (2) control group 1, excluding all university hospitals; (3) control group 2, excluding rivals of the merged hospitals; (4) control group 3, excluding the hospitals with low market power; (5) control group 3, excluding all hospitals with low health insurers concentration; and (6) control group 3, excluding hospitals of a different size to hospitals $M1$ and $M2$. We thus had twelve control groups: six for each hospital. Table 4 summarizes the number of hospitals in the control group.

The reasons behind the various exclusion criteria for the control groups were as follows. Control group 2 excludes all university hospitals because these generally spend more time on research and education and they usually treat patients with more complex problems than general acute care hospitals. This could result in different price trends. Control group 3 excludes the merged hospital's rivals, which were identified as such in the *ex ante* merger review by both the merged hospitals and the ACM. If the merger hospitals exercise their newly acquired market power by raising prices, their rivals may respond by also raising their prices (see e.g. Dafny, 2009; Gaynor & Vogt, 2003). Because of this rival-effect, rivals are excluded from control group 3. Hospitals with limited market power are excluded from control group 4. It is generally assumed that hospitals with a 55 percent market share or higher have significant market power (NZa, 2008; EC, 2004). Both hospital $M1$ and hospital $M2$ have a weighted average market share²⁹ of 55

27 We also estimated the models using the per hospital-product revenue in $t-1$ as a weighting factor for the aggregation over products. The results of these models do not differ from the main model and are therefore not included in this article. The results are available from the authors upon request.

28 We also wanted to know whether our disaggregated model was robust to hospital-specific covariates. As a sensitivity check, we therefore also included hospitals-specific covariates in an additional difference-in-differences model (i.e. the number of patients, the percentage of males, the average (standardized) socio-economic status score, the average age of the patients and the weighted market share per hospital). The results using this model did not differ from the other disaggregated model effects and are therefore not included in this article. The results are available from the authors upon request.

29 Measured by the inverse LOGit Competition Index – see section 6 for more information

Table 4. Number of hospitals in control groups for hospitals M1 and M2

	Hospital M1	Hospital M2
<i>Panel A. Hip replacements</i>		
Control group 1	55	55
Control group 2	50	50
Control group 3	46	46
Control group 4	38	38
Control group 5	41	41
Control group 6	36	40
<i>Panel B. Knee replacements</i>		
Control group 1	60	60
Control group 2	56	56
Control group 3	52	52
Control group 4	46	46
Control group 5	44	44
Control group 6	44	44
<i>Panel C. Cataract surgery</i>		
Control group 1	61	61
Control group 2	55	55
Control group 3	51	51
Control group 4	49	49
Control group 5	42	42
Control group 6	36	45

Notes: Control group 1 includes all Dutch hospitals that provide the product, excluding hospitals that also merged between years $t-2$ and $t+2$ and Independent Treatment Centers; control group 2 is control group 1 excluding all university hospitals; control group 3 is control group 2 excluding rivals of the merged hospitals; control group 4 is control group 3 excluding the hospitals with low market power; control group 5 is control group 3 excluding all hospitals with low health insurers concentration and control group 6 is control group 3 excluding hospitals of a different size to hospitals M1 and M2.

or higher for all three products. In control group 4, we therefore only take into account those hospitals that also have significant market power. We ranked the hospitals from control group 3 according to their weighted average market share and excluded the hospitals in the bottom quintile. Furthermore, to control for the effect of health insurers' concentration in each hospital in control group 5, we ranked the hospitals according to health insurers' HHI and excluded the hospitals in which the insurers' HHI was in the bottom quintile. Finally, in control group 6, we matched the hospitals that were in control group 3 with the volume of the merged hospitals. Hospital M2 had a much higher volume than hospital M1 and this difference in volume may have reflected different costs per unit product. We therefore matched two groups of equally sized hospitals with hospitals M1 and M2. For hospital M2, we ranked the hospitals by volume per product

and excluded the bottom quintile. For hospital M1, we ranked the hospitals by volume for each product and excluded the top quintile (for hip replacements and cataract surgeries) or the bottom quintile (for knee replacements).

6. DATA

We used a comprehensive nationwide patient-level dataset containing all inpatient and outpatient visits at all hospitals in the Netherlands. For each visit, the patient's zip code, age (year of birth), gender, health insurer, and DTC were observed, as well as the price negotiated for each hospital-insurer-product combination between years $t-2$ and $t+2$. Access to all patient-level data including negotiated prices from all insurers makes our dataset unique. The patient-level data that we used came from the insurers' claims administration and hospital registries, and was provided by the Dutch Healthcare Authority.

We focused on three products for which prices are freely negotiable: hip replacements³⁰, knee replacements³¹ (both orthopedics) and cataract surgery³² (ophthalmology). In year $t-1$, these product markets jointly accounted for 47.5 percent of turnover in the free-pricing segment at the merging hospitals³³. We checked for obvious outliers in the negotiated price data by studying the following for each outlier: the average price of the hospital-product combination; the average price of the health insurer-product combination; the price change in the hospital-product combination; the price change in the health insurer-product combination; and the price change in the hospital-insurer-product combination over the years. Only if the price deviated markedly from all the averages excluded the observation from the analysis³⁴. In all other cases, we could not

30 The definition used in the Dutch hospital product classification system is 'joint degeneration of pelvic/hip/upper leg; surgery with clinical admission and joint prosthesis'.

31 The definition used in the Dutch hospital product classification system is 'joint degeneration of knee; surgery with clinical admission and joint prosthesis'.

32 The definition used in the Dutch hospital product classification system is 'cataract; outpatient treatment with intervention'.

33 In hospital M1, hip replacements represented 18 percent, knee replacements represented 27 percent, and cataract surgeries represented 6 percent of the turnover in the competitive segment in year $t-1$. In hospital M2, hip replacements represented 16 percent, knee replacements represented 14 percent, and cataract surgeries represented 14 percent of the turnover in the competitive segment in year $t-1$. By way of comparison: in control group 1, hip replacements represented 15 percent, knee replacements represented 14 percent, and cataract surgeries represented 14 percent of the turnover in the competitive segment in year $t-1$.

34 In total, 73 hip replacements ($n=66,437$ before cleaning), 57 knee replacements ($n=61,404$ before cleaning) and 281 cataract surgeries ($n=47,6205$ before cleaning) were excluded from the dataset.

detect measurement error with certainty and we kept the prices in the dataset. All hospitals where more than 15% of prices were missing for one or more years between $t-2$ and $t+2$ were excluded from the dataset³⁵.

The pre-merger price was based on data from the year preceding the merger ($t-1$) and the post-merger price was based on data from the year after the merger ($t+1$). Table 5 presents summary statistics on the volume and mean prices of the products within hospital M1, hospital M2 and control group 1.

Table 5. Volume and mean prices for hip and knee replacements and cataract surgery in hospitals M1, M2 and control group 1

	Hip replacements		Knee replacements		Cataract surgeries	
	$t-1$	$t+1$	$t-1$	$t+1$	$t-1$	$t+1$
<i>Panel A. Hospital M1</i>						
Volume	172	173	222	282	381	355
Mean price (in €)	9189.58 (348.00)	10188.05 (559.08)	11022.98 (494.94)	11291.41 (651.32)	1405.00 (40.78)	1421.27 (45.08)
<i>Panel B. Hospital M2</i>						
Volume	389	503	271	295	2140	2077
Mean price (in €)	9181.96 (144.25)	8991.34 (109.09)	10959.49 (185.30)	10321.76 (245.90)	1400.10 (20.34)	1313.40 (29.83)
<i>Panel C. Control group 1</i>						
Volume	224	227	189	194	1520	1498
Mean price (in €)	9045.00 (338.64)	9160.96 (620.08)	10592.34 (473.51)	10608.52 (786.32)	1340.94 (72.83)	1349.43 (104.12)

Notes: The hospitals' volume per product in this table slightly deviates from the hospitals' volume per product reported in table 1. In this table we only report the records with a valid price, whereas in table 1 only records with a valid gender, age and SES-score per product per hospital are reported. The mean prices for each hospital are the averaged over all patients. The mean price for control group 1 is the average over the mean prices of the hospitals within control group 1. The standard errors are in parentheses.

Hospitals with limited market power are excluded from control group 4. The weighted average market share that was used to determine the hospitals' market power was based on the LQgit Competition Index (LOCI), developed by Akosa Antwi et al. (2006;

35 For hip replacements, 31 out of 90 hospitals had more than 15% missing prices in one or more years in the period $t-2$ and $t+2$ and were therefore excluded. For knee replacements, 25 out of 89 hospitals had more than 15% missing prices in one or more years in the period $t-2$ to $t+2$ and were therefore excluded. For cataract operations, 25 out of 89 hospitals had more than 15% missing prices in one or more years in the period $t-2$ to $t+2$ and were therefore excluded. The threshold of 15% was arbitrary. As a sensitivity check, we therefore also used other thresholds for the disaggregated model. This had no effect on the overall results or the conclusions of the article. The results are available upon request by the authors.

2009). The application of the method is explained in Gaynor and Town (2012) and NZa (2014). First, we calculated the hospitals' market share for each product in each zip code. The market share of hospital j for product d in zip code z is defined as $s_{jd,z} = \frac{q_{jd,z}}{\sum_{j=1}^J q_{jd,z}}$, where $q_{jd,z}$ is the total number of patients at hospital j ($j=1,...,J$) for product d ($d=1,2,3$) in zip code z ($z=1,...,Z$). Second, for each hospital and product, we calculated a weighted average market share $\bar{s}_{jd} = \sum_{z=1}^Z w_{jd,z} s_{jd,z}$, where we weighted each market share with its share in hospital j , i.e. $w_{jd,z} = \frac{q_{jd,z}}{\sum_{z=1}^Z q_{jd,z}}$.

The insurer's HHI that was used to construct control group 5 is based on the insurer's market shares for each product and ranged from zero to one³⁶. The insurer's HHI for hospital j and product d : insurer's $HHI_{jd} = \sum_{m=1}^M \left(\frac{q_{mjd}}{\sum_{m=1}^M q_{mjd}} \right)^2$, where q_{mjd} is the total number of patients of insurer m ($m=1, \dots, M$) in hospital j for product d .

7. EMPIRICAL RESULTS

To gain a picture of the change in the market structure as a result of the merger, we calculated the market share of the combined entity M1 + M2 for each product and compared it to the weighted average of the separate market shares of hospitals M1 and M2. Both calculations were based on the pre-merger market shares (i.e. from year $t-1$)³⁷. As expected, the weighted average market shares of the hospitals' products increased as a result of the merger. The increase is from 76.7% to 82.5% for hip replacements, from 78.2% to 85.7% for knee replacements, and from 83.5% to 86.6% for cataract surgeries. In table 6, we present the diversion shares of hospitals M1 and M2 that follow from the bargaining model presented in section 2. Diversion shares reflect the degree of substitution between hospitals. As indicated in section 2, a higher value of the diversion share suggests a higher degree of substitution between two hospitals in treating the same disease.

From table 6 it follows that the diversion shares of hospital M1 to hospital M2 are much higher. Hospital M1 is located in a more isolated region with hospital M2 being its strongest competitor pre-merger. As expected, a large share of patients is diverted to hospital M2 once hospital M1 is not available. If the more centrally located hospital M2 would not be available, however, only few patients are expected to be diverted to hospital M1. When comparing the diversion shares over products, we find that the variation in

36 Although it is also possible to calculate the hospitals' HHI, we opted for the weighted average market share that was based on the LOGit Competition Index (LOCI) because market delineation is necessary for the hospitals' HHI (in contrast to the insurers' HHI), but the use of market delineation methods in healthcare markets is the subject of increasing criticism (e.g. Elzinga & Swisher, 2011).

37 Measured by the inverse LOGit Competition Index – see section 6 for more information.

Table 6. Diversion shares TO/FROM hospitals M1 and M2 (in $t-1$)

To \ From	Hip replacements		Knee replacements		Cataract surgery	
	M1	M2	M1	M2	M1	M2
M1	-	0.105	-	0.158	-	0.034
M2	0.735	-	0.663	-	0.850	-

Notes: The diversion shares are calculated using a conditional logit model of hospital choice, following Capps et al. (2003). We used patient-level data from $t-1$ to estimate the model, which included the travel time between the patient's zip code and hospital location, a dummy indicating whether the patient is older or younger than 65, a dummy for the patient's gender and the socio-economic status score for the patient's zip code.

diversion shares across products within each hospital is much smaller than the variation in diversion shares across hospital M1 and M2 for each product. Table 7 shows the average price increases for hip replacements, knee replacements and cataract surgeries for control group 1 and the merged hospitals M1 and M2, indexed on the average price in control group 1 in year $t-1$.

Table 7. Price changes of hospitals M1, M2 and the control group pre- and post-merger (indexed on the average price in control group 1 in year $t-1$)

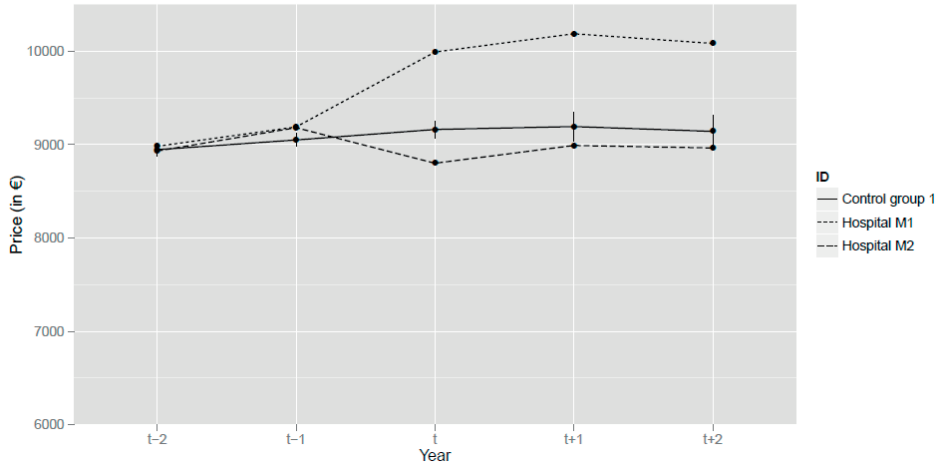
	Hospital					Control group 1				
	$t-2$	$t-1$	t	$t+1$	$t+2$	$t-2$	$t-1$	t	$t+1$	$t+2$
<i>Panel A. Hospital M1</i>										
Hip replacements	99	102	110	113	111	99	100	101	101	100
Knee replacements	101	104	105	107	105	99	100	100	99	101
Cataract surgery	101	103	102	104	100	98	100	99	99	95
<i>Panel B. Hospital M2</i>										
Hip replacements	99	102	97	99	99	99	100	101	101	100
Knee replacements	100	103	95	97	97	99	100	100	99	101
Cataract surgery	99	103	94	96	94	98	100	99	99	95

Notes: Indexed on the average price in control group 1 in year $t-1$; that is, the average price in control group 1 in $t-1$ is 100. The price for the control group is averaged over the mean prices of the hospitals in control group 1.

The table suggests that following the merger, both hospital locations charged different prices. As argued in section 4, the differences in competition intensity between the markets of hospitals M1 and M2, may induce the merged hospital to charge different prices. The prices for hip replacements did not change substantially between years $t-2$ and $t+2$ in control group 1. In comparison to the average control group prices in year $t-1$, the prices for hip replacements in hospital M1 increased by 13 percent after the merger (year t). This was the most substantial deviation from the average prices of control group 1 for year $t-1$.

As explained in section 5, however, price changes only give us a crude indication of the effect of the merger because they do not control for changes in prices that would have occurred anyway. We therefore estimate a model in which price changes at the merging hospitals are compared to price changes at a group of comparison hospitals which were unaffected by the merger (i.e. a difference-in-differences model). We visually investigate the common trend assumption on which the DID model is based. Figures 1-3 suggest that the pre-merger price change in the merged hospital did not deviate substantially from the pre-merger price changes in control group 1.

Figure 1. Average price development hip replacements in Hospitals M1, M2 and control group 1

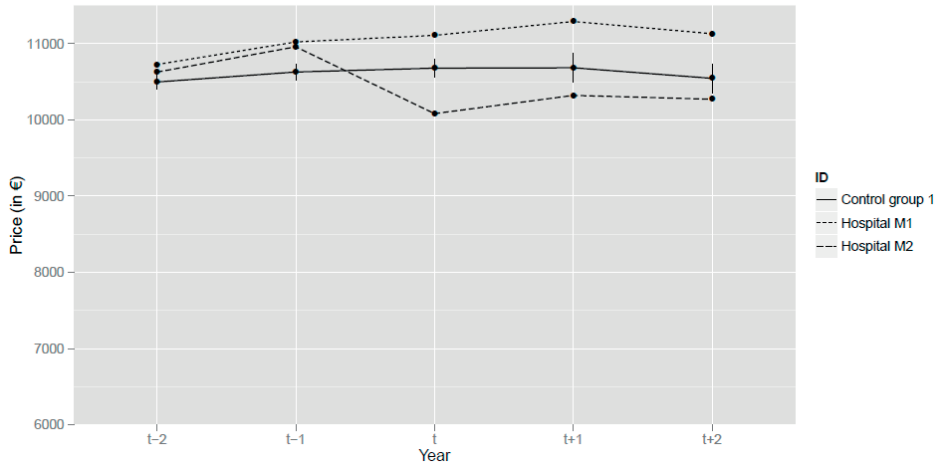


The prices plotted for control group 1 are averaged over all hospitals in control group 1. Control group 1 includes all Dutch hospitals that provide the product, excluding hospitals that also merged between years t-2 and t+2 and Independent Treatment Centers.

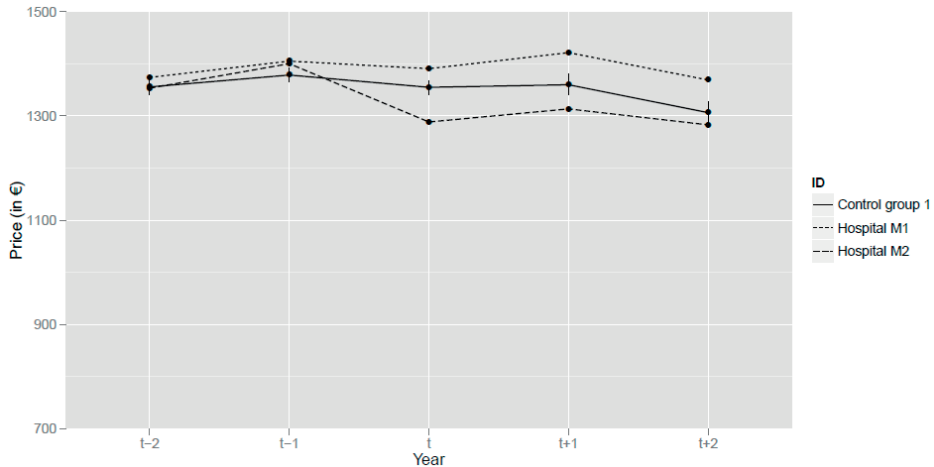
Table 8 presents the results of the difference-in-differences model aggregated over locations, insurers and products. Table 8 shows that no significant merger effect was observed when the result was aggregated over locations, insurers and products.

In table 9, we again show the price effect, aggregated over insurers, products and locations (panel A, column 1) but we then disaggregated the effect by location (panel A, column 2 and 3), by product (panels B to D, column 1), by location and product (panels B to D, columns 2 and 3), by insurer (panel E, column 1), and, finally, by insurer and location (panel E, columns 2 and 3).

If we only disaggregate by location, product or insurer, no significant merger effect is found. However, if we disaggregate by both product and location, we find that the merger led to significantly increased prices for hip replacements in hospital M1, by a total of 9 percentage points. This was the overall price effect of the merger for hip replacements in hospital M1. When the price effect was estimated over hospital locations

Figure 2. Average price development knee replacements in Hospitals M1, M2 and control group 1

The prices plotted for control group 1 are averaged over all hospitals in control group 1. Control group 1 includes all Dutch hospitals that provide the product, excluding hospitals that also merged between years $t-2$ and $t+2$ and Independent Treatment Centers.

Figure 3. Average price development cataract surgery in Hospitals M1, M2 and control group 1

The prices plotted for control group 1 are averaged over all hospitals in control group 1. Control group 1 includes all Dutch hospitals that provide the product, excluding hospitals that also merged between years $t-2$ and $t+2$ and Independent Treatment Centers.

and products, the effect disappeared. Also, if we disaggregated by insurer and location, we found that the merger only resulted in price changes for specific health insurers and only at hospital M1. In table 10, we disaggregate the merger effect by location, product and insurer.

Table 8. Merger effect aggregated over all three products, health insurers and hospital locations^A.

	Hospitals M1 & M2
(intercept)	8.869*** (0.029)
Post-merger price change in the common trend (λ)	0.009 (0.009)
Post-merger price change	-0.017 (0.057)
Observations	54
R-Squared	0.719
Adjusted R-Squared	0.422

Notes: Models estimated by Ordinary Least Squares (OLS) with standard errors in parentheses under coefficients. In this model, hospitals M1 and M2 together are compared to control group 1. Control group 1 includes all Dutch hospitals that provide the product, excluding hospitals that also merged between years $t-2$ and $t+2$ and Independent Treatment Centers. We aggregated the patient-level hospital data to a mean price per hospital. Firstly, we calculated an average price per product for each hospital-insurer pair. Secondly, we aggregated these prices over the insurers to an average price for each hospital-product combination, whereby we weighted the prices with the insurer's specific volume shares in year $t-1$. Thirdly, we aggregated over the products to an average price per hospital, whereby we weighted the hospital-product prices with the market-wide revenue shares for each product in $t-1$. We calculated an average price for the merged entity M1 + M2, by weighting the prices for hospitals M1 and M2 with their corresponding revenue shares in year $t-1$.

^A For clarity reasons, we do not report the hospital dummies here.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

In section 4 we explained that we disaggregated the post-merger price change for each hospital location to see whether the merging hospital differentiated a potential price increase after merger across locations. Table 7 suggested that the hospitals had done so and when we use the difference-in-differences approach we also found that the post-merger increase in prices for hip replacements in hospital M1 varied significantly from the control group, whereas the prices for hip replacements in hospital M2 were unaffected by the merger. Apparently, the merged hospital differentiated its prices across locations.

We also disaggregated the effect of the merger for each product. We found that the price effects of the merger varied significantly between hospital products. Specifically, the merger resulted in higher prices for hip replacements in hospital M1, whereas the prices for knee replacements and cataract care in hospitals M1 and M2 remained unaffected.

Finally, we disaggregated the post-merger price changes for each hospital-insurer combination. For four out of five health insurers that negotiated prices with hospital M1, the post-merger price increases for hip replacements were on average 13 percentage points higher than for the control groups. The merger's price effect varied between health insurers from -12 to 16 percentage points relative to the control groups. Also, the

Table 9. Merger effect for hip and knee replacements and cataract surgery stepwise disaggregation^A

	Hospitals M1 & M2	Hospital M1	Hospital M2
<i>Panel A. Aggregated over insurers & products</i>			
(intercept)	8.869*** (0.029)	8.869*** (0.029)	8.869*** (0.029)
Post-merger price change in the common trend (λ)	0.009 (0.009)	0.008 (0.008)	0.008 (0.008)
Post-merger price change	-0.017 (0.057)	0.053 (0.057)	-0.053 (0.057)
Observations	54	54	54
R-Squared	0.719	0.725	0.720
Adjusted R-Squared	0.422	0.434	0.423
<i>Panel B. Hip replacements: aggregated over insurers</i>			
(intercept)	9.130*** (0.027)	9.130*** (0.026)	9.130*** (0.026)
Post-merger price change in the common trend (λ)	0.014* (0.007)	0.014* (0.007)	0.014* (0.007)
Post-merger price change	0.005 (0.053)	0.090* (0.053)	-0.035 (0.053)
Observations	57	57	57
R-Squared	0.733	0.745	0.734
Adjusted R-Squared	0.452	0.476	0.453
<i>Panel C. Knee replacements: aggregated over insurers</i>			
(intercept)	9.311*** (0.031)	9.311*** (0.031)	9.311*** (0.031)
Post-merger price change in the common trend (λ)	0.003 (0.008)	0.004 (0.008)	0.004 (0.008)
Post-merger price change	-0.021 (0.063)	0.021 (0.062)	-0.064 (0.062)
Observations	57	62	62
R-Squared	0.708	0.709	0.707
Adjusted R-Squared	0.401	0.403	0.399
<i>Panel D. Cataract surgery: aggregated over insurers</i>			
(intercept)	7.249*** (0.029)	7.249*** (0.028)	7.249*** (0.028)
Post-merger price change in the common trend (λ)	-0.015** (0.007)	-0.015** (0.007)	-0.015** (0.007)
Post-merger price change	-0.038 (0.057)	0.027 (0.057)	-0.049 (0.057)
Observations	57	63	63
R-Squared	0.693	0.697	0.697
Adjusted R-Squared	0.371	0.378	0.378
<i>Panel E. Per insurer: aggregated over products</i>			
(intercept)	8.869*** (0.029)	8.869*** (0.029)	8.869*** (0.029)
Post-merger price change in the common trend (λ)	0.008 (0.008)	0.008 (0.008)	0.008 (0.008)
Post-merger price change insurer 1	-0.008 (0.057)	0.074 (0.057)	-0.052 (0.057)
Post-merger price change insurer 2	-0.008 (0.057)	0.049 (0.057)	-0.032 (0.057)
Post-merger price change insurer 3	-0.088 (0.057)	-0.137** (0.057)	-0.070 (0.057)
Post-merger price change insurer 4	0.054 (0.057)	0.115 (0.057)	-0.019 (0.057)
Post-merger price change insurer 5	-0.011 (0.057)	0.106* (0.057)	-0.046 (0.057)
Observations	54	53	53
R-Squared	0.742	0.796	0.728
Adjusted R-Squared	0.430	0.549	0.398

Table 9 Notes: Models estimated by Ordinary Least Squares (OLS) with standard errors in parentheses under coefficients. In this model, hospitals M1, M2 and M1 and M2 together are compared to control group 1. Control group 1 includes all Dutch hospitals that provide the product, excluding hospitals that also merged between years $t-2$ and $t+2$ and Independent Treatment Centers. The data for the model in *Panel A*, columns 2 & 3 is aggregated over health insurers and products: (i) we calculated an average price per product for each hospital-insurer pair, (ii) we aggregated these prices over the insurers to an average price for each hospital-product combination, whereby we weighted the prices with the insurer's specific volume shares in year $t-1$ and (iii) we aggregated over the products to an average price per hospital, whereby we weighted the hospital-product prices with the market-wide revenue shares for each product in $t-1$. The data for the model in *Panels B, C & D*, columns 2 & 3 is aggregated over health insurers: (i) we calculated an average price per product for each hospital-insurer pair, and (ii) we aggregated these prices over the insurers to an average price for each hospital-product combination, whereby we weighted the prices with the insurer's specific volume shares in year $t-1$. The data for the model in *Panel E*, column 2 & 3 is aggregated for the *control group* as in *Panel A*. For the merged hospital entity M1+M2 (*column 1*) the data is aggregated: (i) we calculated an average price per product for each hospital-insurer pair, (ii) we aggregated these prices over the insurers to an average price for each hospital-product combination, whereby we weighted the prices with the insurer's specific volume shares in year $t-1$. We calculated an average price for the merged entity M1 + M2, by weighting the prices for hospitals M1 and M2 with their corresponding revenue shares in year $t-1$.

^A For clarity reasons, we do not report the hospital dummies here.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

largest health insurer – insurer 1, which represented 76 percent of hospital M1's patients – was unable to negotiate lower prices: the prices it paid for hip replacements rose by 11 percentage points as a result of the merger. In contrast, one of the four other much smaller health insurers – insurer 3, which represented only 11 percent of hospital M1's patients – was able to negotiate prices that were much lower than the control groups. These results were robust between the control groups. It is therefore less likely that the merger effect estimated was driven by unobserved characteristics in the control group³⁸.

Hence, what we can deduct from these tables is that aggregating the merger effect over locations, products and insurers masked considerable variations between locations, products and insurers. In other words, failing to disaggregate would prevent us from detecting the price effects of a hospital merger.

8. DISCUSSION

The main finding of our study is that a merger between two hospitals in overlapping geographical markets generated heterogeneous prices effects at the two different hospital locations, for different hospital products and for different health insurers. The theoretical model that was presented in section 2 explains why this might be the case.

38 These findings are not reported in the article, but are available upon request by the authors.

Table 10. Merger effect for hip and knee replacements and cataract surgery per health insurer in hospitals M1 & M2^A

	Hip replacements	Knee replacements	Cataract surgeries
<i>Panel A. Hospital M1</i>			
(intercept)	9.130*** (0.026)	9.311*** (0.031)	7.249*** (0.028)
Post-merger price change in the common trend (λ)	0.014* (0.007)	0.004 (0.008)	-0.015** (0.007)
Post-merger price change insurer 1	0.113** (0.053)	0.049 (0.062)	0.037 (0.057)
Post-merger price change insurer 2	0.099* (0.053)	0.024 (0.062)	-0.053 (0.057)
Post-merger price change insurer 3	-0.118** (0.053)	-0.153** (0.062)	-0.114** (0.057)
Post-merger price change insurer 4	0.157*** (0.053)	0.089 (0.062)	0.067 (0.057)
Post-merger price change insurer 5	0.147*** (0.053)	0.080 (0.062)	0.059 (0.057)
Observations	57	62	63
R-Squared	0.828	0.767	0.740
Adjusted R-Squared	0.617	0.487	0.429
<i>Panel B. Hospital M2</i>			
(intercept)	9.130*** (0.026)	9.311*** (0.031)	7.249*** (0.028)
Post-merger price change in the common trend (λ)	0.014* (0.007)	0.004 (0.008)	-0.015** (0.007)
Post-merger price change insurer 1	-0.032 (0.053)	-0.066 (0.062)	-0.051 (0.057)
Post-merger price change insurer 2	-0.029 (0.053)	-0.035 (0.062)	-0.016 (0.057)
Post-merger price change insurer 3	-0.049 (0.053)	-0.084 (0.062)	-0.074 (0.057)
Post-merger price change insurer 4	-0.021 (0.053)	-0.016 (0.062)	-0.010 (0.057)
Post-merger price change insurer 5	-0.044 (0.053)	-0.049 (0.062)	-0.022 (0.057)
Observations	57	62	63
R-Squared	0.738	0.716	0.706
Adjusted R-Squared	0.417	0.375	0.354

Notes: Models estimated by Ordinary Least Squares (OLS) with standard errors in parentheses under coefficients. In this model, hospital M1 and M2 are compared to control group 1 which includes all hospitals excluding other merging hospitals and Independent Treatment Centers. The data for this model is aggregated for the *control group* as follows: (i) we calculated an average price per product for each hospital-insurer pair, (ii) we aggregated these prices over the insurers to an average price for each hospital-product combination, whereby we weighted the prices with the insurer's specific volume shares in year $t-1$. For the *merging hospitals* the data is aggregated as follows: an average price per product for each hospital-insurer pair.

^A For clarity reasons, we do not report the hospital dummies here.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Different price effects for different products

First, when we compare the price effects of a hospital merger on hip replacements, knee replacements and cataract surgery, we find a significant increase in the post-merger price of hip replacements but not of the other two products. This result was robust across all control groups and model specifications. In section 5, we explained that large

post-merger price increases for the merged hospitals in comparison to prices among a control group could be consistent with at least four hypotheses. By a close consideration of the market under study, we can rule out the possibility that the increase in the post-merger price of hip replacements can be explained by a catching-up effect, or by an increase in quality or case mix severity. This because the pre-merger prices of hip replacements in hospital M1 were no lower than the prices at the comparison hospitals, as table 5 shows. Also, the pre-merger price for hip replacements at hospital M1 corresponds to the pre-merger price for hip replacements at hospital M2. According to the 'learning about demand' explanation, following a merger, a hospital is able to observe the prices paid to one of its former competitors, revealing potentially important information about the willingness of health insurers to pay for hospital services (see Adams & Noether, 2011). This explanation, however, cannot apply here as the pre-merger prices are similar. Furthermore, it is unlikely that the quality of care for hip replacements increased in hospital M1 following the merger. Although the hospital advertised quality increases in other procedures during the study period, this did not include the quality of its hip replacements. Furthermore, if it were the case that hospital M1 increased its quality because it learned from hospital M2 following the merger, we would expect prices to converge between the locations, but this did not happen. Also, an increase in quality that would justify such a large price increase (9 percentage points on average) would most likely also have an effect on patient volume at the expense of patient numbers at hospital M2 or rival hospitals, but this did not occur either. Therefore, we find it unlikely that an increase in quality between $t-1$ and $t+1$ can account for the price increase for hip replacements in hospital M1. From table 1 it also follows that the demographic characteristics of the patients at hospital M1 did not change much following the merger. The number of males increased slightly, but as the number of males increased in almost all hospitals, this cannot explain the increase in the prices for hip replacements at hospital M1. Also, it is more likely that if the patients' case mix had increased post-merger, more complex cases would have gone to hospital M2 rather than to hospital M1 because hospital M2 is a larger general hospital that also provides tertiary care. In view of this, the most plausible explanation out of the four possible explanations that follow from the empirical literature is that the merger enhanced the market power of the hospitals.

However, this raises the question of why the price rise only occurred for hip replacements and not for knee replacements and cataract surgery. It is possible that this was due to a different level of competition intensity for these products. Indeed it followed from the theoretical framework that product d 's price change after the merger in each hospital is increasing in the diversion share between these hospitals, as well as the price-cost margin of the partnering hospital. We found that the diversion shares in hospital M1 of hip replacements were no higher than the diversion shares of other products. In fact, the diversion share of cataract surgeries is higher, whereas the price change for cataract

surgeries in hospital M1 after merger is not significant. Hence, based on the conclusions from the theoretical model, the difference in product-price effects after merger must be explained by other factors, i.e. the pre-merger price-cost margins of hospital M2. Unfortunately, we have no information on the product's price-cost margins of hospitals in this market. However, because the pre-merger prices for hip replacements in hospitals M1 and M2 were remarkably similar according to table 5, the theory suggests that the pre-merger cost of hip replacements at hospital M2 were lower than the pre-merger cost of hip replacements at hospital M1.

Nevertheless, the finding that price effects *are* heterogeneous across hospitals' top-revenue products highlights the importance of using a more disaggregated approach rather than the more aggregated approach, when defining product markets. In practice, it is often assumed that the merger price effect will be the same for all hospital products because acute care, inpatient services can be considered as a single and thus homogeneous hospital product in cases of hospital mergers. Typically, antitrust agencies use a cluster approach to define hospital product markets and most empirical studies follow this approach and look at the aggregated price effects of hospital mergers. Also, the bargaining models that were developed to reflect hospital-insurer bargaining assume that a hospital system and an insurer bargain over a single base price per hospital location. In section 2, we already noted that freeing the product price ratios would more closely correspond to hospital-insurer bargaining in practice. The hospital market is highly complex due to the multiplicity of services offered and the heterogeneity of consumers and therefore many different hospital products exist. Sacher and Silvia (1998) show that using the standard inpatient cluster may mask considerable variability in the concentration statistics across the inpatient categories that make up an overall cluster. They argue that disaggregation can provide a better understanding of the potential competition effects of a merger in a range of market configurations. A similar point is made by Hentschker et al. (2014).

Also, from the theoretical model it followed that price effects after merger may differ between hospital products. For that reason, when we estimated the model parameters, we also disaggregated the effects of the merger by product markets. Like Sacher and Silvia (1998), we find that disaggregation can provide a fuller understanding of the potential competitive effects of the merger. However, if potential competitive effects are not homogeneous over product markets this may have important implications for future antitrust scrutiny. If the rules for market definition that are formulated in the EC

merger guidelines (EC, 1997)³⁹, as well as in the US merger guidelines (FTC, 2010)⁴⁰, were applied strictly, hundreds or maybe thousands of separate hospital product markets would have to be distinguished because many hospital products and services are not demand or supply substitutes as prescribed by these rules. Clearly this would not be a feasible strategy in cases of hospital mergers. Hence, only a certain level of disaggregation would be warranted. Although our theoretical model defines each product d as a treatment of one illness, d may also be understood as a product cluster combining several illnesses based on revenue or volume or specialism or otherwise. Hence, the model conclusions also hold for the situation in which some clustering (aggregation) is applied in order to reduce the number of product dimensions in the analysis or because this more closely corresponds with practice. Sacher and Silvia (1998) show that even a very limited disaggregation of the standard inpatient cluster can lead to a more accurate merger analysis. Zwanziger et al. (1994), too, propose a manageable disaggregation of the standard clusters. Because it is unclear how often antitrust outcomes would be affected by using a different level of aggregation (Sacher & Silvia, 1998), we suggest using both the clustered approach as well as a limited disaggregated approach when defining product markets in the case of hospital mergers. One feasible approach may then be similar to our approach in which at least the top 3 or top 5 of the highest revenue products affected by the merger are analyzed separately. If the initial disaggregated approach gives rise to suspicions, the analysis can be further disaggregated⁴¹.

If antitrust authorities indeed decide to conduct disaggregated analyses, it is an interesting question how an antitrust authority should deal with differences in merger outcomes between products. It is unlikely that the antitrust authority will block a merger if the prospective analysis indicates that the prices for one product will increase, whereas the prices of other products will not be affected. Rather, finding different effects across products may lead to interventions that are specifically addressed only to the product that is found to be affected by the merger. For example, antitrust authorities may im-

39 According to the EC (1997) Commission Notice, 'A relevant product market comprises all those products and/or services which are regarded as interchangeable or substitutable by the consumer, by reason of the products' characteristics, their prices and their intended use'.

40 According to the FTC (2010) Merger Guidelines: 'Market definition focuses solely on demand substitution factors, i.e. on consumers' ability and willingness to substitute away from one product to another in response to a price increase or a corresponding non-price change such as a reduction in product quality or service'.

41 In practice, antitrust authorities occasionally take potential differences between products into account. For example, in one case the UK Competition Commission performed a detailed analysis of the appropriate product markets (CC, 2013) and in the *FTC v. ProMedica Health System* case, the US antitrust authority paid special attention to the inpatient obstetrical services in addition to general acute-care inpatient services (FTC, 2012).

pose remedies requiring the divestiture of a specific product, imposing the obligation to support new entrants (like ITCs) or introducing a price ceiling on particular products at one or more hospital locations.

Different price effects at different locations

Second, the merged hospital raised its price for hip replacements significantly at one location (hospital M1), but not at the other (hospital M2). To establish whether the merging hospitals experienced different price changes after merger, we aggregated the post-merger price change according to hospital location. It followed from the theoretical model that price changes caused by merger are proportional to the merging hospitals' diversion shares and the initial price-cost margins of the merger partner. To date, however, most studies have not controlled for this potential source of heterogeneity. Only Tenn (2011) examines and finds evidence of differential pricing strategies after merger.

In our case study, the merging hospitals' diversion shares were different due to their geographic location. Hospital M1 is located in a more geographically isolated area. Hospital M2 was the strongest competitor to hospital M1 and therefore posed a major constraint on hospital M1's prices prior to the merger. Hospital M2, however, faced additional competition from other hospitals. This difference manifests itself in higher diversion shares for hospital M1 than for hospital M2 before merger (table 6). After the merger, the two hospitals were likely able to internalize this constraint, leading to higher prices at hospital M1. They were able to do this without being penalized by rivals because hospital M1 experienced competitive pressure from only one rival hospital after the merger. By contrast, hospital M2 still experienced significant competitive pressure from five other hospitals after the merger. In this setting, differentiating prices according to the location may be a profitable strategy for the merged hospital: hospital M1 was in a position to raise its prices whereas maintaining a steady flow of patients, whereas hospital M2 maintained its prices at the pre-merger level in order to prevent losing patients to a rival hospital. Our results are consistent with this line of reasoning: the price change after merger was higher for hospital M1 whose diversion shares to hospital M2 were much higher than the diversions shares from hospital M2 to hospital M1.

By means of our empirical analysis we showed that it needs to be recognized that a merger between a rather isolated hospital location and its closest substitute creates opportunities for post-merger price increases that may be overlooked when not taking the disaggregate approach. Our findings suggests that the competition intensity that merging locations experience before and after merger may differ considerably between locations even if the merger entails two neighboring hospitals. Because this difference may result in a heterogeneous merger effects across locations, antitrust agencies should take the difference between locations into account. Given that these hospitals initially function as separate entities, the data that would be needed for the analysis at

the location level should be available. However, then the question remains how antitrust authorities should deal with differences in merger outcomes between locations. We discussed product-specific remedies in the previous paragraph. Likewise, antitrust authorities may think about location-specific remedies in case they predict the merger effect to be differentiated across locations. Like product-specific remedies, location-specific remedies might entail structural remedies or behavioral remedies that are only aimed at the location(s) that is (are) affected by merger⁴².

Different price effects for different insurers

The theoretical model that we presented in this article showed that the price change caused by merger may differ between health insurers. In our empirical analysis we disaggregated the overall results for each hospital-insurer combination which revealed that there is considerable heterogeneity across health insurers in the change in the post-merger negotiated prices. For four out of five health insurers that negotiated prices with hospital M1, the post-merger price increases for hip replacements were on average 13 percentage points higher than the control group. The merger's price effect varied between health insurers from -12 to 16 percentage points relative to the control group. This finding corresponds to the results from an earlier retrospective study from the US (Thompson, 2011), which indicated that two health insurers experienced price increases due to the hospital merger under study, whereas a third insurer experienced a price decrease and a fourth experienced no price effect from the merger.

The theoretical model suggests that the insurer-specific price differences may arise due to differences in the insurers' bargaining abilities. In particular, a health insurer with more bargaining weight or ability is confronted with a higher price increase after the merger.

The source of bargaining ability of health insurers is the topic of many studies. The evidence suggests that idiosyncratic effects such as bargaining skills of the individuals at the negotiating table might have a sizeable impact on the market outcomes (Sorensen, 2003; Halbersma et al., 2010; Grennan, 2014). Thompson (2012) furthermore suggests that the differences between insurers may be attributed to variations in the types of plans that the insurers offer and the services that they provide. Hence, although the bargaining model gives us some ideas on the source of heterogeneity in the post-merger

42 Occasionally, antitrust authorities have opted for imposing remedies at the entire location level. Divestitures of hospital locations were, for example, ordered by the US antitrust authority in the *FTC v. ProMedica Health System* case (FTC, 2012) and by the German antitrust authority in the *Asklepios/LBK Hamburg* case (Bundeskartellamt, 2005), whereas in the *Evanston Northwestern/Highland Park Hospital* case the US antitrust authority imposed a firewall so that the two firms had to negotiate separately with insurers after merger (FTC, 2008). See Gowrisankaran et al. (2015) for a critical review of the latter remedy.

price effects across health insurers, it remains largely unclear why such large differences exist across insurers within markets and why some health insurers experience price increases whereas others experience price decreases after merger. Because this is an issue that has been indicated a few times in research on hospital mergers (Thompson, 2012; Gaynor & Town, 2012), we suggest that further research on hospital-insurer bargaining should aim to establish the source of bargaining ability of health insurers in relation to hospital mergers.

From a policy perspective, the fact that post-merger price effects are not homogeneous across insurers within markets is an interesting finding, however. It is furthermore interesting to note that the heterogeneities are large. In *ex ante* merger reviews in the Netherlands, the Authority for Consumers and Markets (ACM) asks representatives of large health insurers in the region about their expectations regarding competitive effects of the merger. In fact, in the guidelines for assessing mergers and collaborations in healthcare, issued in 2013, the ACM says: 'When assessing a concentration's implications, the arguments put forward by insurers and patient organizations will be central.' (ACM, 2013). Like in most prospective merger cases, the representatives of the two largest health insurers in the region indicated that they did not anticipate negative competitive effects from the consolidation that we studied; and partly because of that reason the merger was cleared. However, the retrospective analysis indicates that the health insurers that believed to be able to counteract post-merger price increases were not both able to do that. We therefore suggest that a more critical assessment of health insurers' bargaining ability in merger cases is warranted.

9. CONCLUSION

In this study, we expanded existing bargaining models to allow for heterogeneous product-price effects and used a difference-in-differences model in which price changes at the merging hospitals are compared to price changes at a group of comparison hospitals. The main finding of our study is that the merger led to heterogeneous price effects for different health insurers, hospital products and hospital locations and that these differences depend on (i) the degree of substitution between hospitals, which may also vary over products, (ii) the relative bargaining ability of hospitals and insurers and (iii) the pre-merger price-cost margins of different products delivered by these hospitals.

The theoretical model provided us with valuable insights on the sources of heterogeneity, whereas our detailed empirical analysis of a hospital merger improved our understanding of the magnitude of differences. The analysis, however, also gives rise to three areas for future research. First, it would be interesting to replicate this study for different hospital mergers to find out which of our findings persist. Second, more insight into the

sources of insurers' bargaining ability would be valuable. Third, analysis of pre-merger price-cost margins will improve our understanding of heterogeneous post-merger price effects across products.

Nevertheless, the fact that price effects of a merger are heterogeneous across products, locations and insurers signals important conclusions for *ex ante* merger scrutiny. First, it highlights the importance of using a disaggregated approach rather than the current cluster approach when defining product markets. Second, it suggests that future prospective merger analyses should take potential differences across hospital locations into account. Finally, it asks for a critical assessment of health insurers' bargaining ability in merger cases.

REFERENCES

- ACM (Authority for Consumers and Markets). 2013. *Assessing mergers and collaborations in hospital care*. Den Haag: Authority for Consumers and Markets (ACM).
- Adams, G. & M. Noether. 2011. 'Comment on "Hospital Mergers and Competitive Effects: Two Retrospective Analyses"'. *International Journal of the Economics of Business*. 18(1):33-40.
- Akosa Antwi, Y.O.D., M. Gaynor. & W.B. Vogt. 2009. 'A bargain at twice the price? California hospital prices in the new millennium'. *Forum for Health Economics and Policy*. 12(1), Article 3.
- Akosa Antwi, Y.O.D., M. Gaynor. & W.B. Vogt. 2006. *A competition index for differentiated products oligopoly with an application to hospital markets*. Carnegie Mellon University.
- Berenson, R.A., T. Bodenheimer & H.H. Pham. 2006. 'Specialty-service lines: salvos in the new medical arms race'. *Health Affairs*. 25(5):337-343.
- Bertrand, M., E. Duflo & S. Mullainathan. 2004. 'How much should we trust difference-in-difference estimates? *Quarterly Journal of Economics*. 119:249-275.
- Binmore, K., A. Rubinstein. & A. Wolinsky. 1986. 'The Nash bargaining solution in economics modelling'. *RAND Journal of Economics*. 176-188.
- Blundell, R. & M. Costa Dias. 2009. 'Alternative Approaches to Evaluation in Empirical Microeconomics'. *The Journal of Human Resources*. 44(3):565-640.
- Bundeskartellamt, 2005. *Beschluss Asklepios Kliniken GmbH / Landesbetrieb Krankenhäuser Hamburg GmbH*. Case number: B 10 – 161/04. Bonn: Bundeskartellamt. Available at: <http://www.bundeskartellamt.de/SharedDocs/Entscheidung/DE/Entscheidungen/Fusionskontrolle/2005/B10-161-04.pdf>
- Capps, C, D. Dranove & M. Satterthwaite. 2003. 'Competition and market power in option demand markets'. *RAND Journal of Economics*. 34(4):737-763.
- Carey, K., J. F. Burgess jr. & G.J. Young. 2011. 'Hospital competition and financial performance: the effects of ambulatory surgery centers'. *Health Economics*. 20:571-581.
- CC (Competition Commission). 2013. *Final Report on The Royal Bournemouth and Christchurch Hospitals NHS Foundation Trust/Poole Hospital NHS Foundation Trust: A Report on the anticipated merger of The Royal Bournemouth and Christchurch Hospitals NHS Foundation Trust and Poole Hospital NHS Foundation Trust*. London: Competition Commission.
- Chernew, M.E., R.E. Mechanic, B.E. Landon & D. Gelb Safran. 2011. 'Private-payer innovation in Massachusetts: the 'Alternative Quality Contract''. *Health Affairs*. 30(1):51-61.
- Collard-Wexler, A., G. Gowrisankaran & R.S. Lee. 2014. *Bargaining in bilateral oligopoly: an alternating offers representations of the "Nash-in-Nash" solution*. Washington: National Bureau of Economics Research.
- CZ. 2015. *Document Selective Contracting Policy in 2016 – Best practice in breast cancer* (Verantwoordingsdocument selectief inkoopbeleid 2016 – Beste zorg bij borstkanker). [In Dutch].<https://www.cz.nl/~media/actueel/inkoop%20van%20zorg/verantwoordingsdocument%20selectieve%20inkoop%20borstkankerzorg.pdf>. Tilburg: CZ.
- Dafny, L. 2009. 'Estimation and Identification of Merger Effects: An Application to Hospital Mergers'. *Journal of Law and Economics*. 52:523-550.
- Delbanco, S. 2014. *The payment reform landscape*. Health Affairs Blog July 2, 2014. <http://healthaffairs.org/blog/2014/07/02/the-payment-reform-landscape-bundled-payment/>.
- Donald, S.G. & K. Lang. 2007. 'Inference with difference-in-difference and other panel data'. *The Review of Economics and Statistics*. 89:221-233.

- Elzinga, K. G., & A. W. Swisher. 2011. "Limits of the Elzinga-Hogarty Test in Hospital Mergers: The Evanston Case" *International Journal of the Economics of Business* 18 (1): 133-146.
- EC (European Commission). 1997. *Commission Notice on the definition of relevant market for the purposes of Community competition law: 97/C 372/03*. Brussels: Official Journal of the European Communities.
- EC (European Commission). 2004. *Guidelines on the assessment of horizontal mergers under the Council Regulation on the control of concentrations between undertakings. (2004/C 31/03)*. 'Official Journal of the European Union': C 31: 5-18.
- FTC (Federal Trade Commission). 2012. *In the Matter of ProMedica Health System*. Washington DC: Federal Trade Commission. Docket number: 9346. Available at: <https://www.ftc.gov/enforcement/cases-proceedings/101-0167/promedica-health-system-inc-corporation-matter>
- FTC (Federal Trade Commission). 2010. *Horizontal Merger Guidelines*. Washington: US Department of Justice and Federal Trade Commission.
- FTC (Federal Trade Commission). 2008 *In the Matter of Evanston Northwestern Healthcare Corporation: Final Order*. Docket number: 9315. Washington DC: Federal Trade Commission. Available at: <https://www.ftc.gov/enforcement/cases-proceedings/0110234/evanston-northwestern-healthcare-corporation-enh-medical-group>.
- Gal-Or. E. 1997. 'Exclusionary equilibria in health-care markets'. *Journal of Economics & Management Strategy*. 6(1):5-43.
- Garmon. C. & D. Haas-Wilson. 2011. 'The Use of Multiple Control Groups and Data Sources as Validation in Retrospective Studies of Hospital Mergers'. *International Journal of the Economics of Business*. 18(1):41-44.
- Gaynor, M., K. Ho & R.J. Town. 2015. The industrial organization of health-care markets. *Journal of Economic Literature*. 53(2):235-284.
- Gaynor. M. & R.J. Town. 2012. *Chapter 9. Competition in Health Care Markets*. In: 'Handbook of Health Economics Volume 2' (eds: McGuire. T., M.V. Pauly & P. Pita Barros). Amsterdam: Elsevier North-Holland. 499-637.
- Gaynor. M. & W.B. Vogt. 2003. 'Competition among Hospitals'. *RAND Journal of Economics*. 34(4):764-784.
- Gaynor. M. & W.B. Vogt. 2000. *Chapter 27. Antitrust and Competition in Health Care Markets*. In: 'Handbook of Health Economics Volume 1' (eds.: Culyer A.J. & J.P. Newhouse). Oxford: Elsevier Science. 1405-1487.
- Gowrisankaran. G., A. Nevo & R. Town. 2015. 'Mergers when prices are negotiated: Evidence from the hospital industry'. *American Economic Review*. 105(1):172-203.
- Grennan, M. 2014. 'Bargaining Ability and Competitive Advantage: Empirical Evidence from Medical Devices'. *Management Science*. 30: 3011-3025.
- Haas-Wilson. D. & C. Garmon. 2011. 'Hospital Mergers and Competitive Effects: Two Retrospective Analyses'. *International Journal of the Economics of Business*. 18(1):17-32.
- Halbersma. R., M. Mikkers, E. Motchenkova. & I. Seinen. 2010. Market structure and hospital-insurers bargaining in the Netherlands. *European Journal of Health Economics*. 12(6):589-603.
- Hentschker. C., R. Mennicken. & A. Schmid. 2014. 'Defining hospital markets – an application to the German hospital sector'. *Health Economics Review*. 4:28.
- Ho. K. 2009. 'Insurer-provider networks in the medical care market'. *American Economic Review*. 99(1):393-430.
- Horn. H. & A. Wolinsky. 1988. 'Bilateral monopolies and incentives for merger'. *RAND Journal of Economics*. 19:408-419.
- Lakdawalla. D., & T. Philipson. 'The nonprofit sector and industry performance'. *Journal of Public Economics*. 90(8):1681-1698.

- Meijer, S., R. Douven. & B. van den Berg. 2010. 'The influence of regulated competition on the Dutch Hospital Market' (Hoe beïnvloedt geregleerde concurrentie de Nederlandse ziekenhuismarkt?). [In Dutch]. *TPEdigitaal*. 4(1):39-49.
- NZa. (Nederlandse Zorgautoriteit / Dutch Healthcare Authority). 2014. *Point of view on the Merger between Stichting Holding Isala Klinieken and Stichting Zorgcombinatie Noorderboog – Appendix 1: Merger simulations*. (Zienswijze Stichting Holding Isala Klinieken en Stichting Zorgcombinatie Noorderboog – Bijlage 1: Fusiesimulatiemodellen). Utrecht: Dutch Healthcare Authority (NZa).
- NZa. (Nederlandse Zorgautoriteit / Dutch Healthcare Authority). 2012. *Independent Treatment Centers: A qualitative and quantitative analysis. (Zelfstandige behandelcentra: een kwalitatieve en kwantitatieve analyse)*. [In Dutch]. Utrecht: Dutch Healthcare Authority (NZa).
- NZa. (Nederlandse Zorgautoriteit / Dutch Healthcare Authority). 2009. *Monitor Hospital care 2009*. (Monitor Ziekenhuizenzorg 2009). [In Dutch]. Den Haag: Dutch Healthcare Authority (NZa).
- NZa. (Nederlandse Zorgautoriteit / Dutch Healthcare Authority). 2008. *Explanatory notes on the policy on Significant Market power in healthcare*. (Toelichting op de beleidsregel Aanmerkelijk Marktmacht in de Zorg). [In Dutch]. Utrecht: Dutch Healthcare Authority (NZa).
- Porter, M.E. 2009. 'A Strategy for Health Care Reform – Toward a Value-Based System'. *The New England Journal of Medicine*. 361(1):109-112.
- Ruwaard, S., R. Douven, J. Struijs & J. Polder. 2014. 'Hospital care purchasing by Dutch health insurers: a study on the hospital-insurer contracts' (Hoe kopen zorgverzekeraars in bij ziekenhuizen: een analyse van de contracten tussen verzekeraars en ziekenhuizen). [In Dutch]. *TPEdigitaal*. 8(2):98-117.
- Sacher, S. & L. Silvia. 1998. 'Antitrust issues in defining the product market for hospital services'. *International Journal of the Economics of Business*. 5(2):181-202.
- Schut, F.T. & W.P.M.M. van de Ven. 2011. 'Effects of purchaser competition in the Dutch health system: is the glass half full or half empty?'. *Health Economics, Policy and Law*. 6: 109-123.
- Schut, F.T. & W.P.M.M. van de Ven. 2005. 'Rationing and competition in the Dutch health-care system'. *Health Economics*. 14:559-574.
- Song, Z., S. Rose, D. Gelb Safran, B.E. Landon, M.P. Day & M.E. Chernew. 'Changes in health care spending and quality: 4 years into global payment'. *New England Journal of Medicine*. 371(18):1704-1714.
- Sorensen, A.T. 2003. 'Insurer-hospital bargaining: negotiated discounts in post-deregulation Connecticut'. *The Journal of Industrial Economics*. L1(4):469-490.
- Tenn, S. 2011. 'The Price Effects of Hospital Mergers: A Case Study of the Sutter-Summit Transaction'. *International Journal of the Economics of Business*. 18(1):65-82.
- Thompson, A. 2011. 'The Effect of Hospital Mergers on Inpatient Prices: A Case Study of the New Hanover-Cape Fear Transaction'. *International Journal of the Economics of Business*. 18(1):91-101.
- Town, R. & G. Vistnes. 2001. 'Hospital competition in HMO networks'. *Journal of Health Economics*. 20(5):733-753.
- Varkevisser, M. & F.T. Schut. 2008. 'Hospital mergers need to be more rigorously reviewed by the Dutch Competition Authority' (NMA moet strenger zijn bij toetsen ziekenhuisfusies). [In Dutch]. *Economische Statistische Berichten*. 93(4532): 196-199.
- Ven, van de. W.P.M.M & F.T. Schut. 2008. 'Universal Mandatory Health Insurance In The Netherlands: A Model For The United States?'. *Health Affairs*. 27(3):771-781.
- Ven, van de. W.P.M.M. & F.T. Schut. 2009. 'Managed Competition in the Netherlands: Still Work-in-Progress'. *Health Economics*. 18:253-255.
- VGZ. 2014. *Your Health Insurance Policy in 2015 (Uw zorgverzekering in 2015)*. [In Dutch]. <https://www.vgz.nl/zorgverzekering/wijzigingen-2015>. Arnhem: VGZ.

- Vogt, W.B. & R. Town. 2006. *How has hospital consolidation affected the price and quality of hospital care?*. Robert Wood Johnson Foundation Research Synthesis Report No. 9.
- Zwanziger, J., G. Melnick. & K.M. Eyre. 1994. 'Hospitals and Antitrust: Defining Markets, Setting Standards'. *Journal of Health Politics, Policy and Law*. 19(2):423-446.

Chapter 3

Back to the Future: Predictive Power of the Option Demand Method in the Dutch Hospital Industry

With Anne-Fleur Roos

ACKNOWLEDGEMENTS

We are very grateful to Erik Schut, Marco Varkevisser, Danny Hughes, Misja Mikkers and Rein Halbersma for their valuable comments to the paper.

ABSTRACT

Antitrust authorities need new approaches to predict the effects of healthcare mergers. Merger simulation models are promising alternatives to highly debated traditional approaches, but they have only been validated to a limited extent. This paper evaluates the predictive power of the Option Demand method, a merger simulation model developed specifically for the US hospital market. We contrast the predictions of the merger simulation model to the estimated price effects of a consummated merger between two Dutch hospitals. We find that the Option Demand method could be a valuable addition to the antitrust agencies' toolkit, but that more research is necessary.

INTRODUCTION

In competitive markets, the aim of preventive merger control is to prohibit anticompetitive consolidation. To determine whether a merger between two (or more) firms will result in anticompetitive price increases (and/or quality decreases), antitrust authorities need to carry out an *ex ante* (prospective) review. Unfortunately, the approaches that are commonly used to prospectively review mergers are problematic. Generally, these methods first define the relevant market for the industry being studied and then use market shares to infer how the merger could affect competition in that market (Shapiro 2010; Werden and Froeb 2006). However, in order to delineate the relevant market, they typically rely on disputed methodologies and the conclusions drawn from the resulting analysis will depend heavily on how that market is defined. Moreover, these measurements are imperfect indicators of market power and so they do not necessarily reveal the actual exercise of market power. Merger reviews in the healthcare sector are subject to an additional difficulty because there are unique factors that render the most commonly used tests for measuring geographic markets less reliable in healthcare than in other sectors (Elzinga and Swisher 2011). Antitrust authorities therefore struggle to delineate the healthcare market effectively (Gaynor and Town 2012; Varkevissers and Schut 2012; Capps et al. 2002).

The most promising alternatives to these traditional approaches to review mergers are Merger Simulation Models (MSMs). The use of MSMs has clear advantages over the traditional approaches. MSMs use structural models to represent specific industries. By calibrating these models to the specifics of the market being studied, they can be used to predict the price effects of a merger directly (Werden 2005). Merger simulations take into account more than just market shares and concentration levels; they provide direct evidence and do not require or depend upon arbitrary market definitions (Argue and Shin 2009). For all these reasons, interest in MSMs is growing, both in the US and the EU (see e.g. Budzinski and Ruhmer 2009; Argue and Shin 2009; Walker 2005; Kalbfleisch 2005). However, the important question of whether MSMs are able to predict anticompetitive price increases accurately has not yet been answered conclusively. So far, MSMs have only been validated to a limited extent; they are always used in combination with traditional approaches and have rarely been subject to public scrutiny (Budzinski and Ruhmer 2009). Only merger simulation models that can produce reliable predictions are useful for merger policies, and the key issue with any merger simulation is its predictive capacity.

This paper contributes to the small, but growing, body of literature relating to the evaluation of merger simulation methods by evaluating the predictive powers of a reduced-form MSM that was developed specifically for hospital markets. The reduced-form MSM that we tested is referred to by its developers as the Option Demand method

(OD method). In the literature, this model is also referred to as the Capps, Dranove and Satterthwaite (CDS) or Willingness-to-Pay (WTP) model. The OD method is designed specifically to model markets in which managed care organizations or health insurers (selectively) contract with hospitals (Capps, Dranove, and Satterthwaite 2003; Town and Vistnes 2001). Recently, this model has been generalized by Gowrisankaran, Nevo and Town (2015)⁴³.

In this paper, we use the OD method to predict the price effects of a hospital merger between a general acute care hospital (hospital M1) and a neighboring general acute care hospital that also provides tertiary hospital care (hospital M2) that took place in the Dutch hospital market⁴⁴. From the viewpoint of the Option Demand method, the current Dutch healthcare system bears evident similarities with the US healthcare system. We explicitly take the multiproduct nature of hospitals into account by examining the price effects of the hospital merger for different hospital products. We also allow for potential differences in bargaining outcomes between neighboring locations by predicting the merger effects for each location. We use an instrumental variable approach to control for potential endogeneity issues. The actual price effects of the merger that we study are determined through a difference-in-difference (DiD) technique (Roos et al. 2017). By contrasting the simulated price effects with the actual price effects of the merger, we are able to evaluate the predictive power of the Option Demand method for hospital mergers in the Dutch context.

This paper is structured as follows. In section 1, we discuss how to identify unilateral effects after a horizontal merger and we consider the small number of available studies that evaluate the accuracy of merger simulation models. Section 2 describes the Option Demand model and discusses the applicability of the Option Demand method to the Dutch healthcare industry. Section 3 focuses on the modeling details of the Option Demand method and in section 4 we focus on the details of the estimation that we carried out. Section 5 describes the data that were used and section 4 presents the results. In section 7 we briefly discuss the findings of the retrospective study and compare the simulation results with the effects of the actual merger. In section 8 we present our conclusions on the predictive power of the reduced-form merger simulation model that we have applied.

43 For an extensive review of the literature on modeling hospital competition, see Gaynor, Ho and Town (2015).

44 For reasons of confidentiality, we anonymize the merged hospitals' and health insurers' names. For the same reason, the merger year is reported as t (t lies in the period 2005 – 2012).

1. MERGER SIMULATION MODELS

A. Identifying unilateral effects after a horizontal merger

According to most national and supranational antitrust laws, mergers must be reported to an antitrust authority prior to being consummated⁴⁵. After notification, the antitrust authorities carry out a review of the proposed merger, in which they make inferences regarding the expected anticompetitive effects of a merger in the relevant market. In general, horizontal mergers may give rise to two types of anticompetitive effects: (i) unilateral and (ii) coordinated effects. Both unilateral and coordinated effects may lead to higher post-merger prices, but prospective merger analyses focus predominantly on predicting the unilateral effects that a merger may cause (Baker 2003). In this paper, we also focus on the potential for unilateral effects.

Two methods are available to determine unilateral effects quantitatively: (i) a market definition approach and (ii) methods to predict unilateral effects directly. The market definition approach first defines the market and then hypothesizes on the merger-effect in that market. However, the market definition approach has several shortcomings, particularly when applied to the hospital industry (Dranove and Ody 2016; Elzinga and Swisher 2011; Gaynor, Kleiner, and Vogt 2011; Kaplow 2011; Shapiro 2010; Varkevisser, Capps, and Schut 2008; Capps et al. 2002). For example, the approach assumes that a product is either inside or outside the market. The products in the market are assumed to be subject to equal competitive pressure, while the products outside the market are not taken into account. However, in a market with differentiated products - which is typically the case for hospital markets - the degree of competition between two products depends on their substitutability and it is often difficult to draw meaningful boundaries between markets (Werden and Froeb 2006). Furthermore, it is only when very specific assumptions are made (e.g. homogeneous goods) that market shares can be translated into unilateral price effects (Kaplow 2011). The Elzinga-Hogarty test, in particular, has also been criticized because of its limited applicability to the hospital industry, mainly because of what has become known as the silent majority fallacy (e.g. Elzinga and Swisher 2011; Capps et al. 2001).

Given the drawbacks of the market definition approach, alternatives such as MSM that screen or predict anticompetitive effects directly and that circumvent the need for market delineation are promising alternatives.

45 See 15 USC §18A for the US and the competition laws of the EU Member States or EC:2004 for the European Union's rules on prior merger notification.

B. Merger Simulation Models

An MSM builds a structural model of the industry being studied. Typically, a structural model consists of (i) a demand model, which models the consumers' decision-making process and (ii) a model of competition, which models the supply-side of the market on the basis of the firm's behavior, the actions of its rivals and the consumer demand model. Having defined the competition model that best fits the industry being studied, the demand model can be estimated and the model of competition should be calibrated with pre-merger data. Next, a merger can be simulated by changing the ownership structure, for example by modeling that the number of competitors in a market decreases from 4 to 3 after merger (Budzinski and Ruhmer 2009).

A major issue with merger simulations is their predictive power and, thus, their credibility as a technique in the prospective merger review process (Budzinski and Ruhmer 2009). Only MSMs that are able to predict the actual effects of mergers accurately are useful for merger policy. Weinberg and Hosken (2013) stipulate that there are two methods for testing structural models: (i) the marginal costs approach, in which the actual (observed) marginal costs are contrasted with the marginal costs calculated by the calibrated simulation model; and (ii) the market structure approach, in which actual (observed) changes in price and/or quality following a merger are contrasted with the changes in price and/or quality simulated by the structural model. Budzinski and Ruhmer (2009), Werden and Froeb (2006) and Davis and Garcés (2010; chapter 8) describe both these methods in detail. Our study employs the second approach. Hence, we use past changes in market structure and the resulting price effects to test the accuracy of a (reduced-form) merger simulation model.

There are a handful of studies that have used the market structure approach to test merger simulation models. In addition to the three studies reviewed by Budzinski and Ruhmer (2009) (i.e. Pinkse and Slade 2004; Peters 2006 and Weinberg and Hosken 2013⁴⁶), Weinberg (2011), Friberg and Romahn (2015), Greenfield, Kreisle and Williams (2015) and Björnerstedt and Verboven (2016) also apply this approach. The studies differ in their efficacy (i.e. whether they are able to accurately predict price effects). In terms of methodology, they most often use a Bertrand model to model market competition. The studies use different demand functions to reflect the differences in industries and data and they also differ in the methodology that they employ to compare the simulated price changes to the actual price changes induced by the merger. Also, none of the previous studies have focused on hospital merger cases, although the problems that arise from using the more traditional market definition approaches are particularly strik-

46 Budzinski and Ruhmer (2009) review an earlier version (working paper: Weinberg, Matthew C. and Daniel Hosken. 2008. 'Using Mergers to Test a Model of Oligopoly'. *Working paper*. University of Georgia) of Weinberg and Hosken (2013).

ing in this sector (e.g. Elzinga and Swisher 2011). A notable exception is a recent FTC working paper (Garmon 2016) that reflects on the accuracy of hospital merger screening methods. The study concludes that the market definition approach is less accurate at predicting post-merger price effects than more recently developed models, including the Option Demand method (Garmon 2016). In contrast to Garmon (2016), we do not focus on contrasting the results of traditional approaches versus MSMs but rather on the predictive powers of one reduced-form MSM that is tailor-made for the healthcare industry: the Option Demand model.

2. THE OPTION DEMAND METHOD AND ITS APPLICABILITY TO THE DUTCH HEALTHCARE SYSTEM

A. What are Option Demand markets?

The Option Demand model that we evaluate in this paper was developed by Town and Vistnes (2001) and further refined by Capps, Dranove, and Satterthwaite (2003) and Gowrisankaran et al. (2015). The papers developed a framework for analyzing bargaining relationships between hospitals and insurers under selective contracting. Under such a healthcare system, consumers buy health insurance from health insurers. The consumers decide on a specific health insurance policy on the basis of the network of hospitals that the insurance contract offers and the premium. Each hospital renegotiates the terms of its contracts with health insurers on a regular basis. The idea is that the (threat of) selective contracting of hospitals may enable insurers to negotiate lower prices and/or higher quality, which may lower premiums (Ho 2009).

The OD method builds on this two-layer model of the hospital industry; that is, it models that (i) consumers buy health insurance from health insurers before fully knowing their medical needs and (ii) health insurers bargain and contract with healthcare providers (here: hospitals) on behalf of their insured. Following Dranove and White (1996), Capps, Dranove, and Satterthwaite (2003) refer to markets that exhibit these two layers as ‘Option Demand’ markets (or OD markets), since the consumer commits to a possibly restricted network of hospitals when he buys health insurance prior to knowing his future healthcare needs and when he is in need of specific care, he has the option of visiting any of the contracted hospitals. The value that a consumer then places on health insurance depends on his expected demand for healthcare and the expected utility that a particular hospital from this network will provide him. This value can be expressed as Willingness-To-Pay (WTP). The notion of WTP gives an estimate of how much consumers are willing to pay *ex ante* to retain access to this hospital in the network. The WTP is therefore a proxy of the hospital’s market power: a hospital with a high WTP score will

be better able to secure higher prices from the health insurer than a hospital with a low WTP score (Capps, Dranove, and Satterthwaite 2003:738).

B. The applicability of the Option Demand method to the Dutch healthcare sector

To date, the OD method has been applied by Capps, Dranove, and Satterthwaite (2003), by Dranove and Sfekas (2009) and by Dranove and Ody (2016) who find a positive relationship between hospital profits and WTP. Garmon (2016) finds an imprecise relationship between prices and WTP. The OD method has also been applied by the US Federal Trade Commission (Dranove and Ody 2016; Garmon 2016).

From the viewpoint of the Option Demand method, the current Dutch healthcare system bears similarities with the US healthcare system. In recent decades, the Dutch healthcare system has moved away from strict governmental supply-side regulation and towards regulated (or 'managed') competition (Van de Ven and Schut 2008; Schut and Van de Ven 2005). Of particular importance to this paper is the gradual introduction of hospital-insurer bargaining since 2005. In 2005, a product classification system for hospital and medical specialist care was introduced. Each activity and/or hospital service associated with a patient's demand for care, including outpatient care, is referred to as a Diagnosis and Treatment Combination (or DTC)⁴⁷. Following the introduction of the DTC system, the room for free negotiations between hospitals and health insurers on prices, volume and quality was gradually increased from 10% of hospitals' revenue in 2005, to 20% in 2008, to 34% in 2009 and to 70% in 2012. The remainder of hospital prices is regulated by the Dutch Healthcare Authority. For those services in the free-pricing segment, each hospital typically renegotiates the terms of its contracts with health insurers on an annual basis. Health insurers are allowed to contract selectively with healthcare providers.

The two-layer model that underlies the OD method seems to reflect the Dutch healthcare system accurately; consumers buy health insurance from health insurers and health insurers bargain and contract with hospitals on behalf of their enrollees. In the early years of the reform selective contracting was limitedly used. However, over the years, the number of health insurers offering contracts with restricted provider networks has increased. Furthermore, the introduction of a new Health Insurance Act has led to strong price competition between health insurers, and health insurers have put increasing pressure on hospitals to charge lower prices (Schut and Van de Ven 2011). The threat of selective contracting, rather than its actual use, may already have had an impact on

47 The DTC system is based on the concept of DRGs (Diagnosis-Related Groups) but constitutes a newly developed classification system.

hospital-insurer bargaining. We therefore consider the OD method applicable to the free-pricing segment of the Dutch hospital industry.

3. THE OPTION DEMAND METHOD: THE MODELING DETAILS

In this section, we describe how to estimate the demand model and Willingness-To-Pay (WTP) (section 3A), how to estimate the supply side and the competition model (section 3B) and how to simulate a merger with the WTP (section 3C). Our paper makes two modifications to the model by Capps, Dranove, and Satterthwaite (2003).

First, we explicitly take into account the multiproduct nature of hospitals by examining the price effects of the hospital merger for different hospital products. Typically, antitrust agencies use a cluster approach to define hospital product markets, assuming that ‘acute care, in-patient services’ can be considered as a single and thus homogeneous hospital product. Most empirical studies follow this approach and examine the aggregated price effects of hospital mergers. However, the hospital market is highly complex due to the multiplicity of services provided and the heterogeneity of consumers, which is in turn caused by differences in medical treatment needs and third-party payer coverage. Sacher and Silvia (1998) show that using the standard in-patient cluster may mask considerable variability in the concentration statistics across the in-patient categories that make up a whole cluster. They show that disaggregation can provide a fuller understanding of the potential competitive effects of a merger in a variety of market configurations. Roos et al. (2017) also find evidence of heterogeneous price effects across products in their retrospective case study. They studied the same merger case as the one simulated in this paper. We therefore also disaggregate the effect of the merger by product markets. We estimate the impact of the merger in three separate product markets that jointly represent 47.5 per cent of the merged hospital's turnover in the segment for which Dutch insurers and hospitals at the time of the merger were allowed to freely negotiate prices. The products included in this study are hip replacements, knee replacements and cataract surgery.

Second, our study allows for potential differences in bargaining outcomes between neighboring locations of merged hospitals by predicting the merger effects for each location. Hitherto, most studies have aggregated the merger effect, thereby disregarding the fact that post-merger differences in market power for each location may lead to opportunities to differentiate pricing strategies. In the case of multiple locations, price differentiation across locations may be a profitable strategy for the merged hospital. In retrospective studies, Roos et al. (2017) and Tenn (2011) find evidence of differential pricing strategies in hospital mergers. However, most previous studies on mergers have not controlled for this potential source of heterogeneity. We disaggregate the predicted

price change for each hospital location. In sections 2C and 4, we will explain in more detail how we handled the modification of the model by Capps, Dranove, and Satterthwaite (2003) in our paper. We also discuss the relationship with the extension of the OD model by Gowrisankaran et al. (2015).

A. Step 1: demand model and Willingness-To-Pay (WTP)

Under the OD method, a consumer's demand for hospital treatment is modeled using a conditional logit demand function (see McFadden 1974). Under this model, patient i seeks treatment after falling ill. His health insurance gives him access to network G of hospitals (all the available hospitals in the market). The expected utility of patient i for receiving treatment at hospital j is given by: $U_{ij} = U(H_j, X_i, \lambda_i) + \varepsilon_{ij}$ where H_j is a vector of hospital j characteristics. X_i is a vector which combines the characteristics and clinical attributes of patient i . The patient's travel time (λ_i) is determined by the distance between the patient's location (e.g. zip code) and the hospital j . Under the conditional logit demand function, we assume that the residuals (ε) are i.i.d. with the double standard exponential distribution (see McFadden 1974)⁴⁸.

Using a logit demand model, the probability that patient i chooses hospital j is given by: $s_j(H_j, X_i, \lambda_i) = \frac{\exp[U(H_j, X_i, \lambda_i)]}{\sum_{g \in G} \exp[U(H_g, X_i, \lambda_i)]}$.

Denote the utility of patient i for access to network G as $V^{ul}(G, X_i, \lambda_i)$. The WTP of patient i for hospital j , denoted by $\Delta V_{ij}^{ul}(G, X_i, \lambda_i)$, is the reduction in V^{ul} due to the exclusion of hospital j from network G . Hence, $\Delta V_{ij}^{ul}(G, X_i, \lambda_i) = V_{ij}^{ul}(G, X_i, \lambda_i) - V_{ij}^{ul}(G/j, X_i, \lambda_i)$, where G/j is the network of hospitals G excluding hospital j . Capps et al. (2003) show that it follows from the logit demand that for the WTP of patient i for hospital j that:

$$\Delta V_{ij}^{ul}(G, X_i, \lambda_i) = \ln \left[\frac{1}{1 - s_j(H_j, X_i, \lambda_i)} \right].$$

The *ex ante* WTP for the entire population (with N ill consumers) of hospital j is the weighted sum of the patients' WTPs (Capps, Dranove, and Satterthwaite 2003:743):

$$W_j = N \int_{X, \lambda} \frac{1}{N} \ln \left[\frac{1}{1 - s_j(H_j, X_i, \lambda_i)} \right] f(X_i, \lambda_i) dX_i d\lambda_i,$$

48 To avoid the IIA property that underlies the conditional logit functions, some studies use the mixed logit model to analyze patient hospital choice (see e.g. Pope 2009; Varkevisser, van der Geest, and Schut 2012). Farrell et al. (2011) find that there is almost no difference in the estimated hospital-level diversions in the patient-level mixed logit compared to the standard patient-level conditional logit model. Recent studies on hospital choice use the conditional logit model (e.g. Chandra et al. 2016; Gaynor, Propper, and Seiler 2016; Gutacker et al. 2016; Frank et al. 2015; Chou et al. 2014; Ho, and Pakes 2014).

where the population density distribution of all ill consumers is given by $f(X_i, \lambda_i)$ and constant γ convert utils into monetary terms. Since we do not observe constant γ , we use WTP up to the unidentified scale factor. For our application this is sufficient, since we are not interested in the exact value of the WTP.

We apply the discrete equivalent of the above equation to calculate the WTP of each hospital (see also Garmon 2016; Balan and Brand 2015; Farrell et al. 2011). Further, following Farrell et al. (2011), we rescale the WTP according to the hospital's expected number of patients. The rescaled discrete WTP equation for hospital j is⁴⁹:

$$w_j = \frac{\sum_{i=1}^N \ln\left[\frac{1}{1 - s_j(H_j, X_i, \lambda_i)}\right]}{\sum_{i=1}^N s_j(H_j, X_i, \lambda_i)}.$$

B. Step 2: supply side and competition model

Under the OD method, the idea is that if a hospital adds a high value to the health insurance network, it will be able to extract more profits from its negotiations and vice versa. Hospitals and insurers thus bargain according to the total value that hospital j adds to the health insurance network, i.e. w_j . Following Capps, Dranove, and Satterthwaite (2003), we model this negotiation with a reduced-form bargaining model:

$$(2) \quad p_j - c_j = \alpha \cdot w_j,$$

where p_j is the revenue per patient and c_j is the variable cost per patient. This equation thus gives the relationship between the margin of hospital j , i.e. the per-patient revenue minus the variable cost per patient, and the WTP per patient for hospital j . The per-patient gain of including hospital j in the network is split between the hospital and the insurer. Parameter α is the proportion that each hospital captures ($0 \leq \alpha \leq 1$). Parameter α is fixed and depends on the parties' relative bargaining abilities (Farrell et al. 2011).

Gowrisankaran et al. (2015) present a structural bargaining model that is more general than the Capps et al. (2003) model that we present here. Gowrisankaran et al. (2015) show that the Capps et al. (2003) model is a special case of their structural bargaining model. An important extension in the model of Gowrisankaran et al. (2015) is that patients face coinsurances. The Capps et al. (2003) model assumes that there is no coinsurance, which simplifies the bargaining model. In the Dutch market, there is no coinsurance. There is a yearly mandatory deductible that the patient pays when he starts using healthcare. However, the deductible is limited to a relatively small fixed amount (220 euro per year

49 The unscaled WTP employed by Capps, Dranove, and Satterthwaite (2003) also increases with the number of patients that a hospital treats. This is undesirable. The rescaled WTP is high only if a hospital does not have close substitutes.

in 2012). Since most hospital prices are higher than this amount, each patient receiving treatment at any hospital would generally pay the same deductible. Hence, deductibles are expected to hardly affect patient hospital choice, which implies that the no out-of-pocket payment assumption is also justifiable in our application of the model. Another extension of Gowrisankaran et al. (2015) is that they take health insurers' costs into account in the bargaining model. However, following Capps et al. (2003) and as is often done in practice (Gaynor, Ho and Town 2015), we regress WTP measures on price, and add marginal cost controls to the regression in our reduced-form merger simulation⁵⁰.

C. Step 3: merger simulation with WTP

In a merger review, antitrust authorities need to make an *ex ante* review to find out whether the merger between two (or more) hospitals will result in anticompetitive price increases. In our model, this means that we are interested in the increase in the post-merger prices of entity $j+k$ compared to the pre-merger prices of hospitals j and k . If we know the demand, WTP and bargaining model, we can calculate the post-merger WTP of the new entity and the post-merger price increase of the merged entity by estimating α (Capps, Dranove, and Satterthwaite 2003).

This works as follows. Let us assume that we want to predict the increase in prices due to a merger between hospitals j and k . With equation (1), we can calculate the pre-merger WTP of hospitals j and k , which we will denote with w_j^{pre} and w_k^{pre} . Post-merger, hospitals j and k form one entity. The weighted joint pre-merger WTP of hospitals j and m is: $w_{j+m}^{pre} := S_j w_j^{pre} + S_k w_k^{pre}$ where S_j is the post-merger revenue share of hospital j in the merged hospital and S_k is the post-merger share of hospital k in the merged hospital. We assume that the merged firm will bargain on an all-or-nothing basis (i.e. the merged hospitals are either in or out of the insurer's network and reimbursement for patients visiting that hospital is therefore either 100% or 0%⁵¹). Thus, post-merger, the WTP of entity $j+k$ is:

$$w_{j+k}^{post} = \frac{\sum_{i=1}^N \ln \left[\frac{1}{1 - s_j(H_j, Y_i, Z_i, \lambda_i) - s_k(H_k, Y_i, Z_i, \lambda_i)} \right]}{\sum_{i=1}^N (s_j(H_j, Y_i, Z_i, \lambda_i) + s_k(H_k, Y_i, Z_i, \lambda_i))}.$$

The increase in WTP due to the merger for the combined entity is then $w_{j+k}^{post} - w_{j+k}^{pre}$.

Given bargaining model (2), we can calculate the increase in the $j+k$ entities' margin with: $(p_{j+k}^{post} - c_{j+k}^{post}) - (p_{j+k}^{pre} - c_{j+k}^{pre}) = \alpha \cdot (w_{j+k}^{post} - w_{j+k}^{pre})$. Using equation (2) the α can be estimat-

50 In a Monte Carlo setting, Balan and Brand (2015) compared the true price effects of more general bargaining models with WTP-based merger simulation methods. They conclude that generally the WTP-based merger simulation methods perform well.

51 In practice, this is the most common negotiating strategy of hospitals. The assumption can, however, be relaxed by adapting WTP to separate bargaining scenarios (Brand and Garmon 2014).

ed and post-merger prices can be predicted. Capps, Dranove, and Satterthwaite (2003) estimate α with an OLS regression of total hospital profits on the unscaled WTP and use the above equation to predict the increase in total profits due to the merger. However, we are interested in the predicted price changes due to the merger. As is common in the MSM literature, we assume that the variable costs per patient do not change due to the merger (i.e. $c_{j+k}^{post} = c_{j+k}^{pre}$) and we can therefore rewrite the latter equation as:

$$(3) \quad (p_{j+k}^{post} - p_{j+k}^{pre}) = \alpha \cdot (w_{j+k}^{post} - w_{j+k}^{pre})$$

Following Balan and Brand (2013), we divide the merged entity's increase in WTP into a per-hospital WTP increase. To this end, we have to determine the post-merger WTP of hospital j and k : w_j^{post} and w_k^{post} . We do this by using two assumptions. The first assumption stipulates that the increase in the joint WTP is divided between the two hospitals according to their revenue share in the merged entity: $w_{j+k}^{post} - w_{j+k}^{pre} = S_j(w_j^{post} - w_j^{pre}) + S_k(w_k^{post} - w_k^{pre})$. But this equation does not yet identify a unique pair (w_j^{post}, w_k^{post}) , since there is an infinite number of combinations that satisfies this assumption. The second assumption therefore stipulates that the increase in the hospitals' WTP is divided in proportion to their diversion ratios:

$$(w_j^{post} - w_j^{pre}) = \frac{D_{jk}}{D_{jk}}(w_k^{post} - w_k^{pre}),$$

where diversion ratio D_{jk} is the share of patients from hospital j that would go to hospital k if hospital j were no longer accessible to them⁵². From the IIA property of the conditional logit model it follows that if patient i can no longer visit hospital j , the diversion of hospital j to hospital k for patient i is equal to $\frac{s_k(H_i, X_i, \lambda_i)}{1 - s_j(H_i, X_i, \lambda_i)}$ (see for example Conlon and Mortimer 2013). We calculated the weighted average diversion of hospital j to hospital k (D_{jk}) by summing over all patients and weighting each patient with their predicted share in hospital j :

$$D_{jk} = \sum_i \left(\frac{s_j(H_i, X_i, \lambda_i)}{\sum_i s_j(H_i, X_i, \lambda_i)} \frac{s_k(H_i, X_i, \lambda_i)}{1 - s_j(H_i, X_i, \lambda_i)} \right).$$

Similarly, diversion ratio D_{kj} is the share of patients from hospital k that would go to hospital j if hospital k were no longer accessible to them. Together, the above assumptions can identify the unique pair (w_j^{post}, w_k^{post}) of hospital specific post-merger WTPs. The hospital-specific increase in WTP for hospitals j and k are $(w_j^{post} - w_j^{pre})$ and $(w_k^{post} - w_k^{pre})$ respectively.

52 The intuition behind this assumption is that the hospital for which the diversion ratio is relatively high, can profit more from the merger.

Following equation (3), the hospital-specific price increase for hospital j is then given by:

$$(4) \quad (p_j^{post} - p_j^{pre}) = \alpha \cdot (w_j^{post} - w_j^{pre}).$$

Similarly, the hospital-specific price increase for hospital k is given by:

$$(5) \quad (p_k^{post} - p_k^{pre}) = \alpha \cdot (w_k^{post} - w_k^{pre}).$$

In the following, we will use equations (4) and (5) to predict the price increases resulting from the merger that we examined in this paper.

4. ESTIMATION

A. Specification of our choice model

Following Capps, Dranove, and Satterthwaite (2003), we first estimated a conditional logit model (see section 3A). Unlike Capps, Dranove, and Satterthwaite (2003), however, we ran the model for each of the products separately (rather than aggregating all the products for each hospital). We used the following specification for patient utility:

$$(6) \quad U_{ij} = \sum_j^{J-1} \alpha_j \cdot D_j + \beta_1 \cdot TRAVELTIME + \beta_2 \cdot TRAVELTIME \cdot D_{AGE} + \beta_3 \cdot TRAVELTIME \cdot D_{FEMALE} + \beta_4 \cdot TRAVELTIME \cdot SESSCORE + \varepsilon_{ij}$$

where *TRAVELTIME* was the travel time in minutes from the patient's home (zip code) to the hospitals, D_{AGE} was a dummy indicating whether the patient is older or younger than 65, D_{FEMALE} was a dummy for the patient's gender and *SESSCORE* was a socio-economic status (SES) score for the patient's zip code. We estimated a fixed-effects conditional logit model. Given that there were J hospitals, the dummy variables in this model would pick up J different sets of undefined attributes (e.g. Farrell et al. 2011; Train 2009). In our data we observed that 99% of the patients will not travel more than 100 minutes for a hip or a knee replacement or cataract surgery. We therefore restricted the choice set of each patient to the hospitals reachable within 100 minutes⁵³. For cataract surgery, we only estimated the conditional logit model for the patient's *first* cataract surgery. Out of all patients, 30% received more than one treatment at the same hospital. It is likely that a patient who received more than one cataract treatment at the same hospital was

53 As a robustness check several alternative patient choice sets were used. Our results are robust to these other assumptions. The results are available from the authors upon request.

treated for both the left and right eyes. In the estimation of the choice model (and the calculation of the WTP), we excluded such repeat choices by the same patient.

B. Specification of our WTP regression

For each product, we used the predicted probabilities that followed from the conditional logit estimation to calculate the WTP for the inclusion of each of the hospitals in the network using equation (1). From the estimated conditional logit (equation (6)), we calculated the per-patient probability for choosing a certain hospital. Patient type i chooses hospital j with probability:

$$\hat{s}_j(H_j, X_{ij}, \lambda_i) = \frac{\exp[\hat{U}(H_j, X_{ij}, \lambda_i)]}{\sum_{g \in G} \exp[\hat{U}(H_g, X_{ig}, \lambda_i)]}$$

We use these probabilities and equation (1) to calculate the WTP for each hospital:

$\hat{w}_j = \frac{\sum_{i=1}^N \ln\left[\frac{1}{1 - \hat{s}_j(H_j, X_{ij}, \lambda_i)}\right]}{\sum_{i=1}^N \hat{s}_j(H_j, X_{ij}, \lambda_i)}$. The calculations were performed in R with the package Merger-Analysis (Halbersma 2013).

The next step was to regress the predicted WTPs on the prices negotiated between hospitals and insurers for hip and knee replacements and cataract surgery. We estimated the following model⁵⁴:

$$(7) \quad PRICE_j = c + \alpha \cdot WTP_j + \beta_1 \cdot INSURER.HHI_j + \beta_2 \cdot SESSCORE_j + \beta_3 \cdot AGE_j + \beta_4 \cdot HOUSEPRICE_j + \beta_5 \cdot HOSPITAL.TYPE_j + \beta_6 \cdot HOSPITAL.SIZE_j + \beta_7 \cdot LIBERALIZED_j + \varepsilon_j$$

where $PRICE$ was the average pre-merger price (per hospital per product), WTP was the WTP following from equation (1) and based on the probabilities from the fixed-effects conditional logit model (equation (6)) (i.e. \hat{w}_j), $INSURER.HHI$ is the insurer's Herfindahl-Hirschmann Index (HHI) for each hospital (based on the insurer's market shares of the total revenue of the hospital, per product). To control for potential differences in hospital costs, we included the average SES-score of the patients ($SESSCORE$) and the average age of the patients (AGE) as proxies for hospitals' casemix differences, the average house price of the hospital's zip code (divided by 100.000) as a proxy for location-specific costs ($HOUSEPRICE$), the hospital type (academic or general hospital⁵⁵) ($HOSPITAL.TYPE$), and the hospitals' size, measured in terms of the total number of beds ($HOSPITAL.SIZE$) to account for potential (dis)economies of scale. Further, we control for the per hospital frac-

54 We examined the robustness of the model by estimating the Huber M-estimator (Huber 1964) and the least trimmed squares (Its) regression (Rousseeuw and Van Driessen 1999). Both methods produced similar results. The results are available from the authors upon request.

55 Due to their low number, we did not distinguish Independent Treatment Centres (ITCs – see also footnote 24) or specialty hospitals separately in this analysis. They were treated as general hospitals.

tion of the liberalized segment (defined by the revenue of the total liberalized segment divided by the total revenue of the hospital) (*LIBERALIZED*). We report the MacKinnon and White (1985) Heteroskedasticity-Consistent standard errors.

C. Instrumental variable approach

It is possible that our predicted WTP is endogenous. There are two important sources of endogeneity. First, performance may feed back into structure, causing a simultaneous equation bias (e.g. lower prices may induce patients to go to a cheaper hospital, which in turn increases the (predicted) WTP of the hospital as derived from observed patient choices). Second, there are attributes that influence both price and patients' choice of a hospital (e.g. quality of care). These are picked up by the conditional logit model's fixed effects, causing an omitted variables bias (see also Evans, Froeb, and Werden 1993).

The common solution to these problems is to use an instrumental variables (IV) approach. Kessler and McClellan (2000), Cooper et al. (2011) and Gaynor, Morena-Serra, and Propper (2013) solve the endogeneity problem by using the predicted patient flows generated from models of patient choice. These only use observable, exogenous characteristics of patients and hospitals (Kessler and McClellan 2000). In our paper, we estimate a WTP instrument (*TRAVELTIME-WTP*) which is based on the predicted probabilities of a conditional logit model that only includes patients' travel times ($U_{ij} = \beta_1 \cdot \text{TRAVELTIME} + \varepsilon_{ij}$). Following Kessler and McClellan (2000) and Gaynor, Morena-Serra, and Propper (2013), we explicitly omit hospital-level fixed effects to prevent predicted choice being based on unobserved attributes of prices.

After determining the *TRAVELTIME-WTP*, we carried out a two-stage least square (2SLS) model where the instrument list consisted of *TRAVELTIME-WTP* (instrument for WTP), *INSURER.HHI*, *SESSCORE*, *AGE*, *HOUSEPRICE*, *HOSPITAL.TYPE*, *HOSPITAL.SIZE*, and *LIBERALIZED* (see section 5 for details on these variables).

5. Data

In this paper, we analyze the price effects of a merger between a general acute care hospital (hospital M1) and a neighboring general acute care hospital that also provides tertiary hospital care (hospital M2). The merger was consummated in the Netherlands in year t . We used pre-merger data ($t-1$ data) to establish what price increases the Option Demand method would have predicted if an antitrust authority had used the model in their review after being notified of the merger. We contrast the predictions obtained using the OD method with the actual price effects of the merger. The latter are determined through a difference-in-difference technique (Roos et al. 2017). In section 3 we explained that we focus on three products for which prices are freely negotiable: hip replacements, knee replacements and cataract surgery. In year $t-1$ these product

markets jointly represent 47.5 percent of the merged hospital's turnover in the segment for which Dutch health insurers and hospitals were allowed to freely negotiate prices.

We use a nationwide patient-level dataset that contains all inpatient and outpatient visits for all hospital locations and Independent Treatment Centers (ITCs)⁵⁶. For each visit, the patient's zip code, age (year of birth), gender, health insurer, diagnosis and treatment were observed, as well as the price negotiated for each hospital location-insurer-product combination in year $t-1$. The patient-level data that we used came from the insurers' claims administration and hospital registries and was provided by the Dutch Healthcare Authority.

For the choice model (see section 4A), we calculated each patient's travel time (in minutes) to the hospitals using a travel time matrix for year $t-1$. Some hospitals have multiple treatment locations, but the data does not reflect which location the patient actually went to. For hospitals with more than one treatment location, we calculated the patient's travel time (in minutes) to the closest hospital location⁵⁷. Additionally, we obtained socio-economics status (SES) scores from the Netherlands Institute for Social Research (SCP). A higher SES score means a higher socio-economic status in the zip code area.

In the WTP regression (see section 4B), we included the average SES score and the average age. Additionally, we included the average house price for the zip code area of the hospital and the hospital types as proxies for location-specific costs. The data on house prices was obtained from Statistics Netherlands (CBS). We differentiated between academic and general hospitals (taking general hospitals as the reference group). ITCs and specialty hospitals were treated as general hospitals. The insurer's HHI was based on the insurer's market shares per product (of the total revenue of the hospital) and ranged from zero to one. Thus, the insurer's HHI for hospital j and product k was calculated as: $INSURER.HHI_{jk} = \sum_{l=1}^n \left(\frac{REV_{jkl}}{\sum_{l=1}^n REV_{jkl}} \right)^2$, where REV_{jkl} is the revenue of insurer l ($l=1, \dots, n$) in hospital j for product k . We also included the per-hospital fraction of the liberalized segment (*LIBERALIZED*), which was defined by the revenue of the whole liberalized segment divided by the total revenue of a hospital (i.e. the regulated and liberalized segments together).

56 ITCs are comparable to the freestanding Ambulatory Surgery Centers (ASCs) that operate in the US and UK healthcare markets.

57 25 hospitals had multiple locations for hip replacements and cataract surgeries. For knee replacements 27 hospitals had multiple locations. As a sensitivity check we also estimated the choice model using the patient's travel time (in minutes) to the main hospital location. This did not affect our WTP estimations. The results are available from the authors upon request.

6. RESULTS

A. Choice model

Table 1 presents summary statistics on the main variables that were included in the conditional logit model of patients' choice of hospital for hip and knee replacement and cataract surgery (panels A).

Table 2 presents the results of our estimation. We estimated two models for each product (hip, knee and cataract). Model 2 is the conditional logit model that includes patients' travel time only (see section 4C), while model 1 is the full fixed effects conditional logit model that also includes other covariates (see section 4A). The results of model 2 clearly show that, as expected, patients dislike travel time. Model 1 also takes patient heterogeneity into account by adding interaction terms. The results show that travel time interacts with age, gender and SES score, indicating that older patients prefer hospitals closer to home than younger patients and that females are less willing to travel further than men, while the higher the SES score, the greater the patients' willingness to travel. All coefficients have the expected sign and correspond with findings from other studies on patient choice in the Netherlands (e.g. Beukers, Kemp, and Varkevisser 2014; Varkevisser, Van der Geest, and Schut 2012; Varkevisser, Van der Geest, and Schut 2010). Furthermore, the goodness of fit measures that are also presented in table 2 show that our models perform well.

B. WTP regression

As discussed in section 4B, we used the estimated coefficients from the conditional logit models to calculate the Willingness-To-Pay for the inclusion of each of the hospitals in the network. We then regressed the predicted WTPs on the observed prices. Equation (7) is an OLS regression model that we estimated with and without instrumental variables (see sections 4B and 4C). Table 1 presents summary statistics on the main variables that were included in the OLS regressions (panels B).

The results of the estimation can be found in table 3. The first model is a simple ordinary least squares model with the WTP and the insurers market power vis-à-vis each individual hospital (measured by the HHI) regressed on price; model 2 adds control variables to model 1; and model 3 is a 2SLS approach with control and instrumental variables. As discussed in section 4C, we use *TRAVELTIME-WTP* as an instrument for the WTP. To determine the relevance of the instrument, we tested its correlation with the possibly endogenous regressor WTP by determining the first-stage F-statistic (Stock and Yogo 2005; Staiger and Stock 1997). Our first-stage F-statistic was 62.617 (p -value = 0.00) for hip replacement, 39.549 (p -value = 0.00) for knee replacement and 181.51 (p -value = 0.00) for cataract surgery. This indicates that our instrument (*TRAVELTIME-WTP*) is strongly correlated with the WTP. The Wu-Hausman statistic was 0.16 (p -value = 0.68) for

Table 1. Descriptive statistics

Variable	Mean	Standard deviation	Minimum	Maximum	Observations
Hip replacements					
<i>Panel A. Patient characteristics</i>					
Age	69.2	10.5	16	99	<i>N</i> = 20846
Age Dummy (>65)	0.66	-	0	1	<i>N</i> = 20846
Gender (female)	0.68	-	0	1	<i>N</i> = 20846
SES score in the zip code area	-0.002	1.000	-5.437	3.813	<i>N</i> = 20846
Travel time (in minutes)	12.60	13.15	0.00	99.96	<i>N</i> = 20846
<i>Panel B. Hospital characteristics</i>					
Patients' average age	69.0	2.7	55.1	0.6	<i>n</i> = 82
Patients' average SES score	-0.023	0.361	-0.909	0.639	<i>n</i> = 82
Price hip replacement (in €)	9092.00	293.29	8527.00	10408.00	<i>n</i> = 82
Willingness-To-Pay	1.813	0.885	1.024	7.234	<i>n</i> = 82
Instrument Willingness-To-Pay (TRAVELTIME-WTP)	1.666	0.676	1.056	5.177	<i>n</i> = 82
Academic hospital	0.09	-	0	1	<i>n</i> = 82
ITC	0	-	0	1	<i>n</i> = 82
Insurers' HHI	0.391	0.134	0.163	0.795	<i>n</i> = 82
Housing price in the zip code area (€1000)	193.9	32.9	134.0	266.0	<i>n</i> = 82
Hospital size (number of beds)	512.7	275.0	138.0	1575.0	<i>n</i> = 82
The hospital's share of the liberalized segment (LIBERALIZED)	0.11	0.04	0.02	0.23	<i>n</i> = 82
Knee replacements					
<i>Panel A. Patient characteristics</i>					
Age	69.0	9.9	20	97	<i>N</i> = 17558
Age Dummy (>65)	0.65	-	0	1	<i>N</i> = 17558
Gender (female)	0.69	-	0	1	<i>N</i> = 17558
SES score in the zip code area	-0.002	1.001	-5.148	2.772	<i>N</i> = 17558
Travel time (in minutes)	13.25	14.15	0	99.71	<i>N</i> = 17558
<i>Panel B. Hospital characteristics</i>					
Patients' average age	69.0	2.0	64.1	74.6	<i>n</i> = 85
Patients' average SES score	0.009	0.357	-0.869	0.791	<i>n</i> = 85
Price knee replacement (in €)	11493.00	390.69	9756.00	10689.00	<i>n</i> = 85
Willingness-To-Pay	1.712	0.795	1.019	6.628	<i>n</i> = 85
Instrument Willingness-To-Pay (TRAVELTIME-WTP)	1.579	0.589	1.045	4.576	<i>n</i> = 85
Academic hospital	0.09	-	0	1	<i>n</i> = 85
ITC					<i>n</i> = 85
Insurers' HHI	0.408	0.127	0.618	0.783	<i>n</i> = 85
Housing price in the zip code area (€1000)	194.1	32.2	134.0	266.0	<i>n</i> = 85
Hospital size (number of beds)	509.3	272.3	140.0	1575.0	<i>n</i> = 85

Variable	Mean	Standard deviation	Minimum	Maximum	Observations
The hospital's share of the liberalized segment (LIBERALIZED)	0.11	0.03	0.02	0.23	$n = 85$
Cataract surgery					
<i>Panel A. Patient characteristics</i>					
Age	73.5	10.4	0.0	110.0	$N = 103750$
Age Dummy (>65)	0.81	-	0	1	$N = 103750$
Gender (female)	0.61	-	0	1	$N = 103750$
SES score in the zip code area	0.000	1.000	-6.171	2.809	$N = 103750$
Travel time (in minutes)	11.30	11.35	0.00	99.99	$N = 103750$
<i>Panel B. Hospital characteristics</i>					
Patients' average age	73.3	2.3	63.5	77.0	$n = 86$
Patients' average SES score	-0.031	0.385	-1.148	0.627	$n = 86$
Price cataract surgery (in €)	1365.00	83.80	1046.00	1547.00	$n = 86$
Willingness-To-Pay	1.875	0.846	1.018	5.795	$n = 86$
Instrument Willingness-To-Pay (TRAVELTIME-WTP)	1.782	0.805	1.056	5.766	$n = 86$
Academic hospital	0.08	-	0	1	$n = 86$
Insurers' HHI	0.421	0.128	0.210	0.694	$n = 86$
Housing price in the zip code area (€1000)	192.5	33.4	134.0	284.0	$n = 86$
Hospital size (number of beds)	488.3	289.35	0.0	1575.0	$n = 86$
The hospital's share of the liberalized segment (LIBERALIZED)	0.12	0.06	0.02	0.44	$n = 86$

Notes: Summary statistics refer to $t-1$, where t is the merger year. N = total number of patients that underwent hip or knee replacements or cataract surgeries. The total number of patients that underwent cataract surgery only includes the patient's first cataract surgery. n = total number of hospitals in the sample. We calculated the patient's travel time (in minutes) to the closest hospital location.

hip replacement, 2.39 (p -value = 0.13) for knee replacement and 0.24 (p -value = 0.63) for cataract surgery. This indicates that the variable WTP is not endogenous.

The average price for hip replacements is €9,092 (table 1). Table 3 indicates that a one-unit increase in WTP will increase prices for hip replacements by €88.69 (model 2). Following Capps, Dranove, and Satterthwaite (2003), we show how to interpret the magnitude of this estimate by considering the hospital with the highest WTP (i.e. WTP: 7.234 –table 1) and the hospital with the lowest WTP (i.e. WTP: 1.024 –table 1). Using the results of model 2, the WTP difference of 6.210 translates into a difference in the price of a hip replacement of €550.76.

For knee replacements and cataract surgeries, the WTP is not significantly related to price. Apparently, in our regression model, WTP does not explain the variation in prices for knee replacement and cataract surgeries. This means that in the market for knee

Table 2. Conditional logit model of patient hospital choice for hip and knee replacements and cataract surgery^a

	Hip replacements		Knee replacements		Cataract surgery	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Time	-0.1266*** (0.0018)	-0.1488*** (0.0010)	-0.1248*** (0.0018)	-0.1504*** (0.0011)	-0.1343*** (0.0010)	-0.1668*** (0.0005)
Time * Age	-0.0590*** (0.0020)		-0.0460*** (0.0019)		-0.0551*** (0.0011)	
Time * Female	-0.0031 (0.0020)		-0.0025 (0.0020)		-0.0088*** (0.0010)	
Time * SES-score	0.0085*** (0.0011)		0.0104*** (0.0011)		0.0072*** (0.0006)	
Observations	20846	20846	17558	17558	103750	103750
McFadden's R ²	0.68	0.64	0.65	0.61	0.72	0.67
Hit-and-miss	0.70	0.69	0.67	0.66	0.70	0.68

Notes: Models estimated by conditional logit model with standard errors in parentheses under coefficients. Model 2 is the conditional logit model that only includes patients' travel time (see section 4C), while model 1 includes a full set of hospital dummies (not reported here) and other covariates (see section 4A). The conditional logit models are estimated on data from $t-1$, where t is the merger year. We restricted the patients' choice sets to the hospitals reachable within 100 minutes.

^a For clarity reasons, we do not report the hospital dummies (fixed effects) here. The results are available from the authors upon request.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

replacements and cataract surgeries using the WTP for the WTP-based merger simulation is less meaningful than in the market for hip replacements.

7. USING THE WTP FOR ANTITRUST PURPOSES

In this section, we contrast the *ex ante* predicted price effects with the actual *ex post* price effects of a merger between a general acute care hospital (hospital M1) and a neighboring general acute care hospital that also provides some types of tertiary hospital care (hospital M2).

The actual price effects were determined through a difference-in-differences technique (Roos et al. 2017). For a detailed discussion of the method, data and results of the difference-in-differences technique, we refer to Roos et al. (2017). In sum, Roos et al. (2017) use data on hospital-insurer negotiated contract prices in the Netherlands for each of the three hospital products considered, to investigate whether the merger between hospitals M1 and M2 has led to price changes. They first estimate an aggregated difference-in-differences model (
$$p_{it} = \alpha + \sum_{j=1}^J \beta_j D_{jt} + \lambda \cdot D_{i+1} + \delta \cdot D_{i+1} \cdot D_{merged} + \varepsilon_{it}$$
)

Table 3. Willingness-To-Pay models for hip and knee replacements and cataract surgery

	Model 1	Model 2	Model 3
Hip replacements			
(intercept)	9238.21*** (100.09)	8075.17*** (1515.84)	8027.90*** (1519.47)
WTP	76.21** (33.28)	88.69** (40.00)	94.40** (42.44)
INSURER.HHI	-727.75** (294.38)	-894.08*** (324.63)	-909.52*** (335.73)
SESSCORE		-91.61 (101.93)	-95.51 (104.86)
AGE		16.81 (20.00)	17.30 (20.11)
HOUSEPRICE		-0.24 (0.98)	-0.18 (0.98)
HOSPITAL.TYPE		151.25 (220.69)	154.75 (220.79)
HOSPITAL.SIZE		0.09 (0.19)	0.09 (0.19)
LIBERALIZED		279.38 (2298.29)	256.06 (2295.96)
IV	NO	NO	YES: TRAVELTIME-WTP
Observations	82	82	82
R-Squared	0.10	0.13	0.13
Adjusted R-Squared	0.08	0.04	0.04
Knee replacements			
(intercept)	10857.12*** (170.90)	10805.00*** (1947.40)	10619.25*** (1917.42)
WTP	14.38 (105.15)	3.00 (125.09)	37.90 (122.47)
INSURER.HHI	-473.89 (433.23)	-5381.80 (450.97)	-613.11 (447.91)
SESSCORE		21.99 (136.89)	-3.59 (138.74)
AGE		8.25 (26.70)	9.88 (26.27)
HOUSEPRICE		-2.12 (1.63)	-1.76 (1.62)
HOSPITAL.TYPE		153.74 (277.65)	156.39 (282.73)
HOSPITAL.SIZE		-0.06 (0.21)	-0.06 (0.21)
LIBERALIZED		-394.23 (2938.10)	-605.29 (2903.94)
IV	NO	NO	YES: TRAVELTIME-WTP
Observations	85	85	85
R-Squared	0.02	0.06	0.06
Adjusted R-Squared	-0.00	-0.04	-0.04
Cataract surgery			
(intercept)	1319.70*** (35.50)	803.25 (899.40)	295.27 (899.51)
WTP	1.26 (9.70)	-2.17 (10.79)	-5.94 (10.77)
INSURER.HHI	100.78 (71.39)	33.65 (89.72)	47.49 (89.22)
SESSCORE		-19.57 (32.78)	-17.03 (33.09)
AGE		10.12 (11.09)	10.30 (11.09)
HOUSEPRICE		-0.34 (0.36)	-0.36 (0.36)
HOSPITAL.TYPE		74.51 (83.95)	75.24 (83.43)
HOSPITAL.SIZE		-0.07 (0.05)	-0.07 (0.05)
LIBERALIZED		-828.47 (331.48)	-822.98** (338.05)

	Model 1	Model 2	Model 3
IV	NO	NO	YES: TRAVELTIME-WTP
Observations	86	86	86
R-Squared	0.03	0.41	0.41
Adjusted R-Squared	0.00	0.35	0.35

Notes: Per product we report three models. The first model is a simple OLS model with the WTP and the insurers' market power vis-à-vis each individual hospital regressed on price. Model 2 adds control variables to model 1 and model 3 is a 2SLS model that adds control and instrumental variables. We report the Mackinnon and White (1985) Heteroskedasticity-Consistent standard errors (in parentheses under coefficients). We used data from $t-1$, where t is the merger year. Note that the R^2 for cataract surgeries is much higher than the R^2 for hip and knee replacements. This is due to a higher number of ITCs in the market for cataract surgeries.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

and then, to show the effect of aggregating over products, locations and insurers, they remove the aggregations stepwise. The pre-merger price was based on data from the year preceding the merger ($t-1$) and the post-merger price was based on data from the year after the merger ($t+1$). Table 4 summarizes the estimated merger effects on prices of hip replacement, knee replacement and cataract surgery for hospitals M1 and M2 in comparison with the average price change pre- and post-merger in a control group. Roos et al. (2017) find evidence of heterogeneous price effects for a merger between neighboring hospitals across hospital products and hospital locations. Their result is robust for different control groups and different model specifications.

The *ex ante* predictions for merger-induced price increases were calculated using the Option Demand method as described in section 3C. Table 5 displays the predicted WTP increases. To see how a merger affects WTP, we looked at the change in the predicted WTPs. In the case of hip replacements, the WTP for hospital M1 increased by 25.7%, and the WTP for hospital M2 increased by 11.7%. Both of these increases were substantial. In general, patients are more willing to pay for the inclusion of hospital M1 than hospital M2. This is not surprising because the merger also had a differential impact on the structure of the market in which the hospitals were competing. Hospital M1 is located in an isolated geographical area and hospital M2 was the largest competitor to hospital M1 pre-merger. Hospital M2, in contrast, is subject to notable competitive pressure from (at least) five other hospitals in the three submarkets studied in this paper. Note that table 4 suggests that the price increases are higher for hospital M1 than for hospital M2, a finding which is only statistically significant for hip replacements.

Next, the increase in the hospital specific prices due to merger can be determined using equations (4) and (5). From equations (4) and (5) it follows that we were able to calculate the predicted increase in prices for hospital j as:

Table 4. Merger effect on prices of hip and knee replacements and cataract surgery for hospitals M1 and M2 in comparison to average price changes pre- and post-merger in a control group (retrospective analysis)

	Merger effect on price (DiD coefficient)
<i>Panel A. Hospital M1</i>	
Hip replacements	0.090* (0.053)
Knee replacements	0.021 (0.062)
Cataract surgery	0.027 (0.057)
<i>Panel B. Hospital M2</i>	
Hip replacements	-0.035 (0.053)
Knee replacements	-0.064 (0.062)
Cataract surgery	-0.049 (0.057)

Notes: Time period is $t-2$ and $t+2$, where t is the merger. Models estimated by OLS with standard errors in parentheses under coefficients. Null hypothesis: difference-in-difference estimator is equal to zero.

Source: Roos et al. 2017.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

$$(8) \quad \frac{\hat{a}(w_j^{\text{post}} - w_j^{\text{pre}})}{\widehat{\text{PRICE}}_j},$$

where $\widehat{\text{PRICE}}_j$ is the fitted pre-merger price of hospital j and \hat{a} is the estimated coefficient of the WTP that is obtained by equation (7). As discussed above, we only estimate the predicted price increases for hip replacements. Table 6 compares the results with the *ex post* estimates.

We constructed 90% and 95% confidence intervals for the predicted and estimated price increases using the student t distribution of \hat{a} and treatment effect, respectively. Care should be taken in interpreting the results as the *ex post* estimates have large confidence intervals.

The merger simulation showed that the prices for hip replacements in hospitals M1 and M2 were likely to increase significantly, although at a different magnitude. The confidence intervals of the predicted price increases are all nested within the confidence intervals of the actual price increases. Given that the confidence intervals of the *ex post* estimates are quite large, however, we should be cautious in interpreting this result as evidence that the OD method is able to accurately predict price increases after merger. If we were to ignore this for a moment, because Roos et al. (2017) showed that the *ex post* estimation was robust for different control groups and different model specifications, table 6 suggest that OD method overestimates the price effects for hospital M2 and underestimates the price effects for hospital M1.

Table 5. Hospital specific change in WTP after merger for hip and knee replacements and cataract surgery

	Pre-merger price (mean in €)	Pre-merger WTP	Absolute increase in WTP after merger
<i>Panel A. Hospital M1</i>			
Hip replacements	9135.66	4.668	1.199
Knee replacements	10693.64	4.706	0.994
Cataract surgery	1393.25	3.655	1.530
<i>Panel B. Hospital M2</i>			
Hip replacements	9064.99	2.296	0.268
Knee replacements	10645.73	2.021	0.322
Cataract surgery	1358.92	2.500	0.157

Notes: For the pre-merger WTP, we used data from $t-1$, where t is the merger year (WTP estimation using the results from table 3, model 2). For the change in WTP after merger, we used data from $t-1$ and $t+1$, where t is the merger year (change in WTP using the results from table 3, model 2).

Table 6. Predicted and estimated price increases for hip replacements due to merger

	<i>Ex ante</i> predictions (by the Option Demand Method)			<i>Ex post</i> estimates (by the difference-in- differences estimates)		
	% price increases	95% CI	90% CI	% price increase	95% CI	90% CI
<i>Panel A. Hospital M1</i>						
Hip replacements	1.16	[0.12 – 2.21]	[0.29 – 2.04]	9.00	[-1.63 – 19.63]	[0.13 – 17.87]
<i>Panel B. Hospital M2</i>						
Hip replacements	0.26	[0.03 – 0.50]	[0.07 – 0.46]	-3.50	[-14.13 – 7.13]	[-12.37 – 5.37]

Notes: The increases in the hospital specific prices due to merger are determined using equations 4 and 5. The *ex post* estimates are obtained using a difference-in-differences technique, which is reported in Roos et al. (2017).

8. CONCLUSION

The aim of this paper is to examine the predictive power of the option demand (OD) method for hospital mergers. Like other merger simulation models (MSMs), the OD method has clear advantages over more traditional market definition approaches because it provides antitrust agencies with direct evidence about the expected effects of the merger and does not require questionable assumptions to be made on the relevant (geographic) market. Also, studies that contrasted the predictions by the OD method and several traditional measures concluded that the OD model outperforms ad hoc measures in predicting prices (Garmon 2016; Dranove & Ody 2016).

Antitrust agencies should aim to use MSMs that are able to explain outcomes in the relevant market reasonable well, for example by demonstrating that the model accu-

rately predicts the effects of mergers in the same industry (Budzinsky and Ruhmer 2009; Werden, Froeb, and Scheffman 2004). We have contrasted the findings of this prospective method of analysis with the findings of a retrospective study involving a consummated Dutch hospital merger (Roos et al. 2017). Our results indicate that there is a relationship between WTP and prices for hip replacements. We were not able to establish a relationship between WTP and prices for knee replacements and cataract surgeries. We therefore only estimated a reduced-form merger simulation for hip replacements. The comparison between the reduced-form merger simulation and *ex post* estimates suggest that the OD method overestimates the price effects for hospital M2 and underestimates the price effects for hospital M1. Yet, the overestimation is not statistically significant.

Garmon (2016) also finds mixed results for the performance of the reduced-form merger simulation in the US. Hence, we conclude that although the OD method could be a valuable addition to the antitrust agencies' toolkit in signaling potentially anti-competitive merger effects, our findings also indicate that more research is necessary. For example, the explanatory power of our regression models is quite low. This may either indicate that the model needs to be reconsidered to find factors that have higher explanatory power or that the model does not (yet) fit the Dutch healthcare market well enough. With respect to the latter, we concluded in section 2B that the OD method is applicable to the free-pricing segment of the Dutch hospital industry. However, the industry is in transition and the number of health insurers offering contracts with restricted provider networks has increased over the years. As the OD method depends upon the bargaining relationship between health insurers and hospitals, we expect that the relationship between WTP and price will get stronger as the threat of selective contracting becomes more credible.

REFERENCES

- Argue, David A., and Richard T. Shin. 2009. "An Innovative Approach to an Old Problem: Hospital Merger Simulation" *Antitrust* 24 (1): 49-54.
- Baker, Jonathan B. 2003. "Why did the antitrust agencies embrace unilateral effects?" *George Mason Law Review* 12 (1): 31-37.
- Balan, David J., and Keith Brand. 2013. "Simulating hospital merger simulations". *Working paper*.
- Balan, David J., and Keith Brand. 2015. "Simulating hospital merger simulations". *Unpublished Draft*. Available at https://works.bepress.com/david_balan/7/
- Brand, Keith, and Christopher Garmon. 2014. *Hospital Merger Simulation*. Member Briefing Antitrust Practice Group: American Health Lawyers Association.
- Beukers, Puck D.C., Ron G.M. Kemp, and Marco Varkevisser. 2014. "Patient hospital choice for hip replacement: empirical evidence from the Netherlands" *European Journal of Health Economics* 15(9): 927-936.
- Björnerstedt, Jonas, and Frank Verboven. 2016. "Does merger simulation work? Evidence from the Swedish analgesics market" *American Economic Journal: Applied Economics* 8 (3): 125-164.
- Budzinski, Oliver, and Isabel Ruhmer. 2009. "Merger Simulation in Competition Policy: a Survey" *Journal of Competition Law and Economics* 6 (2): 277-319.
- Capps, Cory, David Dranove, and Mark Satterthwaite. 2003. "Competition and market power in option demand markets" *The RAND journal of economics* 34 (4): 737-763.
- Capps, Cory, David Dranove, Shane Greenstein, and Mark Satterthwaite. 2002. "Antitrust policy and hospital mergers: recommendations for a new approach" *The Antitrust Bulletin* 47: 677-714.
- Capps, Cory, David Dranove, Shane Greenstein, and Mark Satterthwaite. 2001. *The silent majority fallacy of the Elzinga-Hogarty criteria: a critique and new approach to analyzing hospital mergers*. "NBER Working Paper 8216". Cambridge: National Bureau of Economic Research.
- Chandra, Amitabh, Amy Finkelstein, Adam Sacarny, and Chad Syverson. 2016. "Health Care Exceptionalism? Performance and Allocation in the US Health Care Sector" *American Economic Review* 106 (8): 2110-2144.
- Chou, Shin-Yi, Mary .E. Deily, Suhui. Li, and Yi Lu. 2014. "Competition and the impact of online hospital report cards" *Journal of Health Economics* 34 (1): 42-58.
- Cooper, Zack, Stephen Gibbons, Simon Jones, and Alistair McGuire. 2011. "Does hospital competition save lives? Evidence from the English NHS patient choice reforms" *The Economic Journal* 121: 228-260.
- Conlon, Christopher, and Julie Mortimer. 2013. *An Experimental Approach to Merger Evaluation*. "NBER working paper 19703". Cambridge: National Bureau of Economic Research.
- Davis, Peter J., and Eliana Garcés. 2010. *Chapter 8: Merger Simulation*. In: "Quantitative techniques for competition and antitrust analysis". (eds. Davis. P.J. & E. Garcés). Princeton: Princeton University Press. 382-434.
- Dranove, David, and Christopher Ody. 2016. "Evolving measures of provider market power" *American Journal of Health Economics* 2 (2): 145-160.
- Dranove, David, and Andrew Sfekas. 2009. "The Revolution in Health Care Antitrust: New Methods and Provocative Implications" *The Milbank Quarterly* 87 (3): 607-632.
- Dranove, David, and William D. White. 1996. "Specialization, Option Demand and the Pricing of Medical Specialists" *Journal of Economics and Management Strategy* 5: 277-306.
- Elzinga, Kenneth G., and Anthony W. Swisher. 2011. "Limits of the Elzinga-Hogarty Test in Hospital Mergers: The Evanston Case" *International Journal of the Economics of Business* 18 (1): 133-146.

- European Commission (EC). 2004. *Guidelines on the assessment of horizontal mergers under the Council Regulation on the control of concentrations between undertakings. (2004/C 31/03)*. "Official Journal of the European Union" C 31: 5-18.
- Evans, William N., Luke M. Froeb, and Gregory J. Werden. 1993. "Endogeneity in the concentration-price relationship: causes, consequences and cures" *The Journal of Industrial Economics* 4: 431-438.
- Farrell, Joseph, David J. Balan, Keith Brand, and Brett W. Wendling. 2011. "Economics at the FTC: Hospital Mergers, Authorized Generic Drugs, and Consumer Credit Markets" *Review of Industrial Organization* 39 (4): 271-296.
- Frank, Matthew B., John Hsu, Mary Beth Landrum, and Michael E. Chernew. 2015. "The impact of a tiered network on hospital choice" *Health Services Research* 50 (5): 1628-1648.
- Friberg, Richard, and André Romahn. 2015. "Divestiture requirements as a tool for competition policy: A case from the Swedish beer market" *International Journal of Industrial Organization* 42: 1- 18.
- Garmon, Christopher. 2016. "The Accuracy of Hospital Merger Screening Methods". *Working paper Federal Trade Commission*. Washington DC: Federal Trade Commission.
- Gaynor, Martin, Kate Ho, and Robert J. Town. 2015. „The Industrial Organization of Health-Care Markets" *Journal of Economic Literature* 53 (2): 235-284.
- Gaynor, Martin, Samuel A. Kleiner., and William B. Vogt. 2011. *A Structural Approach to Market Definition with an Application to the Hospital Industry* "NBER working paper 16656". Cambridge: National Bureau of Economic Research.
- Gaynor, Martin, Carol Propper, and Stephan Seiler. 2016. "Free to Choose? Reform, Choice and Consideration Sets in the English National Health Service" *American Economic Review* 106 (11): 3521-3557.
- Gaynor, Martin, Rodrigo Moreno-Serra, and Carol Propper. 2013. "Death by Market Power: Reform, Competition, and Patient Outcomes in the National Health Service" *American Economic Journal: Economic Policy* 5 (4): 134-66.
- Gaynor, Martin, and Robert J. Town. 2012. *Chapter 9. Competition in Health Care Markets*. In: "Handbook of Health Economics Volume 2". (eds: McGuire, T., M.V. Pauly & P.P. Barros). Amsterdam: Elsevier North-Holland. 499-637.
- Gowrisankaran, G., A. Nevo & R. Town. 2015. "Mergers when prices are negotiated: Evidence from the hospital industry" *American Economic Review* 105 (1): 172-203.
- Greenfield, Daniel, Nicholas Kreisle, and Mark Williams. 2015. *Simulating a Homogeneous Product Merger: A Case Study on Model Fit and Performance* "Federal Trade Commission Working paper 327" Washington: Federal Trade Commission.
- Gutacker, Nils, Luigi Siciliani, Giuseppe Moscelli, and Hugh Gravelle. 2016. "Choice of hospital: Which type of quality matters?" *Journal of Health Economics*. 50: 230-246.
- Halbersma, Rein. 2013. "MergerAnalysis: Merger Analysis package". R package version 0.9.6. <http://www.R-project.org/>
- Ho, Katherine, and Ariel Pakes. 2014. "Hospital choices, hospital prices, and financial incentives to physicians" *American Economic Review* 104 (12): 3841-3884.
- Ho, Katherine. 2009. "Insurer-Provider Networks in the Medical Care Market" *American Economic Review* 99 (1): 393-430.
- Huber, Peter J. 1964. "Robust estimation of a location parameter" *The Annals of Mathematical Statistics* 35 (1): 73-101
- Kalbfleisch, Pieter. 2005. *European Merger Control: A Case of Second Mover Advantage?* In: "Modeling European Mergers: Theory, Competition Policy & Case Studies". (eds. Bergeijk, van, P.A.G. & E. Kloosterhuis). Cheltenham: Edward Elgar Publishing Limited. 27-34.

- Kaplow, Louis. 2011. "Market Definition and the Merger Guidelines" *Review of Industrial Organization* 39 (1-2): 107-125.
- Kessler, Daniel, and Mark McClellan. 2000. "Is hospital competition socially wasteful?" *Quarterly Journal of Economics* 115 (2): 577-615.
- MacKinnon, James G., and Halbert White. 1985. "Some Heteroskedasticity-Consistent Covariance Matrix Estimators with Improved Finite Sample Properties" *Journal of Econometrics* 29: 305-325.
- McFadden, Daniel. 1974. *Conditional logit analysis of qualitative choice behavior*. In: "Frontiers in econometrics" (ed. Zarembka, P.). New York: Academic Press. 105-142.
- Peters, Craig. 2006. "Evaluating the Performance of Merger Simulation: Evidence from the US Airline Industry" *The Journal of Law & Economics* 49 (2): 627-649.
- Pinkse, Joris, and Margaret E. Slade. 2004. "Mergers, brand competition and the price of a pint" *European Economic Review* 48 (3): 617-643.
- Pope, Devin G. 2009. "Reacting to rankings: Evidence from "America's Best Hospitals" *Journal of Health Economics* 28 (6): 1154-1165.
- Roos, Anne-Fleur, Ramsis R. Croes, Victoria Shestalova, Marco Varkevisser, and Frederik T. Schut. 2017. *Price effects of a hospital merger: heterogeneity across health insurers, hospital products and hospital locations*. "Working paper April 2017". Rotterdam: Erasmus University Rotterdam.
- Rousseeuw, Peter J., and Katrien van Driessen. 1999. "A fast algorithm for the minimum covariance determinant estimator" *Technometrics* 41: 212-223.
- Sacher, Seth, and Louis Silvia. 1998. "Antitrust issues in defining the product market for hospital services" *International Journal of the Economics of Business* 5 (2): 181-202.
- Schut, Frederik T., and Wynand P.M.M. van de Ven. 2005. "Rationing and competition in the Dutch health-care system" *Health Economics* 14: S59-S74.
- Schut, Frederik T., and Wynand P.M.M. van de Ven. 2011. "Effects of purchaser competition in the Dutch health system: is the glass half full or half empty?" *Health Economics, Policy and Law* 6: 109-123.
- Shapiro, Carl. 2010. "The 2010 Horizontal Merger Guidelines: From Hedgehog to Fox in Forty Years" *Antitrust Law Journal* 7: 701-759.
- Staiger, Douglas, and James H. Stock. 1997. "Instrumental Variables Regression with Weak Instruments" *Econometrica* 65: 557-586.
- Stock, James H., and Motohiro Yogo. 2005. *Testing for Weak Instruments in Linear IV Regression*. In: "Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg". (eds. Andrews, D.W.K., and J.H. Stock). Cambridge: Cambridge University Press. 80-108.
- Tenn, Steven. 2011. "The Price Effects of Hospital Mergers: A Case Study of the Sutter-Summit Transaction" *International Journal of the Economics of Business* 18 (1): 65-82.
- Town, Robert J., and Gregory Vistnes. 2001. "Hospital competition in HMO networks" *Journal of Health Economics* 20 (5): 733-753.
- Train, Kenneth. 2009. *Chapter 2: Properties of Discrete Choice Models*. In: "Discrete choice methods with simulation" (2nd ed.). New York: Cambridge University Press. 11-75.
- Walker, Mike. 2005. "The Potential for Significant Inaccuracies in Merger Simulation Models" *Journal of Competition Law & Economics* 1 (3): 473-496.
- Weinberg, Matthew C., and Daniel Hosken. 2013. "Evidence on the Accuracy of Merger Simulations" *Review of Economics and Statistics* 95 (5): 1584-1600.
- Weinberg, Matthew C. 2011. "More Evidence on the Performance of Merger Simulations" *American Economic Review* 101 (3): 51-55.

- Werden, Gregory J. 2005. *Merger Simulation: Potentials and Pitfalls*. In: "Modeling European Mergers: Theory, Competition Policy & Case Studies" (eds. Bergeijk, van P.A.G., and E. Kloosterhuis). Cheltenham: Edward Elgar Publishing Limited. 37-52.
- Werden, Gregory J., and Luke M. Froeb. 2006. *Chapter 2: Unilateral Competitive Effects of Horizontal Mergers*. In: "Handbook of Antitrust Economics". (ed. P. Buccirossi). Cambridge: MIT Press. 43-104.
- Werden, Gregory J., Luke M. Froeb, and David T. Scheffman. 2004. "A Daubert Discipline for Merger Simulation." *Antitrust Magazine* 18 (3): 89-95.
- Varkevisser, Marco, Cory Capps, and Frederik T. Schut. 2008. "Defining hospital markets for antitrust enforcement: new approaches and their applicability to The Netherlands" *Health Economics, Policy and Law* 3 (1): 7-29.
- Varkevisser, Marco, Stéphanie A. van der Geest, and Frederik T. Schut. 2010. "Assessing hospital competition when prices don't matter to patients: the use of time-elasticities" *International Journal of Health Care Finance and Economics* 10 (1): 43-60.
- Varkevisser, Marco, Stéphanie A. van der Geest, and Frederik T. Schut. 2012. "Do patients choose hospitals with high quality ratings? Empirical evidence from the market for angioplasty in the Netherlands" *Journal of Health Economics* 31 (2): 371-378.
- Varkevisser, Marco, and Frederik T. Schut. 2012. "The impact of geographic market definition on the stringency of hospital merger control in Germany and the Netherlands" *Health Economics, Policy and Law*. 7: 363-381.
- Ven, van de Wynand P.M.M., and Frederik T. Schut. 2008. "Universal Mandatory Health Insurance in The Netherlands: a Model for the United States?" *Health Affairs* 27 (3): 771-781.

Chapter 4

Competition and quality indicators in the health care sector: empirical evidence from the Dutch hospital sector

With Yvonne Krabbe-Alkemade and Misja Mikkers

Published in: The European Journal of Health Economics, 19(1), 5-19. (Open access under Creative Commons 4.0 License, see <http://creativecommons.org/licenses/by/4.0/>)

ACKNOWLEDGEMENTS

We are grateful to Erik Schut, Marco Varkevisser, and two anonymous reviewers for helpful comments and suggestions.

ABSTRACT

There is much debate about the effect of competition in healthcare and especially the effect of competition on the quality of healthcare, although empirical evidence on this subject is mixed. The Netherlands provides an interesting case in this debate. The Dutch system could be characterized as a system involving managed competition and mandatory healthcare insurance. Information about the quality of care provided by hospitals has been publicly available since 2008. In this paper, we evaluate the relationship between quality scores for three diagnosis groups and the market power indicators of hospitals. We estimate the impact of competition on quality in an environment of liberalized pricing. For this research, we used unique price and production data relating to three diagnosis groups (cataract, adenoid and tonsils, bladder tumor) produced by Dutch hospitals in the period 2008–2011. We also used the quality indicators relating to these diagnosis groups. We reveal a negative relationship between market share and quality score for two of the three diagnosis groups studied, meaning that hospitals in competitive markets have better quality scores than those in concentrated markets. We therefore conclude that more competition is associated with higher quality scores.

INTRODUCTION

Many countries are facing high healthcare costs, which are also continuing to rise steadily. As a response to these trends, reforms have been implemented with the aim of controlling healthcare costs and improving the quality of care, and it seems likely that further such reforms will be put in place. The Netherlands is an example of a country that has recently reformed its healthcare system. The reforms took place in 2006 and included essential competitive elements [28].

The main objectives of the reforms to the healthcare system were to reduce costs, increase the quality and accessibility of healthcare and, at the same time, to maintain an equitable healthcare system. Today, the Dutch system can be characterized as a system of managed competition in which health insurers compete for subscribers and healthcare providers compete for contracts with health insurers. A prospective payment system has been implemented to support negotiations between health insurers and hospitals. The prices of the hospital products, called Diagnosis Treatment Combinations, are in part set by the government. Prices for complex and relatively low-volume care are regulated (representing around 30% of hospital production in 2012). Other prices (mainly for elective care) have been liberalized and are set after negotiations between insurers and hospitals [3, 29].

Any healthcare system requires quality information regarding the care provided [33]. For this, an adequate system of outcome and quality measurements is necessary and information based on the insights provided by this system must be available to consumers. In the Netherlands, quality information became publicly available in 2008 via the website <http://www.KiesBeter.nl>. Since then, consumers have been able to compare the quality of treatment for specific medical conditions at all Dutch hospitals [32]. Consumers are obliged to take out basic health insurance covering all basic healthcare including hospital care.

In this paper, we investigate the relationship between competition and quality indicators for hospital products in a market in which prices are negotiated between health insurers and hospitals. We contribute to the literature in several ways. We contribute to our knowledge of the relationship between competition and quality. The majority of the existing literature on this subject focuses on the US and UK healthcare systems (for example [5, 6, 14, 21, 23, 24, 26]). We, however, examine a European country with mandatory health insurance. Secondly, we assess the impact of competition on quality in an environment of liberalized prices. For this research, we used unique price and production data relating to three diagnosis groups (cataract, adenoid and tonsils, bladder tumor) produced by Dutch hospitals and Independent Treatment Centers in the period 2008 to 2011. Some 928,544 claims with a total revenue of 1.3 billion euros were

examined. Thirdly, most other papers do not model competition.⁵⁸ We use a model to measure market power in a framework that is rooted in economic theory [12]. Fourthly, we have also taken into account the Independent Treatment Centers (ITCs), which have entered the market in recent years, when measuring market power; many studies do not take the rising number of ITCs into account [7].

A total of 178 ITCs were licensed to provide health care services in the Dutch hospital market. Particularly in the field of ophthalmology, it is important to include ITCs in this research because ITCs are responsible for more than 10% of the market for cataract treatments [22].⁵⁹ The results reveal a negative relationship between market share and quality scores, meaning that hospitals in competitive markets achieve higher scores on quality than those in concentrated markets. The paper will proceed as follows. In 'Institutional context' we will give a brief introduction of the institutional context of the healthcare system in the Netherlands. In 'Literature', we will provide an overview of the relevant theoretical and empirical literature. We will then proceed to describe our estimation strategy and data in 'Estimation strategy', while the results are given in 'Estimation results' and 'Further examination of the empirical model'. 'Conclusion and discussion' will outline our conclusions.

1. INSTITUTIONAL CONTEXT

In 2006, the healthcare system in the Netherlands underwent extensive reform when managed competition was introduced.⁶⁰ The system is based on two fundamental pillars: competition and solidarity. The basic concept is that insurers compete for policy holders and healthcare providers compete for contracts with insurers. The idea is that health insurers contract with individual health professionals and healthcare institutions, and negotiate terms relating to service delivery, price, quality, and volume of the health care production [27]. The selective contracting of health care providers is permitted. In order to support the negotiations between hospitals and insurers, a prospective payment system was put in place. Quality information for different hospital treatments was also made publicly available.

To guarantee both income solidarity and risk solidarity, the government introduced a mandatory health insurance scheme for the entire population in 2006. The Health Insurance Act (Zorgverzekeringswet 2005) includes several requirements that safeguard equal access to healthcare for everyone. The act obliges citizens to buy a basic benefits

58 As a proxy, many papers use the number of hospitals in a given geographic radius.

59 Unfortunately, we do not have quality information of the ITCs.

60 This paragraph is partially based on [18].

package from a private insurance company and obliges insurers to accept clients without premium differentiation. This basic benefits package is legally defined and includes hospital care, general practitioner care, dental care for children under 18 years, obstetrician care, maternity home care, ambulance services, and curative mental healthcare. The premium for this basic package is roughly 50% of the expected health care cost per capita. The other 50% is paid to the government as an income-dependent premium. The government pays the sum of the income-dependent premiums into a risk adjustment fund. This fund redistributes the money via risk-adjusted payments to the insurers. The system involves virtually no co-payments: the government currently prescribes a mandatory deductible of €375 and an optional deductible (between €0 and €500). In the period that our research relates to (2008–2011), the mandatory deductible increased from €150 to €170, and the optional deductible was between €0 and €500.⁶¹

2. LITERATURE

Theoretical model

Economic theory on quality levels is highly equivocal and quality levels seem to be highly dependent on the market structures that are in place. Hospitals providing comparable health services may vary in the level of price-quality provided. It is important to identify the factors that drive this relationship. In price-regulated markets, quality is the only dimension on which one can compete. In this setting, quality levels depend on whether price-cost margins are positive or negative [12]. In non-price-regulated markets, where providers are able to determine price and quality level, quality levels depend on elasticities of demand for price and quality. If quality information is not transparent, competition will focus on the price dimension [24].

For our setting, which involves managed competition, Gaynor and Town [12] provide a relevant theoretical model. Given that in the Netherlands hospitals and insurers bargain with each other, it is most useful to analyze competition and quality through a bargaining framework. Gaynor and Town [12] (pages 566–568) present a model by which hospital and insurers bargain on price and hospitals are allowed to determine quality levels. In this framework, an insurer constructs a network of hospitals by bargaining with hospitals for their inclusion in that network. Insurers sell this network of hospitals to consumers through a health plan. The desirability of a network to a consumer depends

61 Mandatory deductibles were €150 in 2008, €155 in 2009, €165 in 2010, and €170 in 2011. Since prices for most hospital treatments are higher than the deductible and insurers do not differentiate deductibles on the basis of the hospital chosen, patients are not price-sensitive with respect to hospital treatment.

the value that he/she attaches to the hospitals that are included in that network in the event that he or she needs treatment.

The bargaining model of Gaynor and Town [12] consists of four phases.

1. Each hospital determines its quality level
2. Insurers and hospitals negotiate prices (if they agree, the hospital can become part of the health plan network)
3. Patients choose a health plan based on their preferences
4. In the event that they need treatment, patients choose a hospital from the network

Gaynor and Town [12] find that the impact of competition on quality is generally ambiguous. However, in the event that hospital demand is not responsive to price—which is actually the case in our setting, see ‘Institutional context’—then the impact of competition on quality is positive. Any increase in competition induces all hospitals to increase their quality. On balance, the effect of this increase in competition on hospitals’ prices depends on their relative bargaining position after their increase in quality. If their relative bargaining position is unchanged, then there will be no price effect. However, if there is a change in their relative bargaining position, due to, for example, differences in their marginal costs for achieving quality standards, then hospital with lower marginal costs for achieving quality will choose a relatively higher quality and improve their relative bargaining positions and, therefore, their prices. Using this framework as a guide for our empirical model, we would conjecture that in the Dutch market, we can expect a positive relationship between quality and competition.

Empirical literature

Various studies have examined the relationship between competition and quality in relation to price negotiations between hospitals and health insurers. In their literature review, Gaynor and Town [12] conclude that the impact of competition on quality remains ambiguous, ranging from negative to positive. Most studies undertaken originate in the US or UK. Lyon [19] examined the relationship between competition and quality for the managed care market. Lyon [19] showed that when there is greater price competition, price and quality are both lower than in traditional markets. Lyon’s [19] model shows that competition has a positive impact on quality when patients have a free choice of hospitals. When patients hospital choice is restricted, competition may lead to excessively low or high quality. Also, Gowrisankaran and Town [14] estimate the effect of competition on quality by comparing Medicare patients and HMO patients. Their assumption is that hospitals provide the same quality for all patients and that a change in competition will influence changes in quality for the hospital as a whole. Hospitals cannot influence Medicare prices because the government sets prices. For Medicare patients, the level of reimbursement determines whether hospitals will adjust quality (when the price is lower than cost, hospitals are incentivized to reduce quality and vice versa). For two

diagnoses, pneumonia and acute myocardial infarction, these authors use mortality rates as quality measure. The results show that for HMO patients, competition leads to lower hospital prices and higher hospital quality. For Medicare patients, their results show that competition leads to lower quality, indicating that Medicare margins are (excessively) small. Encinosa and Didem [8] estimated logit regressions to examine the relationship between safety outcomes and hospital profit margins and find a gradual negative relationship, meaning that pressure on hospital finance leads to lower quality. Escarce et al. [9] examine the relationship between competition and quality in three US states with different levels of HMO penetration. Their logistic regression models reveal a positive relationship in states with the highest average market competition measures and HMO penetration.

Propper et al. [24] examine the effect of price competition on quality with a difference-in-difference model in the UK. In this model, competing hospitals are compared with non-competing hospitals based on geographic location. They use AMI mortality as an unobserved quality measure. This study finds a negative relationship between competition and AMI death rate. However, the relationship between an observed quality measure (the waiting list for elective care procedures) and competition is positive.

An important contribution outside the hospital market is Forder and Allan [10], who study the impact of competition on quality and price of English care and nursing homes. They show that competition reduces both quality and prices in nursing homes. A major difference in the institutional setting, however, is that a considerable number of the residents in these nursing homes are 'self-payers', which makes consumers much more price-sensitive than the consumers in studies on the hospital market.

In the Netherlands, two papers have examined the competition-quality relationship [2, 15]. Bijlsma et al. [2] used outcome and process-quality indicators after managed competition was introduced in Dutch hospitals. They used the Basic Data Set of the Health Inspectorate from the period 2004–2008. Most of these indicators were at the hospital level rather than at the diagnosis level. Competition is based on fixed radius measures. Although the relationship between competition and some process indicators was positive and significant, they do not find a quality-competition relationship in the quality of outcomes. Heijink et al. [15] studied the relationship between price, volume, and quality for elective cataract surgery in the Netherlands and found little variation in cataract quality indicators between Dutch hospitals after the introduction of price competition [15]. Other Dutch studies analyzed the relationship between patient hospital choice and quality with publicly available quality data. Both studies found significant patient sensitivity to quality data [33, 1]. Varkevisser et al. [33] examined this relationship specifically in relation to angioplasty treatments, which is a treatment for which hospitals require government permission and which is price-regulated. Even when quality information is noisy (for example because quality data is not adjusted for case

mix), this relationship holds. Beukers et al. [2] studied this relationship with regard to hip replacements in the period 2008–2010. Their logit regressions indicate that although the relationship between quality and patient hospital choice is significant, travel time is a more important indicator of patients choice of hospital.

3. ESTIMATION STRATEGY

Empirical model

To determine the relation between quality and competition, we estimate a panel data model. We consider the following general linear model:

$$y_{it} = x_{it}\beta + v_{it}, \quad (1)$$

where t denotes the year and i denotes the hospital ($i = 1, 2, \dots, N$). The independent variables for hospital i in year t are given by vector x_{it} and the dependent variable is given by y_{it} . In this model $v_{it} = u_{it} + c_i$ is the composite error, where c_i is the unobserved component and u_{it} is the idiosyncratic error (see for example [34]). For each model, we made assumptions on the correlation between the unobserved component c_i and x_{it} . If we allow these to be correlated, then we have a fixed-effects model where c_i is a parameter that we will estimate. If we assume that they are uncorrelated then we have a random-effects model in which we assume a structure for the variance of v_{it} (see [34]).

In our specific application, we estimated a model for each diagnosis. For hospital i in year t ($t = 2008, \dots, 2011$) we denoted its concentration index by ms (see below for the definition) and its quality by $quality$. To control for possible case mix differences between hospitals, we included the fraction of females ($frac_female_{it}$) and the fraction of patients who are 65 years old or older ($frac_65_{it}$). To control for the fact that some hospitals deal with one dominant health insurer, while other hospitals deal with several competing insurers, we calculated the HHI of the insurers that a hospital faces ($HHI_in\ s_{it}$). We calculated this for, say, hospital i in year t , by summing up the squared shares that each insurer has in the total revenue of hospital i in year t . Furthermore, given that teaching hospitals may treat more severe patients, we included a dummy variable for teaching hospitals ($acad_{it}$) and since quality may depend on volume, we included a dummy for hospitals with a low volume of patients ($lowvolum_{it}$). In each year, the 25% of hospitals with the lowest volume were considered as low-volume hospitals.

The estimated model is:

$$quality_{it} = \beta_1 ms_{it} + \beta_2 frac_female_{it} + \beta_3 frac_65_{it} + \beta_4 HHI_in\ s_{it} + \beta_5 lowvolum_{it} + \beta_6 acad_{it} + v_{it} \quad (2)$$

To estimate the relationship between quality and competition, we needed to measure market power. There has been a great deal of debate about market definition and the measurement of market power in the literature. For an overview, see [12]. Many authors use a rather crude measure of competition: for example [24], measure competition as number of hospitals within a 30-min journey controlled for population density. However, although travel time is an important factor in hospital choice, choices can also be influenced by other patient and hospital characteristics.

Gaynor and Vogt [12] propose the use of the Logit Competition index (LOCI) to measure market power in the hospital market. The index is based on a weighted average of a hospitals market share per micro-market. The construction of the competition index starts by modeling the demand with a choice model. The choice model includes a utility function which, given characteristics of the consumer and hospital, depends on the utility that a patient derives from each hospital. The utility depends on both observable and non-observable consumer and hospital characteristics. With the logit choice model, it is possible to calculate the probability that a specific consumer type will choose a specific hospital. Each group of patients with similar characteristics (e.g., zip-code, age, gender, diagnosis, etc.) forms a micro market.

Under a standard price competition model, the competition index (LOCI) of hospital j for consumer type t is given by (see [12])

$$\Lambda_j = \sum_t w_{tj} (1 - s_{tj})$$

where the weights w_{tj} are the relative importance of each consumer type

$$w_{tj} = \frac{N_t s_{tj}}{\sum_t N_t s_{tj}}$$

and N_t is the number of consumers of the type t .

The LOCI Λ_j is a measure of the competitiveness in the market. The index takes on values between 0 and 1, where $\Lambda = 0$ means that hospital j is monopolist and $\Lambda = 1$ means that the market is perfectly competitive.

We interpret the LOCI as 1 minus the weighted market share. To simplify the interpretation of our results with we use the variable *msas* as a shorthand for “market share”, which is constructed as $(1 - \Lambda)$.

For our purposes, we are able to use actual market share with the advantage that all non-observable characteristics are implicitly taken into account. Alternatively, we could have used estimated market share, with the advantage that all consumers are taken into account. In our estimations, we used actual market share. However, the use of actual market share could potentially lead to endogeneity: hospitals providing good quality may have higher market share.

Our approach is similar to other articles about the impact of hospital competition on the quality of healthcare, e.g., Gowrisankaran and Town [13, 14], Kessler and McClellan [16], Gaynor et al. [11], and Cooper et al. [6] all use a market share based on travel distance in their estimations directly or in their IV estimations, in order to avoid potential endogeneity problems. Our approach is different from Forder and Allan [10]: they use an administrative region (Medium-level Super Output Area) as a market to calculate market share. Since they do not use patient choice models, they cannot rely on market share based on travel distance as IV. They therefore use the level of competition in neighboring areas as IV.

To prevent any endogeneity problems, we also estimate our regressions with an estimated market share (based on travel distance only) in an instrumental variable (IV) approach. Our main contribution is that we define our micro markets at the level of the quality indicator. Our micro markets consisted of the group of DTCs that are linked to the quality indicators. For each quality indicator, we estimated the relationship between the indicator and the competition indicator, which meant that we were able to construct a competition indicator for each quality indicator. The micro markets are defined by a four-digit zip code and diagnosis. The narrower a micro market becomes, the more precise the total market share becomes. However, we should not make our micro markets smaller than four-digit zip codes and diagnosis, because then we would have too few observations per zip code. For example, age is highly skewed for each diagnosis: the cataract and bladder tumor diagnosis groups include mainly elderly patients, while the tonsil diagnosis group consists mainly of younger people. This indicates that splitting the micro markets across age categories will not add a great deal of information. Furthermore, there is no reason to assume that choices would depend on gender.

Quality indicators

For the purposes of this research, we used the quality indicators from the 'Dutch Healthcare Transparency Program' (in Dutch: Zichtbare Zorg), which were developed by the Health Inspectorate in order to support various goals such as the provision of information for patients and consumers to help them make their choices, purchase information for health insurers, control information for the Inspectorate and improvement information for providers.

The Dutch Healthcare Transparency Program started in 2007 and in 2008 quality indicators became available for ten diagnosis groups. The hospitals are required to provide the registered quality indicators to the government annually.⁶² Independent Treatment Centers are not obliged to provide quality indicators, and, for this reason, we have no

62 Since 2013 the project organization for the Dutch Health Care Transparency Program has been integrated into the Dutch Quality Institute.

data on the quality indicators of the Independent Treatment Centers. The quality indicators can be divided into process indicators, structure indicators, and outcome indicators.

Structure indicators relate to the organization and are recorded at the hospital level; process indicators measure the process of activities at the patient level and outcome indicators measure outcome values at the patient level. Although outcome indicators are the most important indicators in terms of informational value, the share of outcome indicators for the Dutch Healthcare Transparency Program is less than 16%. According to an evaluation carried out by the Court of Audit (in Dutch: Algemene Rekenkamer), the indicators of the Dutch Healthcare Transparency Program indicators for the hospital sector are stable and it is therefore possible to analyze trends over the years for which records have been kept [25].

We used quality indicator data from 2008–2011. Because our time period is relatively short, we used process and structural indicators (with a ratio scale) rather than outcome indicators because process and structure indicators can be influenced by hospital management and not only by medical specialists, and these types of indicators are also used by insurers [4]. Six selection criteria were applied to include quality indicators for the diagnosis groups in our research sample. (1) The diagnoses are not selected on medical similarity but on whether hospitals can compete for patients. We used quality indicators for the diagnosis groups that made up the market segment. In this segment, the prices for the diagnosis groups are determined by negotiations between insurers and hospitals. This means that hospitals are able to compete on price and quality. This is not the case for all hospital treatments (such as urgent care). The Dutch Healthcare Authority has selected diagnoses for the market segment on criteria such as the transparency of the product definition, price and quality, the existence of market dynamics including entry and exit, the absence of undesirable effects and the absence of high transaction costs [17]. (2) Hospitals have been obliged to record quality indicators for the diagnosis groups since 2008; however, the number of diagnosis groups has grown over the years. We selected only the quality indicators that have been recorded since 2008. This means that the quality indicators that have been developed by the Health Inspectorate since 2008 are not part of our selection. (3) The quality indicators needed to be comparable over the years. (4) We selected diagnosis groups that involved surgical intervention. (5) We selected high-volume diagnosis groups with over 10,000 treatments per year. (6) We excluded indicators with categorical values (yes or no answer). Our final sample consisted of quality indicators for three diagnosis groups: cataract (ophthalmology), adenoid and tonsils (otolaryngology), and bladder tumor (urology). For bladder tumor, quality indicator data were available for 2008 to 2010 and for cataract, and data were available for 2008 to 2011 for adenoid and tonsils. The three diagnoses are elective care treatments that include daycare surgery. It should be noted that this is also the case for the bladder tumor diagnosis because the quality indicator that was used in this study

relates to the low-risk patient group (non-muscle invasive bladder tumor). Table 1 shows the quality indicators that we included in our analysis.⁶³

The indicators measure various aspects of the quality of care. For example, for cataract i02-02 measure the complication rate, while i02-03b measure the diagnosis process. We can indeed observe that (i) i02-02 have a higher average score than i02-03b and (ii) i02-02 have a lower standard deviation than i02-03b. This is unsurprising, since a one percentage-point change in i02-02 has more direct clinical relevance than a one percentage-point change in i02-03b. For this reason, it is more useful to compare changes in terms of one standard deviation. Note that i02-02 is an outlier with respect to the standard deviation. The other indicators have more similar standard deviations and clinical relevance.

In order to compare the quality indicators over the years and with one another, we rescaled and aggregated the indicators as follows: for each year t ($t = 1, 2, \dots, n$) and diagnosis group k ($k = 1, 2, \dots, K$) we calculated a combined quality indicator score per hospital i ($i = 1, 2, \dots, N$). First, each indicator h ($h = 1, 2, \dots, H$) of, say, diagnosis group k is rescaled to a z-score (z-score of indicators have an average value of zero and a standard deviation of 1):

$$z_{ith} = \frac{(P_{ith} - \mu_{th})}{sd_{th}}$$

Where z_{ith} is the value of the indicator h of diagnosis group k for hospital i in year t , μ_{th} is the average value of indicator h in year t and sd_{th} is the standard deviation of indicator h in year t . Note that for all indicators that a high z-score is associated with high quality of care and low z-score is associated with low quality of care.

Secondly, we calculate for each hospital in each year we calculated an average z-scores for diagnosis group k , which we denote by $quality_{itk}$, by averaging over z_{ith} :

$$quality_{itk} = \frac{\sum_{h=1}^H z_{ith}}{H}.$$

We thus interpreted $quality_{itk}$ as the diagnosis group k (combined) quality indicator of hospital i in year t .

⁶³ For a careful assessment of the second eye, there should be enough time between the surgery of the first eye and second eye.

Table 1. Quality indicators

Indicator code	Description	Type	Average	SD
Cataract				
i02-02	Complications: percentage of cataract operations without a posterior capsular rupture (including vitrectomy)	Process	99.63	0.35
i02-03a	Accurate diagnosis of the second eye: percentage of patients undergoing a cataract operation on both eyes with a gap of more than 28 days between the first and second operation. For a careful assessment of the second eye there should be enough time between the surgery of the first and second eye	Process	95.56	8.40
i02-03b	Accurate diagnosis of the second eye: percentage of patients undergoing a cataract operation on both eyes that have (i) a postoperative check of the first surgery before the second operation and (ii) a gap of more than 14 days between the first operation and the last postoperative check for the first operation	Process	87.63	23.21
Bladder tumor				
i01-02	Vesicoclysis: percentage of non-muscle invasive bladder cancer patients with a trans-urethral resection of the tumor (TUR) that have a washing out of the urinary bladder within 24 h after the TUR	Structure	69.98	23.66
Adenoid and tonsils				
i10-02	Preoperative consultation: percentage of tonsillectomy patients screened at an anesthesiology outpatient clinic before tonsillectomy	Process	90.80	23.58
i10-04a	Postoperative pain measurement: percentage of inpatient tonsillectomy patients that have their pain intensity measured every 8 h	Process	84.76	21.61
i10-04b	Postoperative pain measurement: percentage of measured inpatient patients without serious pain (i.e., Visual Analog Scale for Pain < 7 or Numeric Rating Scale < 7)	Process	93.32	14.51
i10-04c	Postoperative pain measurement: percentage of daycare patients that have been telephoned after their operation to monitor their pain intensity	Process	75.47	38.63

This table contains for each quality indicator its average score and standard deviation pooled over the years

4. DATA

This research was based on claims data from 2008 to 2011 from all Dutch hospitals and Independent Treatment Centers (ITCs). The Dutch hospital market consists of 87 hospitals, two specialized hospitals, and eight academic hospitals. The total number of ITCs rose from 189 in 2008 to 282 in 2012. Our unique dataset consists of patient-level data including patient characteristics such as gender, age, zip code, diagnosis and treatment, hospital characteristics and hospital contract prices for our three selected diagnoses. The total number of patients in the period 2008 and 2011 for cataract was 474,410 with

a total revenue of 843 million euros. The total number of patients in the period 2008 and 2011 for adenoid and tonsils was 223,177 with a total revenue of 219 million euros. The total number of patients in the period 2008 and 2011 for bladder tumor was 46,497 with a total revenue of 244 million euros.

The calculation of the weighted market share was based on the claims data from the hospitals and Independent Treatment Centers (ITCs), where we removed those cases with invalid zip codes (which amounted to less than 1% of our sample). Thus, when calculating the market share, we were able to take into account the (potential for) competitive pressure from ITCs. However, as discussed in “Quality indicators”, we had no data on the quality indicators for the ITCs. This implies that the analysis of the quality indicators is restricted to hospitals and excludes the ITCs. For each diagnosis, we removed those hospitals where there was no data on quality indicators. Tables 2, 3, and 4 present the descriptive statistics for the variables used in our empirical model at the diagnosis level. The total number of observations ranges from 191 for tonsils to 286 for cataract. For example, the average actual market share ms for cataract was 0.58 (SD 0.21) with a minimum of 0.06 and maximum of 0.97 meaning that there are large differences between the market share of hospitals. The other diagnoses display comparable actual market share. The average hospital-insurer HHI is moderately strong and ranges from 0.39 for cataract, 0.32 for tonsils, and 0.38 for bladder tumor.

As mentioned in “Quality indicators”, we have used standardized quality scores for each diagnosis. The standardized quality scores ranged between -3.05 and 0.78 for cataract, between -2.17 and 0.71 for tonsils and between -2.85 and 1.42 for bladder tumor.

We included two hospital characteristics in our empirical model. Firstly, we used a dummy variable to control for whether a hospital is a general hospital or academic (university) hospital, a low-volume dummy, and insurance-hospital HHI. For each diagnosis, less than 10% of the hospitals included were academic hospitals. To control for patient characteristics, we included three variables: the fraction of female patients, the fraction of patients that were older than 65 years, and the co-morbidity index. The results differed per diagnosis due to the characteristics of the condition and the set of hospitals included in the analysis. The co-morbidity index, which is defined in this study as the average number of diagnoses, varied considerably between the diagnoses. The co-morbidity for tonsils had the lowest index of 1.04. The average number of additional diagnoses for bladder tumor (3.19) and cataracts (2.18) was higher due to the older population involved.

Table 2. Cataract

Statistic	N	Mean	SD	Min	Max
<i>ms</i>	286	0.58	0.21	0.06	0.97
Quality	286	0.00	0.63	-3.05	0.78
frac_female	286	0.59	0.03	0.51	0.68
frac_65	286	0.83	0.06	0.58	0.91
Com	286	2.18	0.30	1.60	3.17
HHL_ins	286	0.39	0.11	0.19	0.63
Lowvolume	286	0.25	0.43	0	1
Acad	286	0.09	0.28	0	1

This table shows summary statistics for each diagnosis group at the hospital-year level (2008–2011). We report the average, standard deviation, minimum, and maximum of the variables that we included in our regression analysis. We also show the total number of observations for each diagnosis group

Table 3. Adenoid and tonsils

Statistic	N	Mean	SD	Min	Max
<i>ms</i>	191	0.69	0.19	0.12	0.98
Quality	191	0.00	0.59	-2.17	0.71
frac_female	191	0.51	0.03	0.42	0.60
frac_65	191	0.003	0.003	0.00	0.02
Com	191	1.04	0.30	0.53	2.18
HHL_ins	191	0.32	0.09	0.17	0.57
Lowvolume	191	0.25	0.43	0	1
Acad	191	0.07	0.26	0	1

This table shows summary statistics for each diagnosis group at the hospital-year level (2008–2011). We report the average, standard deviation, minimum, and maximum of the variables that we included in our regression analysis. We also show the total number of observations for each diagnosis group

Table 4. Bladder tumor

Statistic	N	Mean	SD	Min	Max
<i>ms</i>	199	0.73	0.15	0.33	0.98
Quality	199	0.00	1.00	-2.85	1.42
frac_female	199	0.22	0.05	0.10	0.45
frac_65	199	0.72	0.06	0.52	0.86
Com	199	2.68	0.46	1.70	4.32
HHL_ins	199	0.38	0.11	0.19	0.71
Lowvolume	199	0.23	0.42	0	1
Acad	199	0.10	0.29	0	1

This table shows summary statistics for each diagnosis group at the hospital-year level (2008–2010). We report the average, standard deviation, minimum and maximum of the variables that we included in our regression analysis. We also show the total number of observations for each diagnosis group

5. ESTIMATION RESULTS

Tables 5, 6 and 7 present the estimation results for each diagnosis group using the pooled OLS estimator, fixed effect estimator, and random-effects estimator. For each estimation method, we present two models: a model with and a model without control variables. Because the estimated standard errors could be heteroskedastic, we used the Mackinnon and White [20] Heteroskedasticity-Consistent standard errors.

We performed the Hausman test to determine whether we should select the fixed or random effect model: given that (i) the fixed effect model is consistent if the unobserved component and observable variables are correlated and (ii) the random effect model is inconsistent, the presence of significant differences in the estimated coefficients

Table 5. Regression results bladder tumor

	Dependent variable					
	Quality					
	Pooled	Pooled	Fixed effects	Fixed effects	Random effects	Random effects
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	1.275*** (0.338)	2.058 (1.347)			1.348*** (0.360)	2.803** (1.265)
<i>ms</i>	-1.731*** (0.492)	-1.764** (0.705)	-2.504* (1.381)	-3.331** (1.695)	-1.841*** (0.510)	-2.241*** (0.674)
<i>frac_female</i>		-1.136 (1.483)		0.066 (1.262)		-0.149 (1.019)
<i>frac_65</i>		-0.173 (1.686)		-1.975 (1.765)		-1.001 (1.425)
<i>Com</i>		0.018 (0.205)		-0.241 (0.159)		-0.106 (0.158)
<i>HHI_ins</i>		-1.019 (0.889)		1.469 (1.217)		-0.134 (0.801)
<i>Lowvolume</i>		-0.165 (0.260)		-0.004 (0.231)		-0.200 (0.188)
<i>Acad</i>		-0.126 (0.271)				-0.286 (0.300)
Observations	199	199	199	199	199	199
<i>R</i> ²	0.069	0.088	0.029	0.073	0.050	0.063
Adjusted <i>R</i> ²	0.068	0.084	0.017	0.040	0.049	0.060

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. We report the results from the pooled, fixed-effects, and random-effects model. For each model, we report two variants: (i) a simple model with the weighted market share regressed on the average scaled quality indicator and (ii) a model with additional control variables. We report the MacKinnon and White Heteroskedasticity-Consistent standard errors (in parentheses under coefficients). We used data from 2008 to 2010

indicates that there is correlation between the unobserved component and observable variables and, thus, that we should discard the random-effects model [34]. For all models, the Hausman test accepts the null hypotheses that there is no difference between the coefficient of the random and fixed-effects model (for each model, the p value is larger than 0.10). This indicates that the random-effects models are consistent and, therefore that we can use the results of the random effect models to determine the relationship between quality indicators and market share.

Table 5 presents the estimation results for bladder tumor using six different models. The adjusted R -squared of these models ranges between 0.017 for the fixed effect model and 0.084 for the pooled model. From all the models shown in Table 5, we can conclude that the weighted market share (1 minus LOCI) is significantly related with scaled quality

Table 6. Regression results cataract

Dependent variable

	Quality					
	Pooled	Pooled	Fixed effects	Fixed effects	Random effects	Random effects
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	0.122 (0.145)	-2.552* (1.550)			0.149 (0.149)	-2.631** (1.237)
<i>ms</i>	-0.199 (0.259)	-0.584** (0.292)	-1.481 (1.023)	-1.923** (0.883)	-0.264 (0.261)	-0.772** (0.325)
<i>frac_female</i>		2.541 (1.925)		1.208 (2.168)		1.718 (1.830)
<i>frac_65</i>		1.398 (1.294)		1.354 (1.703)		1.732 (1.234)
<i>Com</i>		0.072 (0.228)		0.479 (0.377)		0.216 (0.235)
<i>HHL_ins</i>		0.196 (0.680)		3.568* (2.082)		0.467 (0.691)
<i>Lowvolume</i>		0.069 (0.120)		-0.108 (0.154)		-0.017 (0.127)
<i>Acad</i>		-0.196 (0.313)				-0.229 (0.287)
Observations	286	286	286	286	286	286
R^2	0.004	0.072	0.014	0.070	0.004	0.051
Adjusted R^2	0.004	0.070	0.010	0.048	0.004	0.050

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. We report the results from the pooled, fixed-effects, and random-effects model. For each model, we report two variants: (i) a simple model with the weighted market share regressed on the average scaled quality indicator and (ii) a model with additional control variables. We report the MacKinnon and White Heteroskedasticity-Consistent standard errors (in parentheses under coefficients). We used data from 2008 to 2011

score. The estimated coefficients are significant at a level of 1% for the random-effects models and the pooled effect model without control variables. The other models are significant at a level of 5% (fixed effects with control variables and pooled model with control variables) and the 10% level (fixed effect model without control variables). This indicates that hospitals in competitive markets have higher quality scores for bladder tumor, which supports the hypothesis that greater competition leads to higher quality scores. Models 2, 4, and 6 include six control variables which control for the patient characteristics of the hospital (fraction of female patients, fraction of patients older than 65 years and co-morbidity index) and hospital characteristics⁷ including health insurer-hospital HHI, the type of hospital and low-volume dummy.

Table 7. Regression results adenoid and tonsils
Dependent variable

	Quality					
	Pooled	Pooled	Fixed effects	Fixed effects	Random effects	Random effects
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	-0.008 (0.236)	0.292 (0.902)			-0.089 (0.283)	-0.064 (0.804)
<i>ms</i>	0.011 (0.323)	-0.068 (0.562)	0.723 (1.537)	0.382 (1.639)	0.099 (0.383)	-0.034 (0.514)
<i>frac_female</i>		-0.388 (1.452)		-0.191 (1.158)		-0.101 (1.215)
<i>frac_65</i>		-19.167 (25.528)		4.575 (19.921)		-17.188 (23.471)
<i>Com</i>		0.257 (0.198)		0.379 (0.422)		0.304 (0.201)
<i>HHI_ins</i>		-0.579 (0.602)		1.394 (1.915)		-0.233 (0.608)
<i>Lowvolume</i>		-0.185 (0.231)		0.036 (0.182)		-0.156 (0.178)
<i>Acad</i>		-0.224 (0.311)				-0.228 (0.314)
Observations	191	191	191	191	191	191
R ²	0.00001	0.066	0.002	0.015	0.002	0.038
Adjusted R ²	0.00001	0.063	0.001	0.008	0.002	0.036

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. We report the results from the pooled, fixed-effects, and random-effects model. For each model, we report two variants: (i) a simple model with the weighted market share regressed on the average scaled quality indicator and (ii) a model with additional control variables. We report the MacKinnon and White Heteroskedasticity-Consistent standard errors (in parentheses under coefficients). We used data from 2008 to 2011

Table 8. Disaggregated regression results for market share (ms)

	Dependent variable		
	Quality		
	Pooled (2)	Fixed effects (4)	Random effects (6)
Cataract: i02-02 ms	0.040	-2.812	-0.173
Cataract: i02-03a ms	-1.024**	-0.327	-1.038**
Cataract: i02-03b ms	-0.775**	-2.707*	-0.906**
Adenoid and tonsil: i10-02 ms	0.361	-2.316	0.018
Adenoid and tonsil: i10-04a ms	0.901	-3.107	0.744
Adenoid and tonsil: i10-04b ms	-0.369	3.018	-0.334
Adenoid and tonsil: i10-04c ms	-1.150*	3.782	-0.556

* $p<0.1$; ** $p<0.05$; *** $p<0.01$. We report the estimated market-share coefficient from the pooled, fixed-effects, and random-effects models, which also include the control variables. The estimated coefficient of the control variables are available from the authors by request. We report the MacKinnon and White Heteroskedasticity-Consistent standard errors (in parentheses under coefficients). We used data from the period 2008–2011

Tables 6 and 7 show the results for the diagnosis groups cataracts and adenoid and tonsils. The results of the models for cataract are comparable to the results for bladder tumor and reveal a negative relationship between market concentration and quality scores. Contrary to what we expected, the models for the diagnosis group adenoid and tonsils do not show any significant estimations.

For the cataract and bladder tumor diagnosis groups, we can show how to interpret the magnitude of the estimated relationship between quality scores and market share through an example. Consider the estimated difference in the quality scores for a hospital with the lowest market share (i.e., a market share of 0.06 for cataract, see Table 2) and the hospital with the highest market share (i.e., market share 0.98 for bladder tumor, see Table 4). Using the results of the random-effects model for bladder tumor presented in Table 6, the market share difference of 0.92 translates into a shift in quality scores of -2.06 standard deviations from the mean (95% confidence-interval [-3.39; -0.73]). Similarly, using the results of the random-effects model for bladder tumor presented in Table 5, the market share difference of 0.92 translates into a shift in quality scores of -0.71 standard deviations from the mean (95% confidence-interval: [1.35; -0.069]).

Robustness check

To check the robustness of our model with aggregated quality indicators, we also estimated the pooled, fixed-effects, and random-effects model for each individual quality indicator separately (not relevant for bladder tumor). Table 8 shows the estimated market-share coefficients for each quality indicator. For the sake of clarity, we only show the estimated market-share coefficients. For each diagnosis, the sign of the estimated

Table 9. Result: conditional logit model

	Dependent variable		
	Hospital choice		
	Bladder cancer (1)	Cataract (2)	Adenoid and tonsil (3)
2008 traveltime	-0.220*** (0.002)	-0.232*** (0.001)	-0.195*** (0.001)
2009 traveltime	-0.221*** (0.002)	-0.235*** (0.001)	-0.183*** (0.001)
2010 traveltime	-0.231*** (0.002)	-0.236*** (0.001)	-0.179*** (0.001)
2011 traveltime	-0.221*** (0.002)	-0.230*** (0.001)	-0.182*** (0.001)
Number of patients 2008	9436	53,449	38,621
Number of patients 2009	9854	56,412	38,211
Number of patients 2010	10,122	57,391	39,032
Number of patients 2011	10,319	53,144	40,432

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. For each year and diagnosis group, we report the results from the conditional logit model with travel time as the only predictor. For cataract we take a random sample. In each year, the sample size is 50% of the total patient population

market-share coefficients are generally consistent with the aggregated model: negative for cataract and mixed for adenoid and tonsils.

Further examination of the empirical model

In our empirical estimations, we control for differences in patient and hospital characteristics. However, there is a potential source of bias in the estimation of the relationship between weighted market share and quality indicators: Firstly, as discussed in, "Empirical model" hospitals with a high-quality indicators may attract more patients than hospitals with low-quality indicators. This means that our regression models may suffer from an endogeneity problem and therefore give a biased result.

A potential issue is that hospitals may change their supply in response to their own or their competitors quality indicators. However, there were no indications that hospitals significantly restructured their supply:

- no new hospitals or new hospital locations were opened during the period of our study.
- no hospitals or hospital locations were closed during the period of our study.

Table 10. Result: pooled instrumental variable model

	Dependent variable		
	Quality		
	Bladder tumor (1)	Cataract (2)	Adenoid and tonsil (3)
Constant	2.094 (1.349)	-2.380 (1.560)	0.617 (0.889)
<i>ms</i>	-1.829*** (0.707)	-0.753** (0.294)	-0.603 (0.545)
<i>frac_female</i>	-1.134 (1.485)	2.412 (1.922)	-0.317 (1.433)
<i>frac_65</i>	-0.159 (1.691)	1.432 (1.311)	-18.521 (26.590)
<i>Com</i>	0.016 (0.205)	0.046 (0.229)	0.193 (0.202)
<i>HHL_ins</i>	-0.993 (0.889)	0.293 (0.681)	-0.285 (0.612)
<i>Lowvolume</i>	-0.175 (0.261)	0.055 (0.120)	-0.235 (0.229)
<i>Acad</i>	-0.137 (0.271)	-0.243 (0.325)	-0.384 (0.290)
Observations	199	286	191
<i>R</i> ²	0.088	0.071	0.055
Adjusted <i>R</i> ²	0.084	0.069	0.053

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. For each diagnosis group, we report the results from the pooled model, where we use the simulated weighted market share as an instrument for the weighted market share. The simulated weighted market share is based on a multinomial logit model. We report the MacKinnon and White (1985) Heteroskedasticity-Consistent standard errors (in parentheses under coefficients). We used data from 2008 to 2011

- virtually no hospital increased or decreased their volume (number of patients) significantly.⁶⁴
- the presence and entry of independent treatment centers (ITC) occurred mainly in the cataract diagnosis group. There was no ITC in the bladder tumor diagnosis group and only four ITCs in the adenoid and tonsils diagnosis group.

Secondly, hospitals and insurers negotiate the prices of the products in the diagnosis groups that we examined in a liberalized setting. An insurer is likely to be interested in the quality–price ratio and not the quality or price in isolation. Furthermore, as we noted

⁶⁴ For each year, we classified all hospitals into four quartiles, based on the number of patients treated. For cataract and bladder tumor, there was only one hospital that increased its volume so much that it moved more than two quartiles.

in “Literature”, there may be a relationship between competition and prices through quality. In this section, we will examine both these issues.

Endogeneity

To test whether our (pooled) regression model suffers from the endogeneity problem, we used the Wu–Hausman test. In this test the result of an instrumental variable (IV) model is compared to the result if the non-IV model.

To estimate an IV model, we must find an instrument for market share that meets two requirements: (1) it should be strongly correlated to market share and (2) it should be uncorrelated to the error term. We took an approach that has been commonly used in the health economics literature to deal with the endogeneity problem, namely to deploy IV instruments based on predicted patient flows (see for example [6, 16], and [11]). We estimated a conditional logit model that was based only on the travel time between the patient and hospital. By excluding the fixed-effects and other hospital and patient characteristics, we ensured that our predicted patient flows were exogenous to all patient preferences other than travel time. This implies that our patient flows were not influenced by the potential preferences that patients may have a hospital with a certain level of quality indicators.

In the conditional logit model, we estimated the utility that patient t would receive when he or she chose hospital j (j, \dots, n), which we denote by U_{tj} . In our case, where we only use travel time as a predictor, the conditional logit model is given by:

$$U_{tj} = \alpha \text{traveltime}_{tj} + \epsilon_{tj},$$

where the residuals ϵ are i.i.d. extreme value (see for example [31]). We used the same data as in “Estimation strategy”, with the addition of the travel times between the patient’s home and the hospital. The travel time was calculated as the driving time between the patient’s home and the hospital zip code. We estimated this model for each diagnosis group and year (see Table 9 for the estimation results). For cataract treatment, we took a random sample of the data for computability. In each year, the sample size was 50% of the total patient population. As expected, we found that patients dislike traveling (negative coefficient for time travel).

Given the fitted utilities \hat{U}_{tj} from the estimated model, we can calculate for each patient the probability \hat{P}_{tj} that this patient (t) chooses hospital j :

$$\hat{P}_{tj} = \frac{\hat{U}_{tj}}{\sum_k^n \hat{U}_{tk}}.$$

We can now calculate the simulated weighted market share for each hospital in each year by replacing s_{tj} with P_{tj} in “Empirical model” in (4.1). To determine the relevance

of the instruments, we determined first-stage F-statistic, see for example [30]. For each model we found that the first-stage F-statistic was higher than 56 (p -value <0.001). This indicates that our instruments were not weak.⁶⁵ Subsequently, we re-estimated the pooled regressions in “Estimation results” by employing the simulated weighted market share as the instrumental variable (see Table 10 for the results). Given the IV-estimated coefficients, we could now carry out the Wu-Hausman test, in which we tested the hypothesis that there was significant difference between the coefficients from the pooled model and the coefficients from the IV model. For each diagnosis group we accepted the hypothesis that there was no difference (p -value >0.90). We performed an additional IV estimation of the random-effects model (which is preferred to the fixed-effects model, see above). For bladder tumor, the estimated market-share coefficient is -2.45 (p -value=0.062). For cataract, the estimated market-share coefficient is -0.94 (p -value=0.107). For adenoid and tonsils, the estimated market-share coefficient is -0.55 (p -value=0.35). These IV random-effects coefficient are very similar to the non-IV random-effects coefficient presented in the paper. The only difference is that the standard errors are larger, which is no surprise, since it is much more difficult to explain the differences than the levels of the quality indicators with an (IV) regression model. We were therefore able to conclude that our regression did not suffer from endogeneity bias.

Prices

As discussed above, there may be a relationship between prices and quality indicators. However, a comprehensive examination of the relationship between prices and quality indicators is beyond the scope of this paper. Instead, in this section we will provide a brief outline of this relationship. In “Literature”, we discussed the bargaining model by Gaynor and Town [12]. From this model, it follows that the effect of an increase in competition—through quality—on price depends on the effect that the increase in quality has on the relative bargaining position and therefore the prices charged by a hospital. We do not observe the factors that determine this effect such as the marginal cost of quality, see “Literature”. Generally, we would expect that hospitals and insurers to be interested in the quality-to-price ratio and not price or quality in isolation.

To provide an indication of the relationship between quality and prices, we estimated the same model as our quality indicator-market share model in “Empirical model”, where we included the relative prices as the independent variable rather than market share. We then estimated the pooled, fixed-effects, and random-effects models. In general, we

⁶⁵ Furthermore, a simple correlation analysis shows that the simulated market share used in the IV are correlated with the original market share. In each year, the correlation coefficient is between 0.67 and 0.71 for bladder tumor, between 0.72 and 0.90 for adenoid and tonsils, and between 0.69 and 0.77 for cataract.

found no significant results for the price variable.⁶⁶ Our results indicate that, for the hospital products examined, having a relatively higher (or lower) quality indicators was not related to higher (or lower) prices during the period that we studied. We can therefore conclude that hospitals with better quality indicators are not compensated by higher prices.

6. CONCLUSIONS AND DISCUSSION

In this study, we have examined the impact of competition on the quality of healthcare for the Dutch hospital market. The Dutch government reformed the health care system in 2006 introducing managed competition in a context where income and risk solidarity are guaranteed. With this system, the government aims to reduce costs and increase quality of care. Health insurers compete for policy holders and healthcare providers compete for contracts with health insurers. We used a unique data set including individual patient-level claim data and information on quality indicators for three diagnosis groups—cataracts, adenoid, and tonsils and bladder tumor—produced by Dutch hospitals and Independent Treatment Centers.

For cataract and bladder tumor, the relationship between market share and quality scores was found to be significant. The robustness checks confirmed these results. The regression estimators for adenoid and tonsils were not significant. One possible explanation is that the patient group for adenoid and tonsils is less complex. It is mainly children younger than 11 years who are treated for adenoid and tonsils. This type of patient is less complicated and has fewer additional diagnoses compared to patients with, for example, bladder tumor (the fraction of patients older than 65 years is 0.72 for bladder tumor, 0.83 for cataract, and 0.01 for adenoid and tonsils, respectively).

Because of endogeneity, we could not include price as independent variable in our quality indicator models. To give an indication of this relationship, we replaced the market-share variable with the relative prices as the independent variable. We conclude from these models that there is no relationship between price and quality scores, which means that hospitals with higher-quality scores do not have higher prices.

Overall, we conclude that more competition leads to better published quality scores. This research does have some limitations however. For this research, we used observed quality information that did not include mortality rates (an outcome indicator). For our research period data, no data on mortality rates was available. However, since 2014 hospitals are obliged to publish standardized mortality ratios on their websites. For future

66 Only in the random-effects model for bladder tumor we find a significant (positive) result. The results are available from the authors on request.

research, it would be interesting to examine the relationship between mortality rates and competition within the Dutch hospital market.

REFERENCES

1. Beukers, P.D., Kemp, R.G., Varkevisser, M.: Patient hospital choice for hip replacement: empirical evidence from the Netherlands. *Eur. J. Health Econ.* 15(9), 927–936 (2014)
2. Bijlsma, M.J., Koning, P.W., Shestalova, V.: The effect of competition on process and outcome quality of hospital care in the Netherlands. *De Economist* 161(2), 121–155 (2013)
3. Boonen, L.H., Schut, F.T., et al.: Preferred providers and the credible commitment problem in health insurance: first experiences with the implementation of managed competition in the Dutch health care system. *Health Econ. Policy Law* 6(2), 219 (2010)
4. Botje, D., Klazinga, N.S., Wagner, C.: To what degree is the governance of Dutch hospitals orientated towards quality in care? Does this really affect performance? *Health Policy* 113(1), 134–141 (2013)
5. Brekke, K.R., Siciliani, L., Straume, O.R.: Hospital competition and quality with regulated prices. *Scand. J. Econ.* 113(2), 444–469 (2011)
6. Cooper, Z., Gibbons, S., Jones, S., McGuire, A.: Does hospital competition save lives? Evidence from the English NHS patient choice reforms. *Econ. J.* 121(554), F228–F260 (2011)
7. Courtemanche, C., Plotzke, M.: Does competition from ambulatory surgical centers affect hospital surgical output? *J. Health Econ.* 29(5), 765–773 (2010)
8. Encinosa, W.E., Bernard, D.M.: Hospital finances and patient safety outcomes. *INQUIRY J. Health Care Organ. Provis. Financ.* 42(1), 60–72 (2005)
9. Escarce, J.J., Jain, A.K., Rogowski, J.: Hospital competition, managed care, and mortality after hospitalization for medical conditions: evidence from three states. *Med. Care Res. Rev.* 63(6 suppl), 112S–140S (2006)
10. Forder, J., Allan, S.: The impact of competition on quality and prices in the English care homes market. *J. Health Econ.* 34, 73–83 (2014)
11. Gaynor, M., Moreno-Serra, R., Propper, C.: Death by market power: reform, competition, and patient outcomes in the National Health Service. *Am. Econ. J. Econ. Policy* 5(4), 134–166 (2013)
12. Gaynor, M., Town, R.J. Competition in health care markets. In: Pauly, M.V., McGuire, T.G., Barros, P.P. (eds.). *Handbook of Health Economics, Volume 2 of Handbook of Health Economics*, Chapter 9, pp. 499–637. Elsevier, Amsterdam (2011)
13. Gowrisankaran, G., Town, R.J.: Estimating the quality of care in hospitals using instrumental variables. *J. Health Econ.* 18(6), 747–767 (1999)
14. Gowrisankaran, G., Town, R.J. Competition, payers, and hospital quality. *Health Serv. Res.* 38(6p1), 1403–1422 (2003)
15. Heijink, R., Mosca, I., Westert, G.: Effects of regulated competition on key outcomes of care: cataract surgeries in the Netherlands. *Health Policy* 113(1), 142–150 (2013)
16. Kessler, D.P., McClellan, M.B.: Is hospital competition socially wasteful? *Q. J. Econ.* 115(2), 577–615 (2000)
17. Krabbe-Alkemade, Y., Groot, T., and Lindeboom, M.: Competition in the Dutch hospital sector: an analysis of health care volume and cost. *Eur. J. Health Econ.* 1–15 (2016)
18. Krabbe-Alkemade, Y.J.F.M.: The Impact of Market Competition and Patient Classification on Dutch Hospital Behavior. Vrije Universiteit Amsterdam (2014)
19. Lyon, T.P.: Quality competition, insurance, and consumer choice in health care markets. *J. Econ. Manag. Str.* 8(4), 546–580 (1999)
20. MacKinnon, J.G., White, H.: Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *J. Econom.* 29(3), 305–325 (1985)

21. Noether, M.: Competition among hospitals. *J. Health Econ.* 7(3), 259–284 (1988)
22. NZa: Monitor zelfstandige behandelcentra. NZa publicaties (2012)
23. Propper, C., Burgess, S., Gossage, D.: Competition and quality: evidence from the NHS internal market 1991–9. *Econ. J.* 118(525), 138–170 (2008)

Chapter 5

Evidence of selection in a mandatory health insurance market with risk adjustment

With Katalin Katona, Misja Mikkers and Victoria Shestalova

ACKNOWLEDGEMENTS

his paper reflects the personal views of the authors, which are not necessarily those of their employers. We would like to thank Erik Schut and Marco Varkevisser (Erasmus University) and Cédric Argenton (TILEC, Tilburg University) for their valuable comments.

ABSTRACT

This paper aims to identify selection separately from moral hazard in a mandatory health insurance market where enrollees can freely choose their deductible scheme. The empirical analysis uses a unique dataset for the period 2010-2013 covering the whole population of the Netherlands at enrollee level, allowing us to use prior health expenses of the enrollees to demonstrate the selection effect separately from the potential moral hazard effect. Our estimates show that the enrollees who opt for deductibles are both healthier and have a higher risk-adjusted result (i.e. the difference between the compensation from the risk-adjustment fund and the actual health care cost) under the prevailing risk-adjustment system. Compared to enrollees who have chosen the lowest available deductible level, enrollees who have chosen the highest deductible level have an average risk-adjusted result that is approximately €450 higher per enrollee. An option that the Dutch government could consider to fully eliminate the risk-adjustment gain of the deductibles is to include the choice of a voluntary deductible in the risk-adjustment system as one of the characteristics of the consumer. Our detection of substantial selection effect of deductibles suggests the need of further research to understand in greater detail the relationship between premium discounts and the expected gains on the risk-adjustment for enrollees with a voluntary deductible.

1. INTRODUCTION

In order to mitigate market failures and increase affordability in health insurance markets, regulators in several countries, including the Netherlands, set rules for consumers and insurers. Most importantly, these rules oblige insurance companies to accept all applicants (open enrollment) and to offer health insurance policies at the same price to all persons (community rating). An other important rule is to oblige all citizens to enroll in health insurance⁶⁷.

Generally, a risk-adjustment system is in place to provide a level playing field for the insurers. A perfect risk-adjustment system eliminates the predictable profits (losses) on low-risk profile (high-risk profile) enrollees emanating from community rating. In that case, insurers have no incentives to select low-risk enrollees or to offer products that are attractive to low-risk enrollees⁶⁸.

The Dutch risk-adjustment system began simply and has gradually been improved. However, even a state-of-the-art system cannot remove incentives for exploiting the mechanism of selection fully (van Kleef, McGuire, van Vliet, & van de Ven, 2017). There are still incentives for the insurers to sort low-risk and high-risk enrollees into different health plans by designing different health plans for different risk profiles. Because health plans may be priced independently, this would result in premium differentiation according to risk profiles and risk-based sorting across plans. Such selection decreases solidarity and induces social welfare losses (Cutler, Zeckhauser, et al., 1997). In the literature on selection, there is a focus on (i) the relationship between risk-type and the demand for contracts with generous reimbursement (i.e. adverse selection) and (ii) the incentives for health insurers to select low-risk individuals by tailoring insurance options to attract them (Newhouse, Price, Hsu, McWilliams, & McGuire, 2015; Newhouse, Price, Huang, McWilliams, & Hsu, 2012; Einav & Finkelstein, 2011; Rothschild & Stiglitz, 1976).

Empirical identification of selection is challenging because of the interaction of selection with moral hazard (Cohen & Siegelman, 2010). Using a unique dataset on the Dutch health insurance market, this paper contributes to the existing literature by showing empirically that the possibility for consumers to opt for voluntary deductible can result in selection, even in a health insurance market with a risk-adjustment system in place. By taking account of the past health costs and focusing on whether consumers opt for deductibles, we are able to identify the selection effect separately from moral hazard

67 For a description of this general idea see e.g. van de Ven and Ellis (2000).

68 However, a recent theoretical insight from Bijlsma, Boone, and Zwart (2014) suggests that selection could exist even in a market with optimal risk-adjustment. When the health insurance market is imperfectly competitive and healthy consumers have a higher price elasticity than the high-risk consumers, enrollees can still be 'sorted' into health plans with different prices and coverage.

effect. Further, we examine the selection effect separately in terms of adverse selection and in terms of the incentives for health insurance to select low-risk enrollees.

With our study, we improve the understanding of selection in managed competition settings. Our result is therefore also relevant for other health care systems with managed competition, such as Medicare Advantage and Part D in the United States.

In the remainder of this paper we will discuss the previous literature in section 2 explain the organization of the Dutch health insurance sector in section 3. This is followed by the exposition of our empirical approach in section 4, after which we describe the data and our empirical findings in sections 5 and 6. Section 7 ends with concluding remarks.

2. LITERATURE

In this section we focus on the previous literature on voluntary deductible and the empirical identification of selection in health insurance markets.

Voluntary deductibles

In the Netherlands there is the possibility to choose from five different levels of voluntary deductible in exchange for a premium discount on the basic benefit package. Although less than 11% of the Dutch insured chose voluntary deductibles in 2014, the group of enrollees with voluntary deductibles is growing every year (NZa, 2014; Vektis, 2017).

Looking at the potential effects of voluntary deductibles from the perspective of the economic literature, the following opposite effects can be highlighted. On the one hand, deductibles reduce the effect of moral hazard. Individuals incur lower health care cost if they are enrolled in a health plan with a higher deductible (Newhouse, 2004; Aron-Dine, Einav, & Finkelstein, 2013). Most studies attribute this cost reduction to lower health care utilization (Keeler, 1992; Gern & Schellhorn, 2006; Wharam et al., 2007), rather than to choosing cheaper hospitals. Lower health care utilization is desirable when it applies to care valued below its cost⁶⁹.

On the other hand, deductibles are an instrument of selection. Rothschild and Stiglitz (1976) shows that due to the private information available to consumers, selection results in the underinsurance of low-risk enrollees. Moreover, since deductibles are only attractive for enrollees with lower expected costs, they allow for selection by insurers (van Kleef, Beck, van de Ven, & van Vliet, 2008; Tollen, Ross, & Poor, 2004).

69 Note that enrollees who have to pay deductibles (or a co-insurance rate) may reduce or postpone necessary health care treatment (Brot-Goldberg, Chandra, Handel, & Kolstad, 2015; Fronstin & Collins, 2008; Lohr et al., 1986; Davis, Doty, & Ho, 2005; Galbraith et al., 2011).

More specific to our context, van Winssen, van Kleef, and van de Ven (2016) argue that in the Netherlands voluntary deductibles mitigate moral hazard, but it also involves an (adverse) selection component. The Dutch Healthcare Authority has found that enrollees who choose a voluntary deductible have a higher risk-adjusted result than enrollees who choose no voluntary deductible (NZa, 2016). However, the mere fact that enrollees who choose a voluntary deductible have a higher risk-adjusted result does not, in itself, prove selection. It also may be due to the fact that these enrollees consume less health care because they wish to avoid paying the higher deductible (moral hazard). Therefore, the finding suggests that a more elaborated analysis is still needed in order to disentangle the two effects.

Identification of selection

There is a growing body of empirical literature on identifying selection in health insurance markets. For example, Panthöfer (2016) finds adverse selection in the German public health insurance market. Olivella and Vera-Hernández (2013) test for asymmetric information in the UK private health insurance market and find evidence for adverse selection. Dardanoni and Donni (2012) find significant adverse and advantageous selection in the US Medigap insurance market. Bolhaar, Lindeboom, and van der Klaauw (2012) find that information asymmetry is present in the supplementary health insurance in the Netherlands. For an extensive review of the empirical literature on the relationship between coverage and risk see Cohen and Siegelman (2010) and Aarbu (2017).

We conduct a version of the positive correlation test, which is described in Chiappori and Salanie (2000), to determine the extent of selection on the market. A positive correlation test assesses, conditional on observables, if there is a correlation between the choice for a contract and the occurrence or severity of an accident, which is in our case health care expenditure.

We take broadly a similar approach to Abbring, Heckman, Chiappori, and Pinquet (2003), who suggest testing for selection by studying the relationship between behavior under a contract and subsequent amendments to that contract. In this paper, we focus on an amendment to a contract in the light of past behavior. To disentangle moral hazard and selection in our paper, we will look at the history of health care consumption by enrollees *prior* to their decision to choose a voluntary deductible. At that point, enrollees' behavior with respect to health care expenditure is not influenced by the subsequent uptake of deductible, which can be seen as a contract amendment in the context of Abbring et al. (2003).

3. INSTITUTIONAL CONTEXT AND RISK-ADJUSTMENT MODEL IN THE NETHERLANDS

In the Netherlands, health insurance is provided by private insurers that compete mainly on premiums. Since 2006, all citizens have been legally required to take out insurance in the form of a standardized basic benefits package, which is defined by the government. Insurers are obliged to offer this basic package in any health plan they offer. Community ratings are applied, meaning that insurers may not differentiate premiums among enrollees of the same health plan⁷⁰. Insurers can offer a voluntary deductible to its enrollees. Deductible options of €100, €200, €300, €400 and €500 are permitted. Here, too, insurers are not allowed to differentiate the price discounts associated with voluntary deductible options between enrollees of the same health plans. Note also, that the deductible option is added on top of the mandatory deductible (€350 in 2013), which was introduced to cope with moral hazard in health care consumption and to reduce public expenditure on health care. We analyze only the effect of optional voluntary deductible taking the mandatory deductible as given.

The insurance system is funded as follows. Approximately 50% of the total insurance revenue is raised from the premiums paid directly to the insurers and the out-of-pocket expenses falling under the deductibles. The other 50% is raised through an income-dependent premium determined by the government and collected by the tax office. The system of income-dependent premiums is meant to guarantee income solidarity and to keep insurance affordable. The core of this system is the risk-adjustment model: the tax office transfers the income-dependent premiums into the risk-adjustment fund, which in turn distributes these in the form of risk-adjusted capitation payments to the insurers.

Risk adjustment

The risk-adjustment model works at the level of individuals. The payments from the risk-adjustment fund to an insurer are based on a number of characteristics of the insurer's enrollees in order to compensate for differences in expected health costs of the enrollees. The used characteristics define risk classes related to the health status and other characteristics of an enrollee. These risk classes group individuals into health cost categories that are deemed predictive of their health care costs in the subsequent year.

The past health status of individuals is captured by diagnosis-cost classes for both physical and mental care, pharmacy-cost classes, and multi-year high-cost classes,

⁷⁰ Insurers may offer a discount on group contracts up to a maximum of 10%. In 2015, these discounts were on average 4.4%. Insurers with less than 850000 enrollees may also offer regional policies accessible only to inhabitants of a particular region.

based on the past health care consumption of the individual.⁷¹ The additional individual characteristics are captured by age-gender classes, income-source risk classes (benefit-receivers, self-employed and a rest category), socio-economic status risk classes (grouping individuals into three income levels and a separate category for enrollees who reside at an address with more than 15 others, for example a nursing home), region risk classes (grouping individuals into geographic clusters), and a risk class for one-person households.

The risk-adjustment model works as follows. Each year, the normative marginal cost values for each risk-adjustment class are determined by means of a regression of health care costs on the individual characteristics listed above. This estimation is done at the individual level. As a result, the expected health care cost of each enrollee can be estimated. The risk-adjustment fund is distributed among the insurers based on the predicted cost of their population minus an administrative premium that is set by the government.

The difference between the estimated costs as determined by the risk-adjustment fund and the actual health care cost of each individual equates to 'the risk-adjusted result' of the enrollee. The risk-adjusted result in the whole population is normalized to zero. Since the risk-adjustment model includes adjustment for mandatory deductible payments which the insurer receives directly from the enrollees, the variable 'risk-adjusted result' is adjusted for the profits and losses due to the mandatory deductible⁷².

Ideally, predictable health-related cost differences between enrollees should be fully eliminated by this system, leaving the insurers no incentive for selection. Significant effort is therefore devoted to improving the risk-adjustment model. Yet, not all predictable cost differences can be eliminated due to some private information, which leaves scope for selection.

Until 2012, the system also included significant ex-post additional compensations that applied when the predicted and realized health care costs diverged substantially. However, these ex-post adjustments have been phased out in recent years. Due to this decrease in risk-sharing, the incentive for selection has increased if the improvements in the risk-adjustment system are not sufficient to counteract it.

71 Consumption in the previous year except the multi-year high-cost classes where the consumption of the previous three years is considered.

72 In particular, a separate model within the risk-adjustment system predicts the amount of out-of-pocket payment related to the mandatory deductible that each enrollee would pay in the current year. This amount is subtracted from the predicted health care cost of the enrollee. However, the predicted out-of-pocket payment may diverge from the realized payment (just like there may be a difference between the predicted and realized health care costs) which means that the insurer may have a profit or a loss in this part of the system as well.

Health plans

In 2013, 10 insurers offered 67 different health plans. These health plans cover the same basic insurance package, but differ in some details with respect to both their coverage and pricing. While traditional health plans have only minor differences with respect to the choice of health care providers for enrollees, since 2008 there have also been health plans with a restricted network of health care providers (these plans require out-of-pocket payment for visiting non-contracted providers). The market share of these health plans was about 8% in 2015 (NZa, 2015).

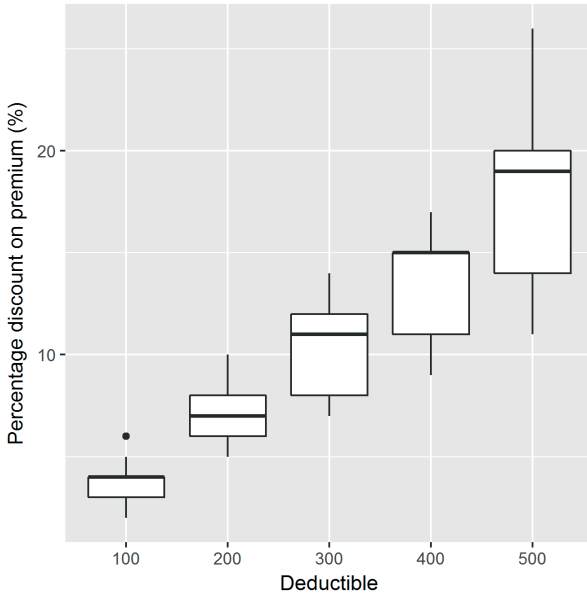
Each health plan offers discounts for voluntary deductibles. Enrollees who opt for a contract with a voluntary deductible benefit from premium discounts, depending on the size of the deductible chosen. A higher deductible is associated with a greater discount. In 2013, the average annual premium was €1269. The maximum voluntary deductible of €500 corresponded to an average discount of €230, which is 18% of the average premium. The situation was similar in the preceding years⁷³. Figure 1 illustrates the discrepancies in the deductible discount schedules over the different health plans, showing that there are quite large differences between health plans in terms of the premium discounts available for each deductible level.

Insurers can offer for each deductible level a discount on the health plan's premium. Insurers may not differentiate the discount among enrollees of the same health plan. In this graph, for each deductible level the average discount is calculated by taking the average of the health plans' discounts.

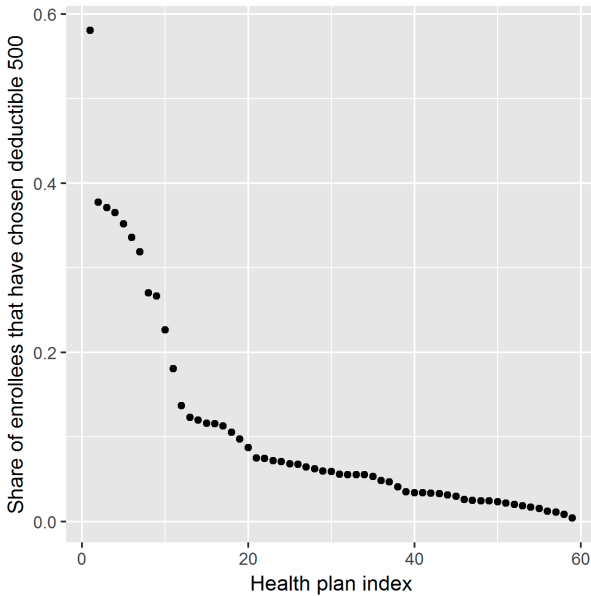
About 9% of enrollees over the age of 18 (the age at which the deductibles may be applied) chose a contract with a voluntary deductible in 2013, and the majority of them chose the maximum level of €500. The distribution of the share of enrollees who chose the maximum level of €500 in their health plans in 2013 is shown in Figure 2. We can observe large variation in the share of enrollees who chose the maximum deductible level of €500. Some health plans have a share that is lower than 5%, while one health plan has a share that is around 60%.

For each health plan, we calculated the share of enrollees that have chosen the maximum deductible level in 2013. We excluded health plans that have less than 1000 enrollees.

73 For the other voluntary deductible levels, the average discounts were €45 (€100 deductible), €88 (€200 deductible), €131 (€300 deductible) and €175 (€400 deductible).

Figure 1. Boxplot of discounts offered at different deductible levels in 2013

Insurers can offer for each deductible level a discount on the health plan's premium. Insurers may not differentiate the discount among enrollees of the same health plan. In this graph, for each deductible level the average discount is calculated by taking the average of the health plans' discounts.

Figure 2. Share of enrollees with deductible level €500 in 2013

For each health plan, we calculated the share of enrollees that have chosen the maximum deductible level in 2013. We excluded health plans that have less than 1000 enrollees.

4. EMPIRICAL STRATEGY

In this paper, we are interested to determine (i) how strong the relationship is between risk-type and the choice for deductibles (i.e. the degree of adverse selection) and (ii) the incentives for insurers to attract enrollees into healthplans with a voluntary deductible in terms of the risk-adjusted result. Note that the most important channel for a health insurer to attract enrollees into healthplans with a voluntary deductible is by offering substantial (community-rated) premium discounts.

To determine both the degree of adverse selection and the incentives for insurers, we perform the conditional correlation test that was proposed by Chiappori and Salanie (2000) and further developed by Chiappori, Jullien, Salanié, and Salanie (2006) and Finkelstein and McGarry (2006). Our analysis follows the application of this test by Aarbu (2017), who examined the presence of asymmetric information in the home insurance market.

In short, to determine the degree of adverse selection we estimate an OLS model on the individual enrollee level to determine if enrollees with a high deductible level have low health care costs. To determine the incentives for insurers to offer deductibles to enrollees, we also estimate an OLS model on enrollee level. However, in this model we determine if enrollees with a voluntary deductible have high risk-adjusted results compared to enrollees with no voluntary deductible.

A challenge with the conditional correlation test is to disentangle the moral hazard effect from the adverse selection effect. Finding that there is a relation between deductible level and health care costs can be explained by both adverse selection (the relationship between risk-type and the demand for deductible) and moral hazard (the hypothesis that enrollees who have chosen a deductible consume less health care because they face out-of-pocket payments). The same holds for the relation between deductible level and risk-adjusted results.

To disentangle moral hazard and adverse selection in our paper, we will look at the history of health care consumption by enrollees *prior* to their decision to choose a voluntary deductible. At that point, enrollees' behavior with respect to health care expenditure is not influenced by the subsequent uptake of deductible.

What follows is a more formal exposition of the conditional correlation test for the examination of adverse selection and the incentives for insurers.

The health care cost of enrollee i , cost_i depends on the health status and other characteristics of the enrollee, X_i . The deductible chosen by enrollee i , deductible_i , is also a function of characteristics of the enrollee. Therefore, we obtain a system of equations of the form

$$\text{cost}_i = g(X_i, \mu_i) \tag{1}$$

$$deductible_i = h(X_i, v_i) \quad (2)$$

where μ_i and v_i are the error terms from the cost equation and the deductible choice equation.

If there is no asymmetric information, and vector X contains all the relevant characteristics used by the insurer, then the error terms μ and v will be uncorrelated. However, if there is a variable with a positive impact on health care costs that is not included in the list of characteristics X , then the error term μ will pick up the effect that this variable has on health care costs, $cost$. According to the literature, see section 2, high-risk enrollees will self-select in a contract with more generous coverage. Hence, a higher value of μ will be observed together with a lower value of the deductible, and thus, with a lower value of v . A significant negative correlation between the error terms μ and v will, therefore, demonstrate the presence of asymmetric information.

Under a conditional independence assumption, this test can be conducted by using reduced-form OLS equation in which $cost$ is the dependent variable and $deductible$ is the independent variable:

$$cost_i = \alpha deductible_i + \beta X_i' + \gamma Z_i' + \epsilon_i \quad (3)$$

where $cost$ denotes the health care costs, $deductible$ is the choice of deductible, X is a vector of consumer characteristics, Z contains all other health plan related relevant variables, and ϵ is the error term. The letters α , β and γ are parameter vectors. We expect a negative relationship between the cost and the deductible. However, it is important to control for enrollees' risk aversion, since there may be a bias in the single-equation OLS model if (i) the choice of deductible is related to risk aversion and (ii) risk aversion is related to enrollee cost type (Finkelstein & McGarry, 2006; Aarbu, 2017). In our application we will include proxies for risk aversion based on several characteristics of the enrollees.

Since the current costs in each year contains both selection and moral hazard effects, one more step needs to be done in order to separate the selection effect from the moral hazard effect. To remove the moral hazard effect from the equation, we focus on the enrollees who had contracts without voluntary deductibles during three years before the year for which we conducted the estimation, year t . Additionally we replaced the health care costs incurred in that year with the health care costs of the previous year. This replacement results in the following specification for the reduced model, in which the dependent variable does not depend on the choice of a contract in year t :

$$cost_{i,t-1} = \alpha deductible_{it} + \beta X_{it}' + \gamma Z_{it}' + \epsilon_{it} \quad (4)$$

In the presence of risk-adjustment, the same reasoning also holds for the risk-adjusted result variable, *result*. Therefore, this selection effect could also be demonstrated by using an alternative model specification with the past *result* as the dependent variable and *deductible* as the independent variable, in which a higher risk-adjusted result would be associated with a higher deductible level chosen.

However, this approach could introduce selection bias in our estimate of selection. By focusing on enrollees who had contracts without deductibles in $t - 1$, $t - 2$ and $t - 3$, we have a subsample of the population with different characteristics compared to the whole population. If these characteristics affect costs or risk-adjusted results, then we have potential bias in our estimation. We are not able to quantify this bias, however, we can try to give a direction of the bias. van Winssen, van Kleef, and van de Ven (2015) showed that in the Netherlands for a large share of enrollees who has not chosen any voluntary deductible would have been better off if they would have chosen a voluntary deductible. This means that our subsample may include an over-representation of enrollees that are less interested and/or less shrewd when choosing their deductible. If this is the case, then we may underestimate the adverse selection effect.

5. DATA DESCRIPTION

Data and defining the relevant subset for the empirical analysis

The data came from two sources: risk-adjustment data and health plan choice data. Both datasets are panel datasets, covering the entire population of the Netherlands, which exceeds 16 million enrollees per year, over the period 2010-2013.

Both datasets were provided by the Dutch Healthcare Authority. The risk-adjustment dataset comprises the characteristics of the enrollees included in the risk-adjustment system and the actual costs incurred by the enrollees. The individual characteristics and the cost types included were described in detail in section 2, where we also defined the concept of 'risk-adjusted result'. The health plan choice dataset is a complementary dataset that includes the insurance enrollment. These records provide the health plan details on each enrollee, including the deductible level chosen. As explained, enrollees can opt for or zero voluntary deductible or a voluntary deductibles of €100, €200, €300, €400 or €500. To conduct the analysis, both dataset were merged at the enrollee level.

Table 1 provides an overview of our dataset coverage (in insured years) and the amounts of costs included in the dataset in billions of euros. Since physical and mental

Table 1. Size of the dataset

year	Insured Years mln	total physical health care costs bln euro	total mental health care cost bln euro
2010	16.3	25.958	3.010
2011	16.4	26.508	3.299
2012	16.5	27.837	2.982
2013	16.5	29.736	3.071

health care costs are the major costs of health care⁷⁴, the cost variable that we use in our analysis are defined as the sum of these two cost components⁷⁵.

The table excludes observations with missing values and the observations relating to individuals who reside abroad (approximately 1% of all observations). In the analysis that follows, we also excluded enrollees younger than 18 years, since this group of enrollees does not have to pay any deductible (this implies excluding 20% of the population). Furthermore, we only selected enrollees that appeared in our dataset every year, so that we could follow each enrollee over the whole period (this implied excluding 11% of the population).

In order to conduct the test described in section 4 for the year 2013, we selected enrollees who did not choose a voluntary deductible in the previous years (2010 till 2012) in order to estimate the equations for the health care costs and risk-adjusted results for the year before the enrollees chose a voluntary deductible (i.e. 2012). By looking at the outcomes for health, costs and risk-adjusted results for 2012, before the enrollee chose a deductible, we were able to exclude the possibility that the voluntary deductible had affected the health, costs and risk adjusted results that we were examining. Next section provides some descriptive analysis of the selected subset of enrollees.

Descriptive analysis of the selected subset

For the selected subset of enrollees, we consider the relationship between their deductible choice in 2013 and some variables of interest, such as the health status, costs, risk-adjusted result, and expected out-of-pocket expenses.

Table 2 shows the number of enrollees for each deductible choice in 2013, together with the share of healthy individuals in each subgroup. Here we classified an enrollee as 'healthy' if he or she was not included in any diagnosis-cost class, pharmacy-cost class or multi-year high cost class in 2012. These enrollees are considered as having no substantial health care costs in the period 2010-2012, which is deemed predictive of

⁷⁴ Some fixed cost components are not included in risk-adjustment.

⁷⁵ Note the difference between the risk-adjusted result and the cost variable: the cost variable that we use represents the incurred cost; this variable is neither adjusted nor normalized.

Table 2. For each deductible category the number of enrollees and the share of healthy in 2012. Only enrollees that have not chosen any voluntary deductible in the period 2010-2012 are taken into the calculation.

deductible 2013	enrollees	share enrollees classified as 'healthy'
000	10828669	0.57
100	25386	0.82
200	41442	0.83
300	23242	0.82
400	7925	0.82
500	265489	0.83

2013 health care costs according to the risk-adjustment model. The share of healthy enrollees among enrollees who did not choose any deductible in 2010-2012 and 2013 was 57%. The share of healthy enrollees was much higher among enrollees who did not choose a voluntary deductible in 2010-2012 but who did choose a voluntary deductible of €500 in 2013: 83%. This descriptive analysis suggests that enrollees who choose a higher deductible level are more likely to be healthy.

Table 3 compares the mean values of health care costs, risk-adjusted results and counterfactual €500-deductible out-of-pocket expenses for 2012 (before choosing a deductible) and 2013 (after choosing a deductible) for each deductible category in 2013. We will explain below how we calculated the counterfactual deductible out-of-pocket expenses. The columns on costs provide insight in the health care use by the individuals; the columns on risk-adjusted results show profitability of each subgroup in risk-adjustment; and the counterfactual expenses characterize the attractiveness of a higher deductible option from the enrollees' perspective.

First, looking at the columns headed 'mean cost' and 'mean result' in this table, we observe that the mean 2012 health costs (mean 2012 result) for enrollees who did not choose a voluntary deductible in either 2010-2012 or 2013 are higher (lower) than the mean 2012 health costs (mean 2012 result) for the enrollees who did not choose a voluntary deductible in 2010-2012 but did choose a voluntary deductible of €500 in 2013: €2,261.68 vs. €600.29 (€-26.05 vs. €399.578)⁷⁶. From this comparison, we can conclude that insurers have lower costs for enrollees who have chosen a voluntary deductible (higher risk-adjusted results). This applies to the year before they chose the voluntary deductible (2012) and it remains true for the year after they chose the voluntary deductible (2013). Although the costs in 2013 reflect not only the selection but also the moral

⁷⁶ The risk-adjusted result was normalized such that its mean is equal to zero. However, the mean results presented in the table differs from zero because we use a subset of the population in our analysis.

Table 3. For each deductible category the mean costs, mean risk-adjusted results and mean counterfactual voluntary deductible 500 out-of-pocket expenses in 2012 and 2013. Only enrollees who have not chosen any voluntary deductible in the period 2010-2012 are taken into the calculation.

deductible 2013	mean cost		mean result		mean expenses	
	2012	2013	2012	2013	2012	2013
000	2262	2487	-26	-27	467	452
100	772	906	286	257	254	243
200	643	810	384	314	231	228
300	653	796	405	349	232	227
400	627	826	445	362	232	230
500	600	703	420	400	227	203

Combining the findings from Tables 2 and 3, we can therefore conclude that enrollees who chose a voluntary deductible in 2013 are: healthier and have lower counterfactual out-of-pocket expenses than enrollees who did not choose a voluntary deductible in 2013. At the same time, insurers incur lower costs and have higher risk-adjusted results for enrollees who chose a voluntary deductible in 2013.

hazard effect, the lower costs and positive risk-adjusted results in 2012 are a strong indication for the presence of selection and also indicates that the selection effect is much larger than the moral hazard effect.

The last two columns of Table 3 headed 'mean expenses' show the mean counterfactual values of out-of-pocket expenses that would have been paid by the enrollees under the €500 deductible option. Since there is also a mandatory deductible, which was €350 in 2013, we simulated the counter-factual situation in which each enrollee had chosen a voluntary deductible of €500 on top of the mandatory deductible of €350. In this case, each enrollee would have to pay any costs incurred up to €850. In the Netherlands, the deductible applies to almost all health care costs, but there are some exceptions; for example, payments for general practitioners (GP) are not included. In our dataset, for each enrollee we knew the exact costs that were taken into account for the deductible. Using these 'deductible costs' and limiting them to €850 euro, we were able to calculate the total counterfactual out-of-pocket expenses that each enrollee would have had to pay if he or she had chosen the maximum voluntary deductible of €500 euro in 2013. We calculated these out-of-pocket expenses based on the 2012 'deductible costs' (that is, before the choice of any voluntary deductible for 2013) and the 2013 'deductible costs' (after the choice of a voluntary deductible for 2013).

Focusing on the column headed 'mean expenses 2012' in Table 3, which shows these counterfactual values for the year 2012, we can see that the mean counterfactual out-of-pocket expenses in 2012 for enrollees who did not choose a voluntary deductible in either 2010-2012 or 2013 are higher than the same figure for those enrollees who opted for a voluntary deductible of €500 in 2013: €466.53 vs. €227.21. This holds for the year

before they chose the voluntary deductible (2012), and a similar result holds for the year after they chose the voluntary deductible (2013).

Table 4 shows the summary statistics of the variables discussed above in Table 3. In addition, Table 5 provides insight on how the risk-adjusted result varies in relation to the choice of a deductible in 2013 by the subset of enrollees who did not choose a voluntary deductible in 2010-2012, which we use in our empirical analysis. The table shows both the mean and the distribution of the risk-adjusted result over percentile groups for each deductible category. For example, we can see that 2.05% of the enrollees that did not choose a deductible in 2013 are in the $[0th, 2th]$ percentile group, while only 0.46% of the enrollees who chose a deductible of €500 in 2013 are in the $[0th, 2th]$ percentile group. Table 5 also includes the 2012 mean risk-adjusted result for each deductible category. The mean risk-adjusted result in the $[0th, 2th]$ percentile group is €-27925.10, while the mean risk adjusted result in the $(10th, 100th]$ percentile group is €1081.41. Thus, compared to the other deductible categories, a relatively large share of the enrollees who did not have any voluntary deductibles in 2013 fall into the lowest result percentile group $[0th, 2th]$. This last observation suggests that enrollees without a voluntary deductible are more likely to be loss-making in risk-adjustment.

Table 4. Summary statistics for cost, counterfactual (cf) out-of-pocket expenses and risk-adjusted result in 2012 and 2013. Only enrollees who have not chosen any voluntary deductible in the period 2010-2012 are taken into the calculation.

Statistic	N	Mean	St. Dev.	Min	Max
cost 2012	11192153	2208	5589	0	645275
cost 2013	11192153	2431	6664	0	806561
result 2012	11192153	-12	4761	-627524	81361
result 2013	11192153	-14	5769	-780855	88989
cf expenses 2012	11192153	459	367	0	850
cf expenses 2013	11192153	444	367	0	850

6. ESTIMATION

In this section we estimate the relationship between (i) costs in the previous year and the choice of a deductible for the subsequent year and (ii) between the risk-adjusted result in the previous year and choice of a deductible for the subsequent year. These are the test of our two hypotheses derived from section 4. We have tried to control as much as possible for enrollee heterogeneity and health plan heterogeneity. Enrollee heterogeneity entails differences in enrollee characteristics, in particular those that are relevant to risk-preferences. Health plan heterogeneity includes differences at the level of the insurer's health plan, for example the possibility that a particular health plan

Table 5. For each deductible category the distribution of population percentiles. Only enrollees that have not chosen any voluntary deductible in the period 2010-2012 are taken into the calculation.

percentile	result mean	result SD	share of deductible in 2013					
			000	100	200	300	400	500
[0th,2th]	-27925	23382	0.021	0.006	0.005	0.005	0.005	0.005
(2th,4th]	-9138	1486	0.021	0.007	0.007	0.007	0.006	0.005
(4th,6th]	-5722	665	0.020	0.011	0.010	0.008	0.009	0.007
(6th,8th]	-3878	423	0.020	0.012	0.010	0.011	0.013	0.009
(8th,10th]	-2685	283	0.020	0.013	0.011	0.011	0.010	0.010
(10th,100th]	1081	1787	0.898	0.950	0.957	0.958	0.958	0.965

may have a more cost-effective way of purchasing health care (lower costs) than the average health plan. We estimated the following equations separately for the costs and risk-adjusted result:

$$cost_i = \alpha + \sum_{k=1}^{n=6-1} \beta_k deductible_{ki} + \sum_{k=1}^{n=7-1} \theta_k age_{ki} + \delta male_i + \sum_{k=1}^{n=3-1} \mu_k income_src_{ki} + \sum_{k=1}^{n=4-1} \lambda_k incom_{ki} + \sigma oneperson_hh_i + \sum_{k=1}^{n=59-1} \gamma_k healthplan_{ki} + \epsilon_i \quad (5)$$

$$result_i = \tilde{\alpha} + \sum_{k=1}^{n=6-1} \tilde{\beta}_k deductible_{ki} + \sum_{k=1}^{n=7-1} \tilde{\theta}_k age_{ki} + \tilde{\delta} male_i + \sum_{k=1}^{n=3-1} \tilde{\mu}_k income_src_{ki} + \sum_{k=1}^{n=4-1} \tilde{\lambda}_k incom_{ki} + \tilde{\sigma} oneperson_hh_i + \sum_{h=1}^{n=59-1} \tilde{\gamma}_h healthplan_{ki} + \tilde{\epsilon}_i \quad (6)$$

In model (5) $cost_i$ is the cost of enrollee i in 2012 and in model (6) $result_i$ is the risk-adjusted result for enrollee i in 2012. In both models, $deductible_{ki}$ is a dummy variable which is equal to 1 if enrollee i has chosen deductible category k in 2013 (leaving out category '000').

To control for possible heterogeneity in risk preference (i.e. risk aversion proxies), we included age dummy variables age_{ki} (7 categories), gender dummy variables $gender_{ki}$ (2 categories), income source dummy variables $income_src_{ki}$ (3 income source categories: benefit-receivers, self-employed and a rest category), average household income variables $incom_{ki}$ (4 categories), one-person household dummy variables $oneperson_hh_{ki}$ (2 categories). These variables are taken from the risk classes from the risk-adjustment system. See section 3 for a description of these variables. From the literature we know that there is a potential relation between gender, age, family status, income and occupation/education characteristics of individuals and their risk aversion (). However, note that the risk aversion proxies may also pick up other effects, such as health behavior or ability to pay. We have to keep this in mind when interpreting the effect of the risk aversion proxies on the estimations.

In both models dummy variable $healthplan_{ki}$ is equal to 1 if enrollee i has chosen healthplan k in 2013, which we added to control for potential unobserved factors that affect the risk-adjusted result and are related to the specific health plan that an enrollee has chosen in 2013. Lastly, α and $\tilde{\alpha}$ are the constants, and ϵ_i and $\tilde{\epsilon}_i$ are the error terms.

In our estimation, we did not include all riskclasses that is used in the risk-adjustment system. Our goal is not to predict as well as possible the health care costs of each enrollee. However, we do want to control for enrollees' risk aversion, since, as discussed in section 4, there may be a bias in the OLS model if (i) the choice of deductible is related to risk aversion and (ii) risk aversion is related to enrollee cost type. It should also be noted that the risk-adjusted result is the residual of the risk-adjusted payment and costs. Therefore the risk-adjusted result has already been 'adjusted' for the effect of the observable consumer characteristics that are included in the risk-adjustment.

We estimated models (5) and (6) with the ordinary least squares (OLS). Table 6 shows the results of the OLS estimation of the cost model (5) and Table 7 shows the results of the OLS estimation of the risk-adjusted result model 6. For both models, we estimated the model in steps: (i) only the deductible dummy variables, (ii) adding health plan fixed effects and (iii) adding the risk aversion proxies. As with our descriptive analysis in section 5.2, we estimated the model using the dataset of enrollees who chose no voluntary deductible in the period 2010-2012.

Table 6 econometrically confirms our conclusions from the descriptive analysis in section 5.2: enrollees who chose a voluntary deductible in 2013 had significantly lower costs in 2012 than enrollees who did not choose a voluntary deductible in 2013. For example, in the model with only voluntary deductible dummy variables, enrollees who chose a voluntary deductible of €500 in 2013 had, on average, costs that were €1,661.50 lower in 2012 than enrollees who did not choose a voluntary deductible in 2013. We found similar effects for the other deductible categories (€100, €200, €300 and €400). Adding fixed effects only slightly reduced the estimated decreases related to the choice of a voluntary deductible in 2013. When the risk aversion proxies were added to the model, we found that the decrease in costs associated with a voluntary deductible is lower and that there is a increase in the R^2 . However, the reduction in costs associated with having a voluntary deductible remains substantial. For example, enrollees who chose a voluntary deductible of €500 in 2013 had, on average, costs of €891.70 lower in 2012 compared to enrollees who did not choose a voluntary deductible in 2013.

Similarly, Table 7 confirms our conclusions from the descriptive analysis in section 5.2: enrollees who chose a voluntary deductible in 2013, had a significantly higher risk-adjusted result in 2012 than enrollees who did not choose a voluntary deductible in 2013. For example, in the model with only voluntary deductible dummy variables, enrollees who chose a voluntary deductible of €500 in 2013 had, on average, a risk-adjusted result in 2012 of €445.86 higher compared to enrollees who did not choose a voluntary

Table 6. OLS estimation of cost 2012 on deductibles 2013.

	Dependent variable: cost2012		
	(1)	(2)	(3)
Constant	2,261.756*** (1.696)	3,578.001*** (13.733)	3,391.542*** (20.448)
deductible_2013 100	-1,490.718*** (35.081)	-1,270.968*** (35.103)	-669.262*** (34.263)
deductible_2013 200	-1,618.965*** (27.475)	-1,482.874*** (27.486)	-852.448*** (26.834)
deductible_2013 300	-1,608.693*** (36.662)	-1,491.396*** (36.646)	-871.760*** (35.768)
deductible_2013 400	-1,634.726*** (62.731)	-1,514.096*** (62.673)	-957.012*** (61.157)
deductible_2013 500	-1,661.473*** (10.969)	-1,412.512*** (11.371)	-891.700*** (11.120)
age[30,40)			277.571*** (6.258)
age[40,50)			134.474*** (5.953)
age[50,60)			551.633*** (6.079)
age[60,70)			1,365.584*** (6.200)
age[70,70+)			2,834.631*** (6.341)
male			-200.765*** (3.277)
oneperson_hh			218.082*** (4.455)
incomegroup1			-1,717.903*** (15.044)
incomegroup2			-1,804.444*** (14.865)
incomegroup3			-2,018.974*** (15.066)
income_src benefitsreceiver			2,430.191*** (5.772)
income_src selfemployed			-387.520*** (7.473)
plan fixed effects	no	yes	yes
Observations	11,189,716	11,189,716	11,189,716
R ²	0.003	0.006	0.054
Adjusted R ²	0.003	0.006	0.054
Residual Std. Error	5,582.036 (df = 11189710)	5,572.685 (df = 11189652)	5,437.474 (df = 11189640)
F Statistic	6,078.346*** (df = 5; 11189710)	1,081.476*** (df = 63; 11189652)	8,466.543*** (df = 75; 11189640)

Note: * p<0.1; ** p<0.05; *** p<0.01.

Table 7. OLS estimation of risk adjusted result 2012 on deductibles 2013.

	Dependent variable: result 2012		
	(1)	(2)	(3)
Constant	-26.029*** (1.447)	-105.554*** (11.730)	-582.410*** (17.893)
deductible_2013 100	312.270*** (29.916)	300.549*** (29.984)	345.379*** (29.982)
deductible_2013 200	409.890*** (23.429)	403.887*** (23.478)	445.349*** (23.481)
deductible_2013 300	431.057*** (31.264)	430.744*** (31.302)	469.269*** (31.298)
deductible_2013 400	471.235*** (53.494)	471.247*** (53.534)	501.459*** (53.516)
deductible_2013 500	445.857*** (9.354)	446.350*** (9.713)	475.240*** (9.730)
age[30,40)			-15.396*** (5.476)
age[40,50)			66.972*** (5.209)
age[50,60)			104.224*** (5.320)
age[60,70)			149.685*** (5.425)
age[70,70+)			343.263*** (5.549)
male			56.165*** (2.867)
oneperson_hh			71.606*** (3.899)
incomegroup1			391.714*** (13.164)
incomegroup2			349.325*** (13.008)
incomegroup3			309.668*** (13.184)
income_src benefitsreceiver			-132.834*** (5.050)
income_src selfemployed			55.798*** (6.539)
plan fixed effects	no	yes	yes
Observations	11,189,716	11,189,716	11,189,716
R ²	0.0003	0.0003	0.001
Adjusted R ²	0.0003	0.0003	0.001
Residual Std. Error	4,760.164 (df = 11189710)	4,760.030 (df = 11189652)	4,758.072 (df = 11189640)
F Statistic	583.033*** (df = 5; 11189710)	57.188*** (df = 63; 11189652)	171.068*** (df = 75; 11189640)

Note: * p<0.1; ** p<0.05; *** p<0.01.

deductible in 2013. Again, we found similar effects for the other deductible categories (€100, €200, €300 and €400). Adding health plan fixed effects and risk aversion proxies only affect the results slightly.

7. DISCUSSION AND CONCLUDING REMARKS

Our empirical results point to the presence of selection under mandatory health insurance with open enrollment in a managed care setting. The uniquely rich dataset covering the entire population of the Netherlands over a period of several years allowed us to demonstrate the presence of the selection effect of deductibles separately from the potential moral hazard effect that arises simultaneously. From our analysis it follows that offering contracts with voluntary deductibles results in self-selection by healthier enrollees, who are overcompensated by the risk-adjustment system.

We observe the selection effect even after controlling for a large set of control variables for risk aversion, healthplan fixed-effects and possible other confounding effects. The expected gains on the risk-adjustment per enrollee with a voluntary deductible of €500 are estimated at around €450 on average. On top of this, the enrollee pays a larger share out-of-pocket, resulting in lower costs for the insurer to be reimbursed. In return, the insurer offers the enrollee a premium discount for taking a voluntary deductible.

It should also be noted that the literature on effects of voluntary deductibles shows that (i) enrollees are not always able to choose the optimal deductible options; (ii) enrollees are not always able to make the optimal choices in relation to these deductibles; (iii) there is little insight into the size of the reduction of the moral hazard effect on costs (because it is difficult to distinguish delayed costs from avoided costs and the reduction in excessive costs).

Thus, our paper provides convincing evidence that individuals have substantial private information that are not captured by the risk-adjustment system. When insurers have the possibility or obligation to diversify their health plans in terms of generosity, individuals can self-select into health plans. For health plan generosity in terms of deductibles, we show that this self-selection results in substantial differences in costs which are only partly compensated by the risk-adjustment system.

Given the substantial expected gains on the risk-adjustment for enrollees with a voluntary deductible, it is important to study in greater detail how these gains are related to the premium discount for voluntary deductible that insurers set. For example, it would be interesting to determine if insurers are able to attract enrollees with large gains from risk-adjustment by offering large premium discounts. However, it may also be the case that by offering higher discounts the insurers may attract on the margin enrollees that are relatively high-risk. If we find that not all gains are translated into premium

discounts, it would be interesting to examine if this is due to the current regulation that stipulate that insurers are not allowed to differentiate the price discounts associated with voluntary deductible options between enrollees of the same health plans.

An option that the Dutch government could consider to fully eliminate the risk-adjustment gain of the deductibles is to include the choice of a voluntary deductible in the risk-adjustment system as one of the characteristics of the consumer. In this way, consumers with a voluntary deductible are not profitable any more in the risk-adjustment system, which decreases the incentive for selection. This may lead to a lower number of enrollees opting for a voluntary deductible. Consequently, the potential moral hazard effect would be lower too. Because this potentially results in a trade-off between improving solidarity (due to better risk adjustment) and cost reduction (due to moral hazard), more research in this area is definitely needed to understand the potential welfare effects.

REFERENCES

- Aarbu, K. O. (2017). Asymmetric information in the home insurance market. *Journal of Risk and Insurance*, 84 (1), 35-72.
- Abbring, J. H., Heckman, J. J., Chiappori, P.-A., & Pinquet, J. (2003). Adverse selection and moral hazard in insurance: Can dynamic data help to distinguish? *Journal of the European Economic Association*, 1 (2-3), 512-521.
- Aron-Dine, A., Einav, L., & Finkelstein, A. (2013). The rand health insurance experiment, three decades later. *The Journal of Economic Perspectives*, 27 (1), 197-222.
- Bijlsma, M., Boone, J., & Zwart, G. (2014). Competition leverage: how the demand side affects optimal risk adjustment. *The RAND Journal of Economics*, 45 (4), 792-815.
- Bolhaar, J., Lindeboom, M., & van der Klaauw, B. (2012). A dynamic analysis of the demand for health insurance and health care. *European Economic Review*, 56 (4), 669-690.
- Brot-Goldberg, Z., Chandra, A., Handel, B., & Kolstad, J. (2015). What does a deductible do? the impact of cost-sharing on health care prices, quantities, and spending dynamics (Tech. Rep.). National Bureau of Economic Research.
- Chiappori, P.-A., Jullien, B., Salanie, B., & Salanie, F. (2006). Asymmetric information in insurance: General testable implications. *The RAND Journal of Economics*, 37 (4), 783-798.
- Chiappori, P.-A., & Salanie, B. (2000). Testing for asymmetric information in insurance markets. *Journal of political Economy*, 108 (1), 56-78.
- Cohen, A., & Siegelman, P. (2010). Testing for adverse selection in insurance markets. *Journal of Risk and Insurance*, 77 (1), 39-84.
- Cutler, D. M., Zeckhauser, R. J., et al. (1997). Adverse selection in health insurance (Tech. Rep.). National Bureau of Economic Research Cambridge, Mass.
- Dardanoni, V., & Donni, P. L. (2012). Incentive and selection effects of medigap insurance on inpatient care. *Journal of Health Economics*, 31 (3), 457-470.
- Davis, K., Doty, M. M., & Ho, A. (2005). How high is too high? Implications of high-deductible health plans (publication). Commonwealth Fund.
- Einav, L., & Finkelstein, A. (2011). Selection in insurance markets: Theory and empirics in pictures. *The Journal of Economic Perspectives*, 25 (1), 115-138.
- Finkelstein, A., & McGarry, K. (2006). Multiple dimensions of private information: Evidence from the long-term care insurance market. *The American Economic Review*, 938-958.
- Fronstin, P., & Collins, S. R. (2008). Findings from the 2007ebri/commonwealth fund consumerism in health survey. EBRI Issue Brief (315).
- Galbraith, A. A., Ross-Degnan, D., Soumerai, S. B., Rosenthal, M. B., Gay, C., & Lieu, T. A. (2011). Nearly half of families in high-deductible health plans whose members have chronic conditions face substantial financial burden. *Health Affairs*, 30 (2), 322-331.
- Gern, M., & Schellhorn, M. (2006). Nonparametric bounds on the effect of deductibles in health care insurance on doctor visits-swiss evidence. *Health economics*, 15 (9), 1011-1020.
- Handel, B. R. (2013). Adverse selection and inertia in health insurance markets: When nudging hurts. *The American Economic Review*, 103 (7), 2643-2682.
- Keeler, E. B. (1992). Effects of cost sharing on use of medical services and health.
- Lohr, K., Brook, R., Kamberg, C., Goldberg, G., Leibowitz, A., Keesey, J., . . . Newhouse, J. (1986). Use of medical care in the rand health insurance experiment: diagnosis-and service-specic analyses in a randomized controlled trial. *Medical care*, 24 (9), S1-S87.

- Newhouse, J. (2004). Consumer-directed health plans and the rand health insurance experiment. *Health Affairs*, 23 (6), 107-113.
- Newhouse, J., Price, M., Hsu, J., McWilliams, J., & McGuire, T. (2015). How much favorable selection is left in medicare advantage? *American journal of health economics*.
- Newhouse, J., Price, M., Huang, J., McWilliams, J., & Hsu, J. (2012). Steps to reduce favorable risk selection in medicare advantage largely succeeded, boding well for health insurance exchanges. *Health Affairs*, 31 (12), 2618-2628.
- NZa. (2014). Zorgverzekeringsmarkt 2014: Weergave van de markt 2010-2014 (Tech. Rep.).
- NZa. (2015). Marktscan zorgverzekeringsmarkt 2015 deel a (Report). Nederlandse Zorgautoriteit.
- NZa. (2016). Kwantitatief onderzoek naar risicoselectie en risicosolidariteit op de zorgverzekeringsmarkt (Report). Nederlandse Zorgautoriteit.
- Olivella, P., & Vera-Hernandez, M. (2013). Testing for asymmetric information in private health insurance. *The Economic Journal*, 123 (567), 96-130.
- Outreville, J. F. (2014). Risk aversion, risk behavior, and demand for insurance: A survey. *Journal of Insurance Issues*, 158-186.
- Panthöfer, S. (2016). Risk selection under public health insurance with opt-out. *Health economics*, 25 (9), 1163-1181.
- Rothschild, M., & Stiglitz, J. (1976). Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *The Quarterly Journal of Economics*, 90 (4), 629-649.
- Tollen, L. A., Ross, M. N., & Poor, S. (2004). Risk segmentation related to the offering of a consumer-directed health plan: A case study of humana inc. *Health services research*, 39 (4p2), 1167-1188.
- van de Ven, W. P., & Ellis, R. P. (2000). Risk adjustment in competitive health plan markets. *Handbook of Health Economics*, 1 (Chapter 14), 755-845.
- van Kleef, R., Beck, K., van de Ven, W., & van Vliet, R. (2008). Risk equalization and voluntary deductibles: a complex interaction. *Journal of Health Economics*, 27 (2), 427-443.
- van Kleef, R., McGuire, T., van Vliet, R., & van de Ven, W. (2017). Improving risk equalization with constrained regression. *The European Journal of Health Economics*, 18 (9), 1137-1156.
- van Winssen, K., van Kleef, R., & van de Ven, W. (2015). How pro table is a voluntary deductible in health insurance for the consumer? *Health Policy*, 119 (5), 688-695.
- van Winssen, K., van Kleef, R., & van de Ven, W. (2016). A voluntary deductible in health insurance: the more years you opt for it, the lower your premium? *The European Journal of Health Economics*, 1-18.
- Vektis. (2017). Zorgthermometer verzekerden in beeld 2017 (Report). Author.
- Wharam, J. F., Landon, B. E., Galbraith, A. A., Kleinman, K. P., Soumerai, S. B., & Ross-Degnan, D. (2007). Emergency department use and subsequent hospitalizations among members of a high-deductible health plan. *Jama*, 297 (10), 1093-1102.

Chapter 6

Is adverse selection effectively mitigated by consumer inertia? Empirical evidence from the Dutch health insurance market

With Frederik Schut and Marco Varkevisser

ABSTRACT

This paper examines whether consumer inertia can effectively reduce adverse selection in health insurance markets. To this end, we investigate consumer choice of deductible in the Dutch health insurance market, using panel data on all insured individuals in the Netherlands over period 2010-2013. The Dutch health insurance market offers a unique setting to study adverse selection, because during annual open enrollment periods all adults are free to choose an extra deductible up to 500 euros per year. By focusing on deductible choices of those who do not switch health plans, we are able to examine the 'pure' adverse selection effect (i.e. not distorted by other health plan attributes). We estimate a logit model to determine how previous and future costs are related to individuals' deductible choice. We find that almost 70% of the individuals with the highest deductible level are likely to stick with this choice even after incurring very high costs during the previous year. Furthermore, only 3.5% of those having no deductible and very low costs are likely to choose a 500 euro deductible next year. These findings suggest that consumer inertia can act as a powerful brake on adverse selection.

1. INTRODUCTION

The presence of adverse selection is a well-known impediment to an efficient health insurance market (Einav and Finkelstein 2011). Adverse selection occurs when enrollees choose health plans with more coverage because they have private information about being likely to incur high costs. Rothschild and Stiglitz (1976) show that adverse selection can result in the under-insurance of low-risk enrollees or even a market with no equilibrium. As shown by Handel (2013) and Handel et al. (2015), however, adverse selection may in practice be counteracted by individuals' suboptimal decision making. There is ample empirical evidence that optimal consumer choice in health insurance markets is hampered by "frictions" like inertia, search and switching cost, and a lack of knowledge ("health insurance literacy"); see for example Samuelson and Zeckhauser (1988), Abaluck and Gruber (2011), Bhargava et al. (2015), Handel (2013), Handel and Kolstad (2015), Handel et al. (2015), Heiss et al. (2016), Ho et al. (2015), and Marzilli Ericson (2014).

Beforehand, it is not clear whether adverse selection or friction is more important to market outcomes and how these two factors interact with each other. As Pauly (1984) noticed already more than three decades ago: "One of the things that theory does say here is that only a little bit of adverse selection may cause market equilibrium to unravel. But then only a little bit of consumer inertia is needed to reinstate it."

The interaction of adverse selection and consumer inertia has recently been studied in the context of US markets for health insurance (Handel 2013, Handel and Kolstad 2015, Handel et al. 2015, Polyakova 2016). Handel and Kolstad (2015) measure inertia as the implied monetary costs of switching plans when a default option is present. They identify inertia by comparing health plan choice of the same consumers over time in both clearly active and clearly passive choice environments. In the context of an employment-based insurance setting of a large US firm, they show that both adverse selection and inertia are important. Furthermore, they show that reducing frictions is welfare decreasing (increasing) when the mean and variance of surplus from risk protection compared to its costs are relatively low (high).

In this paper, we aim to determine to what extent adverse selection is mitigated by consumer inertia. In our setting – the Dutch market for mandatory basic health insurance – the choice environment is stable, since we focus on individuals staying with the same health plan during the study period. These individuals, however, can freely adapt their choice of deductible (i.e. coverage level) resulting in a lower or higher premium. Hence, no health plan attributes other than the deductible level and its corresponding premium difference play a role in the consumer choices examined here. Given that only monetary trade-offs are involved in this choice setting, we are able to truly identify inertia as defined by Handel (2013); i.e. the implied monetary costs of choice

persistence. Using detailed data for the total Dutch adult insured population who were enrolled in the same health plan during the entire study period 2010-2013 (about 9.5 million individuals), we first constructed 16 possible deductible choice paths to examine the relationship between health care costs and deductible choice. In a second step we used the same dataset for estimating a logit model. The model estimates reveal to what extent individuals' choice of deductible in 2013 can be explained by their previous and future health care cost. Although we find clear evidence of adverse selection, we also find that the extent of adverse selection is strongly mitigated by the presence of substantial consumer inertia.

The remainder of this paper is organized as follows. In the next section, we briefly describe the context of the Dutch health insurance market in which people annually have free choice of deductible. Section 3 informs about the data and descriptive statistics. In section 4, deductible choice paths are defined and analyzed. The empirical model is formulated in section 5, after which the estimation results are presented in section 6. Section 7 concludes.

2. CONTEXT

In the Netherlands, universal mandatory health insurance is offered by competing private health insurers.⁷⁷ During our study period (2010-2013), the number of insurers decreased from 11 to 9 due to two mergers, whereas the number of basic health plans (or health insurance policies) offered by these insurers increased from 57 to 67. All Dutch citizens are required to buy a basic health plan and health insurers have to accept all individuals applying for enrollment in such a plan. The basic benefit package is comprehensive and standardized by law. Hence, each basic health plan covers the same benefits. In addition, health plan premiums have to be community-rated. That is, all people enrolling in the same health plan face the same premium (except that in case of a group contract insurers are allowed to offer a premium discount up to 10%).⁷⁸ For all adult enrollees (18 years and older) there is a mandatory deductible. Its level is

77 For a more detailed description of the Dutch health insurance market and the system of regulated (or managed) competition, see Van de Ven and Schut (2008), and Douven et al. (2017).

78 For each enrollee insurers receive a risk-adjusted premium subsidy from a risk-equalization fund that is filled with income-related contributions. This risk-adjusted premium subsidy is equal to the enrollee's predicted costs minus a fixed amount that is annually determined by the government (about 1,000 euro per year). Hence, to break even health insurers must at least charge a community-rated premium equal to this fixed amount (McGuire and Van Kleef, 2018).

annually set by the government.⁷⁹ On top of this mandatory deductible, adults can opt for a (voluntary) deductible in return for a premium discount. The deductible levels are restricted by the government to 0, 100, 200, 300, 400 or 500 euro per year. For each deductible level, health insurers are free to determine a community-rated premium discount. During our study period the cost of maternity care and family care (provided by GPs) were exempted from the deductible.

Each year, individuals can switch health plans during the annual open enrollment period (December/January). Health plans differ from each other in terms of premium, (preferred) provider networks and premium discounts for the various deductible levels. Enrollees can adjust the deductible level every year by notifying their health insurer during the open enrollment period. This typically requires only one phone call or ticking another box at the insurer's website. Changing deductible levels does not require changing health plans. Hence, after having increased the deductible level people can easily lower it again during the next open enrollment season if they have acquired a chronic disease or otherwise expect higher medical costs in the year(s) to come.

As mentioned in Section 1, depending on the fundamentals of the market, having substantial consumer inertia in the health insurance market could be welfare increasing or decreasing (Handel et al. 2015). In the Netherlands, there is a sophisticated system of risk equalization, which reduces the potential negative welfare effects of adverse selection (Van Kleef et al. 2017). Further, as described above, the premium discount in return for a deductible is community rated, thereby reducing the opportunities for insurers to attract favorable risks (relative to the risk adjusted capitation payments) by higher discounts. Nevertheless, even after sophisticated risk equalization those opting for the highest deductible level appear to be profitable to insurers at the prevailing discount levels (Croes et al. 2017). Increasing the uptake of voluntary deductibles by reducing consumer inertia may reduce profitability for insurers if people with a higher risk profile opt for it, or if it results in higher premium discounts because of enhanced competition. However, the welfare implications of stimulating the uptake of voluntary deductibles require further empirical research.

3. DATA AND DESCRIPTIVES

We use individual level panel data on the total Dutch population over the period 2010-2013 obtained from the Dutch Healthcare Authority (NZa). The dataset includes information on (i) individual health care expenses – including out-of-pocket costs – for

⁷⁹ During our study period the mandatory deductible was gradually raised from 165 euro in 2010 to 170 euro in 2011, 220 euro in 2012, and 350 euro in 2013.

benefits covered by mandatory health insurance, (ii) individuals' choice of health plan and deductible level, and (iii) a number of individual characteristics. In our analysis, we included age (5 classes), gender and household income (3 classes) as potentially relevant individual characteristics. In addition, we obtained data from the Dutch Healthcare Authority (NZa) about the community-rated premium discount offered by each health plan in return for a higher deductible. Throughout the study period, the average discount was 222 euro for a 500 euro deductible with a standard deviation of about 40 euro. We linked the information in the dataset about each individual's health plan to the health plan specific discount.

The full dataset includes about 16.6 million individual records in 2010. After excluding individuals that were not in the dataset for all four years, the number of observations reduced to 15.9 million. We also excluded a small number of people with more than one health plan per year (for example because they enrolled in another group contract after changing jobs), and people with incomplete information on household income (both groups comprise less than 1% of the total population). Furthermore, we excluded all individuals younger than 18 years in 2010, since these individuals did not face any deductible during part or the entire study period. The remaining dataset includes about 11.8 million individual records with data for each of the four years.

For this balanced panel, Table 1 shows the distribution of individuals over the various deductible levels. The proportion of enrollees that opted for a deductible other than zero increased from 5% to 9% from 2010 to 2013. This change is almost entirely due to the increasing share of individuals choosing the deductible of 500 euro (from 2% to 5%).

Table 1. Distribution of enrollees over deductible levels, 2010-2013 (N x 1000).

Deductible level	2010		2011		2012		2013	
	N	share	N	share	N	share	N	share
000	11226	0.95	11172	0.94	11097	0.94	10816	0.91
100	174	0.01	167	0.01	166	0.01	170	0.01
200	102	0.01	101	0.01	107	0.01	136	0.01
300	77	0.01	86	0.01	92	0.01	80	0.01
400	16	0.00	17	0.00	17	0.00	24	0.00
500	227	0.02	279	0.02	344	0.03	596	0.05
Total	11822		11822		11822		11822	

As explained above, from the perspective of adverse selection we are primarily interested in the question to what extent enrollees' choice of deductible is related to their past, current and expected future health costs. Enrollees who switch health plans, are

probably also reconsidering their deductible choice during this switching process.⁸⁰ In our dataset, 46% of all enrollees who changed their deductible, also switched health plans. The decision to switch health plans depends on multiple factors (Boonen et al. 2016). Because these factors may be correlated with the choice of deductible, we restricted our analysis to those enrollees that did not switch health plans during the study period. For this subsample, it is most likely that the change of deductible level is driven by past or anticipated health care expenses. Furthermore, to keep the analysis concise we also excluded the small minority of people who chose intermediary deductible levels (100-400 euro).⁸¹ As a result the final sample consist of about 9.4 million individuals who in the period 2010-2013 (i) did not switch health plans and (ii) each year have either a zero or a 500 euro deductible. Given the two deductible choice options (0 and 500) and 4 years (2010-2013) we can distinguished 16 possible deductible choice paths. Each individual can be assigned to one of these paths. Table 2 exhibits the distribution of individuals over the 16 possible choice paths.

As shown in Table 2 the vast majority (96%) of all enrollees included in the dataset sticks with the default option of a zero deductible during the entire period. The second largest subgroup (1%) includes the enrollees who stick with a once chosen 500 euro deductible. The remaining 3% changed their deductible at least once during the study period.

When calculating each individual's health care costs, we excluded the costs of GP and maternity care because these health care services are exempted from the deductible. The average costs per individual increased from 2,085 euro in 2010 to 2,508 euro in 2013 during the study period. As typically is the case, the distribution of individual health care costs is highly skewed. For instance, in 2013 25% of all enrollees had costs below 45 euro while 50% and 87.5% had costs below 477 euro and 4464 euro, respectively. Therefore, we transformed the cost data by taking the natural logarithm of one plus cost, i.e. $\log(1+cost)$. Figure 1 shows that the resulting log transformed cost distribution has the familiar bimodal shape, with local maxima at 0-0.5 (about zero euro) and 6.5-7 (about 900 euro). The log transformed cost distribution is convenient for our application, since we want to categorize enrollees into cost classes.

For the empirical analysis, we distinguished 6 different log transformed costs categories. Table 3 displays the proportion of people with either a zero or 500 euro deductible in 2013 for each of the 6 cost categories in 2012. As expected, the share of people opting

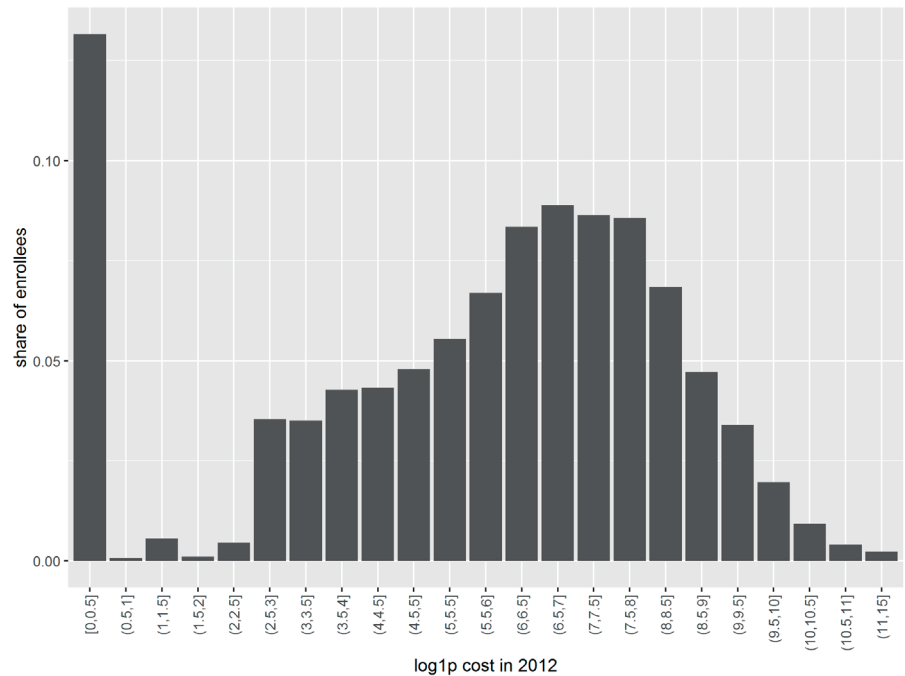
80 See Boonen et al. (2016) for a study of the factors that impact the propensity to switch health insurers in the Dutch health insurance market.

81 Note that – in contrast to intermediary deductible levels – the share of enrollees who have chosen the highest (500 euro) deductible level has grown significantly from 2010 to 2013. Furthermore, this growth continued after 2013 (Vektis 2017).

Table 2. Distribution of individual enrollees in the final sample over the possible deductible choice paths (2010-2013)

Path Nr.	Deductible level 2010-2011-2012-2013	N	share
1	500_500_500_500	134977	0.01
2	500_500_500_000	5471	0.00
3	500_500_000_500	1279	0.00
4	500_500_000_000	5969	0.00
5	500_000_500_500	1325	0.00
6	500_000_500_000	220	0.00
7	500_000_000_500	4524	0.00
8	500_000_000_000	8359	0.00
9	000_500_500_500	20838	0.00
10	000_500_500_000	1546	0.00
11	000_500_000_500	589	0.00
12	000_500_000_000	2130	0.00
13	000_000_500_500	26329	0.00
14	000_000_500_000	2245	0.00
15	000_000_000_500	121941	0.01
16	000_000_000_000	9025461	0.96
Total sample		9363203	1.00

Figure 1. Distribution of log transformed individual health care expenses in 2012



for the highest deductible in 2013 is negatively related to their health care expenses in the previous year. In the lowest cost category 7.8% of the individuals opted for a 500 euro deductible, whereas this holds for only 0.3% of the individuals in the highest cost category. However, note that even among those in the lowest cost category more than 90% of the individuals preferred not to have an extra deductible.

Table 3. Number and percentage of individuals with a zero and a 500 euro deductible in 2013 for each 2012 log transformed cost class

$\log(1+\text{cost}2012)$	Number of individuals		Percentage individuals
	deductible 000	deductible 500	with deductible 500
[0,2]	1200177	101563	7.8%
(2,4]	1033134	70097	6.4%
(4,6]	1921228	79647	4.0%
(6,8]	3174945	49024	1.5%
(8,10]	1575631	11017	0.7%
(10,15]	146286	454	0.3%
Total	9051401	311802	3.3%

The high percentage of people with low costs choosing a zero deductible indicates that many people may not make a (financially) optimal deductible choice. Indeed, as shown by Van Winssen et al. (2015) an uptake of a 500 euro deductible would have been retrospectively financially profitable for 48% of the Dutch insured population in 2014, whereas only about 11% actually acted in this way.⁸² We calculated the share of people that with hindsight made the most profitable deductible choice. We find that, given their actual health care expenses, about 50% of the people with a zero deductible made the best choice during the study period (varying between 47 and 54%).⁸³ By contrast, 82 to 85% of the people with a 500 euro deductible with hindsight made the financially optimal decision. These findings are similar to those of an earlier study by Douven and Remmerswaal (2016). Of course, an ex-post non-profitable deductible choice does not necessarily mean this choice is also non-profitable ex-ante, since the ex-ante profitability depends on the distribution of risk as well as the existing risk preferences.⁸⁴ Therefore the key question is: do people, from a financial perspective, make the optimal deduct-

82 Handel et al. (2015) find similar differences in a US employment-based insurance context, where, based on ex post spending 60% of employees are better off financially in a high deductible health plan, though only 15% of employees actually choose that plan.

83 Individuals with a zero deductibles who retrospectively made a suboptimal choice, could have gained on average about 200 euros if they had opted for a 500 euro deductible.

84 Note that the rationale of insurance is that ex-post the majority of the insured make a loss because their premiums are used to pay the claims of an unlucky minority.

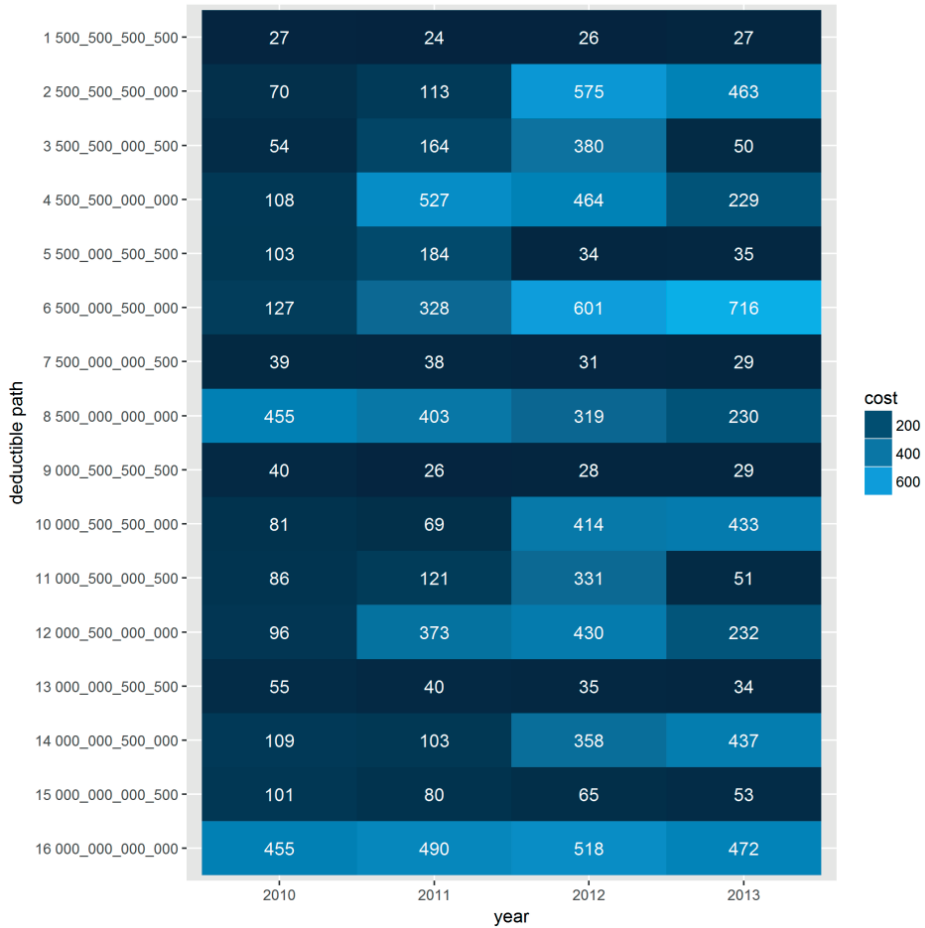
ible choice given their expected health care expenses? Before exploring this question in-depth, we first examine enrollees' annual deductible choices in relation to their previous and future health care expenses.

4. DEDUCTIBLE CHOICE PATHS

To examine the relation between enrollees' deductible choices and their health care costs, we have distinguished all possible deductible choice paths over the period 2010-2013. For each of these paths, the aggregate annual median cost are then calculated. The resulting median cost patterns for each of the 16 deductible choice paths are presented in a heatmap (Figure 2). The first eight paths include all enrollees with a 500 euro deductible in 2010, while the last eight paths include all enrollees with a zero deductible in 2010. The heatmap confirms our initial expectation. Enrollees in path 1 () consistently have the lowest annual median costs. In contrast, enrollees in path 16 () in each year have among the highest median costs compared to the other deductible paths. These median cost patterns are consistent with the presence of adverse selection: high (low) risks sorting themselves into the low (high) deductible plans.

Of particular interest are the choice paths of enrollees who experience a substantial increase in health care expenses over the years. In Figure 2, this involves the paths with a profound change of color. As the heatmap shows, for all the choice paths in which enrollees experience a strong increase in costs – i.e. choice paths 2, 4, 6, 10, 12, and 14 – we observe that the cost jump is followed by a change in deductible from 500 to zero in the next year. This is also consistent with the presence of an adverse selection effect. Enrollees following choice paths 3, 5 and 11 exchanged a zero for a 500 euro deductible in the only year they had higher health care cost (and changed back in the subsequent year). This is both consistent with adverse selection (choosing a zero deductible in anticipation of health care expenses) and moral hazard (a zero deductible resulting in higher health care expenses). Choice paths 7, 9, 13, and 15 show that enrollees with a zero deductible who are experiencing low health costs over the entire period eventually opt for a 500 euro deductible. This demonstrates that adverse selection may not take place immediately because consumers may learn more over time about their health risk – as well as the corresponding health care expenses – and the available deductible choice options.

In sum, all the possible choice paths seem to have cost patterns that are consistent with the presence of adverse selection. The patterns also suggest that (i) adverse selection may take time to arise and (ii) some healthy people are able to effectively anticipate an increase in next year's health care costs.

Figure 2. Heatmap of annual median cost (in euros) per deductible choice pattern over the period 2010-2013

5. EMPIRICAL MODEL

We estimate a logit model to examine the determinants of individual choice of deductible (0 or 500) for the year 2013 based on individual level data over the period 2010-2013. The logit model specifies the probability that individual i chooses a 500 euro deductible in 2013 as:

$$\pi_i = \frac{\exp(\mu_i)}{1 + \exp(\mu_i)}$$

where μ_i denotes the set of relevant characteristics for individual i , which can be specified as follows:

$$\mu_i = c + \beta deductible_{2012,i} + \sum_{k=1}^5 \rho_k cost_{2012,ki} + \sum_{k=1}^5 \sigma_k cost_{2013,ki} + \sum_{k=1}^5 \omega_k cost_{2012,ki} deductible_{2012,i} + \sum_{k=1}^5 \varphi_k cost_{2013,ki} deductible_{2012,i} + \delta stablelowcost_i + \sum_{k=1}^2 \gamma_k income_{ki} + \sum_{k=1}^4 \theta_k age_{ki} + \vartheta gender_i + \tau discount_i$$

The main variables of interest are *deductible2012*, *cost2012*, *cost2013* and *stablelowcost* because these four variables may provide evidence of adverse selection. The dummy variable *deductible2012* is equal to one if the individual had a 500 euro deductible in 2012. If there is choice persistence, we expect that having a 500 euro deductible in 2012 has an positive impact on the probability of choosing a 500 euro deductible in 2013. *Cost2012* and *cost2013* represent an individual's health care cost in 2012 and 2013, categorized in 6 categories ($k=1,..,6$), which are included as dummy variables in the regression (excluding the lowest cost class as reference dummy). If adverse selection is present, we expect to find a negative relationship between the uptake of a 500 euro deductible in 2013 and the individual health care expenses in 2012. Since people may not only base their expectations on last year's health care costs but also on expected future expenses, we also included *cost2013* as an indicator of adverse selection. It should be noted, however, that this *cost2013* variable may be endogenous, since these costs may depend on the choice of deductible because of moral hazard. Hence, the effect of *cost2013* on the choice of deductible may be overestimated. For this reason, we estimated the model with and without the *cost2013* variable. We also interacted both cost variables with the dummy variable *deductible2012*. This is because we expect that among individuals having a 500 euro deductible in 2012 those who experience (or expect) high costs are less likely to opt again for a 500 euro deductible in 2013. Hence, if these interaction effects are negative for higher cost categories, this will provide further evidence of adverse selection. For individuals who have chosen zero deductible in 2010, 2011 and 2012, we included the variable *stablelowcost* to investigate the possibility of a delayed adverse selection effect due to learning. We have defined *stablelowcost* as a dummy variable taking the value of 1 if (i) the individual's health care expenses did not exceed the mandatory deductible for three consecutive years (165 euro in 2010, 170 euro in 2011 and 220 euro in 2012) and (ii) if the individual has chosen zero deductible in 2010, 2011 and 2012. If there is a delayed adverse selection effect because for some individuals learning takes time, we expect to find a positive coefficient for this variable. Note that this variable also provides an indication of consumer inertia, as a positive coefficient implies that people are incurring monetary costs of choice persistence.

To control for risk preferences, we include for each individual his/her household income (*income*) (three categories, with the lowest income category as reference dummy), age (5 dummies, with the youngest age category as reference dummy), and *gender* (male=1). Furthermore, to control for health plan specific differences in the monetary

value of a 500 euro deductible, for each enrollee we include his specific health plan's community rated premium *discount* for a 500 euro deductible in 2013.

6. ESTIMATION RESULTS

Table 4 presents the estimation results for two models, one excluding and one including 2013 cost variables.⁸⁵ First, the results provide evidence of a high degree of choice persistence. Individuals already having a 500 euro deductible in 2012 are very likely to opt for the same deductible in 2013. The weighted average probability of choosing a 500 euro deductible in 2013 is 85 percentage points higher for individuals already having a 500 euro deductible in 2012 than for individuals having a zero deductible in 2012.⁸⁶

Next, there is a significant negative relation between the propensity of choosing a deductible in 2013 and the costs in the previous and current year. This result provides evidence of adverse selection.⁸⁷ Moreover, the interaction effects between cost and deductible choice show that this relationship is particularly strong for people with a 500 euro deductible in 2012.⁸⁸ Below, we provide a detailed discussion and illustration of the marginal effects of prior and current year's health care cost on deductible choice.

In model 1, having stable low costs during the 2010-2012 period is positively related to the propensity of having a 500 euro deductible in 2013. However, the marginal effect of *stablelowcost* in this model is close to zero. Furthermore, in model 2 where we include the cost of the current year, this effect disappears. This suggest that the learning effect that is measured by the *stablelowcost* variable is captured by the *cost2012* and *cost2013*

85 We determined the accuracy rate of both models based on a down-sampled dataset, including all people choosing a 500 euro deductible and an equally sized randomly selected subsample of people choosing a zero deductible. Assigning a 500 euro deductible to individuals when the predicted probability of choosing this deductible exceeds 0.5, both models predict 29 percentage point better than a 50/50 random prediction (79% vs. 50%).

86 For each possible combination of covariates, we calculated the difference in probability of choosing a 500 euro deductible in 2013 with and without having a 500 euro deductible in 2012. The weighted average of these differences in probability is then calculated by weighing each covariate combination with the number of individuals. The marginal effects of the other variables are calculated in the same way.

87 As discussed earlier, the effect of people's health care cost in 2013 on deductible choice may be over-estimated, because these costs are likely to be influenced to some extent by the choice of deductible as a result of moral hazard. Given the broad cost categories we distinguished, however, the moral hazard effect is most likely not large enough to have a substantial impact on the classification of individuals into the various cost classes. Therefore the potential endogeneity problem seems small.

88 Notice that the interaction effects for the highest cost category are not significant (2012) or only at 10% level (2013), which is probably due to the very small subgroup (only 454 individuals, see Table 3).

Table 4. Regression results

	<i>Dependent variable: 500 euro deductible in 2013 (Y/N)</i>	
	(1)	(2)
deductible2012	7.460 (0.028)	7.638 (0.033)
cost2012(2,4]	-0.004 (0.008)	0.051 (0.009)
cost2012(4,6]	-0.217 (0.009)	-0.060 (0.009)
cost2012(6,8]	-1.074 (0.010)	-0.655 (0.010)
cost2012(8,10]	-1.851 (0.016)	-1.150 (0.017)
cost2012(10,15]	-3.339 (0.097)	-2.312 (0.098)
income2	0.190 (0.008)	0.191 (0.008)
income3	0.525 (0.008)	0.519 (0.008)
age[30,40)	0.085 (0.010)	0.095 (0.010)
age[40,50)	0.090 (0.009)	0.126 (0.009)
age[50,60)	0.069 (0.009)	0.152 (0.009)
age[60,60+)	-0.467 (0.010)	-0.283 (0.010)
gender	0.092 (0.006)	0.071 (0.006)
discount	0.008 (0.0001)	0.008 (0.0001)
stablelowcost	0.079 (0.008)	-0.004 (0.008)
cost2013(2,4]		-0.048 (0.008)
cost2013(4,6]		-0.259 (0.008)
cost2013(6,8]		-0.882 (0.010)
cost2013(8,10]		-1.444 (0.018)
cost2013(10,15]		-1.579 (0.055)
deductible2012:cost2012(2,4]	-0.215 (0.041)	-0.119 (0.043)
deductible2012:cost2012(4,6]	-0.531 (0.037)	-0.363 (0.039)
deductible2012:cost2012(6,8]	-0.725 (0.034)	-0.637 (0.037)
deductible2012:cost2012(8,10]	-0.598 (0.043)	-0.600 (0.046)
deductible2012:cost2012(10,15]	-0.028 (0.133)	0.019 (0.137)
deductible2012:cost2013(2,4]		-0.160 (0.043)
deductible2012:cost2013(4,6]		-0.375 (0.039)
deductible2012:cost2013(6,8]		-0.488 (0.038)
deductible2012:cost2013(8,10]		-0.261 (0.049)
deductible2012:cost2013(10,15]		-0.114 (0.111)
constant	-5.770 (0.019)	-5.637 (0.020)
Observations	9,363,203	9,363,203
Log Likelihood	-667,808.6	-659,057.7
Akaike Inf. Crit.	1,335,659.0	1,318,177.0

Note: * = $p < 0.01$, standard errors in brackets

variable, which is plausible since for most of the individuals having low costs in 2010 and 2011 the cost in 2012 and 2013 are expected to be low as well. All other explanatory variables are significant and have the expected signs (e.g. Borghans et al. 2009). The propensity of choosing a 500 euro deductible is positively related to *income* and *gender* (being male). For *income*, however, the weighted average marginal effect of a change in income category is only 1 percentage point for individuals with zero deductible in 2012, and 3 percentage points for individuals with 500 euro deductible in 2012. For *gender*, the weighted average marginal effects are less than 1 percentage point.

The propensity of choosing a 500 euro deductible is also positively related to being enrolled in a health plan offering a higher premium *discount* in return for a 500 euro deductible. When comparing the probability of choosing a 500 euro deductible in 2013 between having the maximum (311) and minimum (150) *discount* in our sample, the weighted average marginal effect is 2 percentage points for individuals with zero deductible in 2012 and 6 percentage points for individuals with 500 euro deductible in 2012. Furthermore, the propensity of choosing a 500 euro deductible increases with *age* up to an age of 40-50, and then decreases, particularly among those older than 60. When comparing the probability of choosing a 500 euro deductible in 2013 between age 40-50 and age older than 60, the weighted average marginal effect is -1 percentage point for individuals with zero deductible in 2012 and -3 percentage points for individuals with 500 euro deductible in 2012.

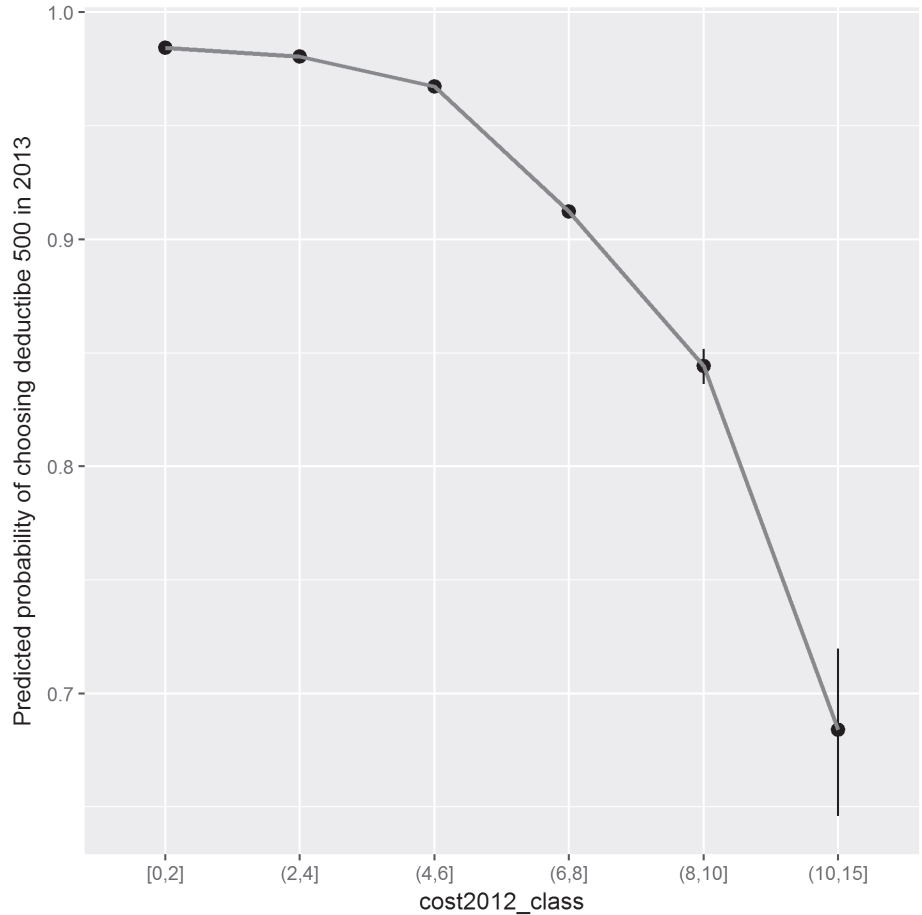
We illustrate the marginal effects of changes in 2012 cost classes on deductible choice in model 1 by Figures 3 and 4. Both figures refer to a 'median individual', i.e. the median of each covariate of a 500 euro deductible chooser (as specified below the figures).⁸⁹ Figure 3 shows for each 2012 cost category the probability that people with a 500 euro deductible in 2012 choose the same deductible in 2013.

As illustrated, individuals within the lowest cost category in 2012 have a 98% probability of continuing their 500 euro deductible choice in 2013, whereas this probability drops to 68% for individuals in the highest cost category. The 30 percentage point reduction in probability provides evidence of a substantial adverse selection effect. Despite extremely high costs (exceeding 20,000 euro) in 2012, 68% of the individuals do not lower their 500 euro deductible in 2013. This provides evidence of substantial consumer inertia, because it is unlikely that these costs were just incidental for most individuals. For the whole population, the weighted average reduction in the probability of choosing a 500 euro deductible for people moving from the lowest to the highest cost category is 37 percentage points.

Figure 4 shows for each 2012 cost class the probability that people with a zero deductible in 2012 choose a 500 euro deductible in 2013. Clearly, the probability that people

89 Using alternative values for *income*, *stablelowcost age*, *gender* and *discount* show similar patterns.

Figure 3. Probability of choosing a 500 euro deductible in 2013, for individuals with a 500 euro deductible in 2012 in six different 2012 cost classes

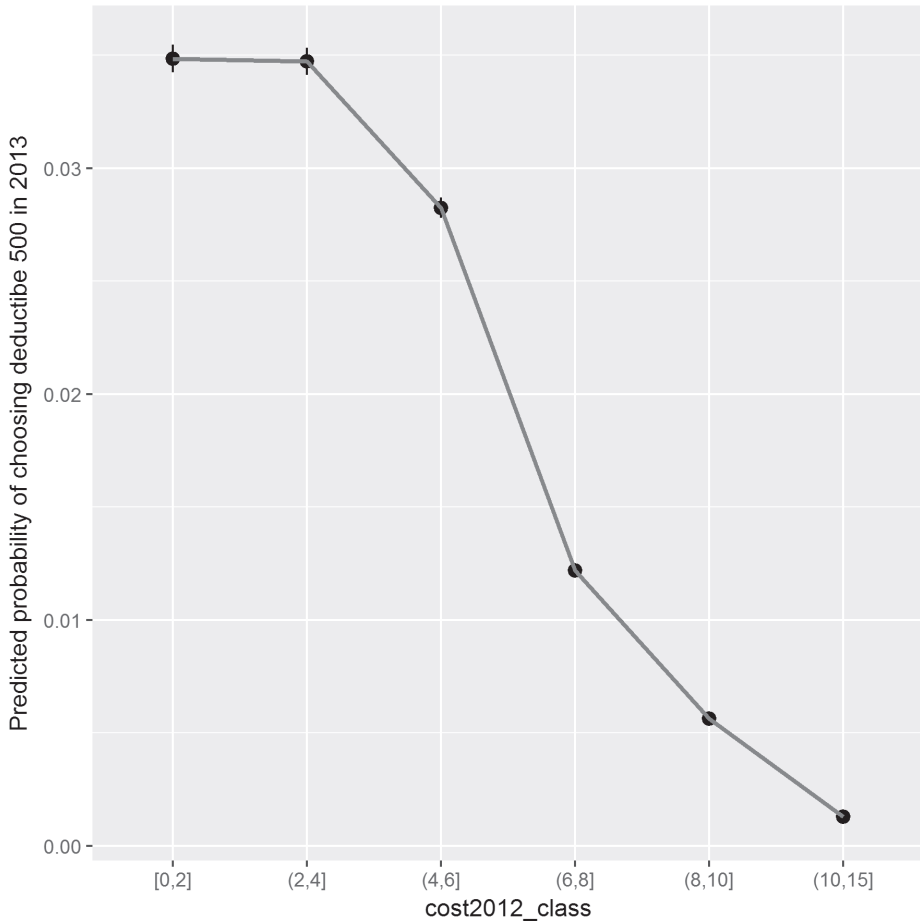


Based on model 1 for individuals with the following characteristics: income=category 3, stablelowcost=0, age=[40-50], gender=male, discount=222 euro.

switch from a zero to a 500 euro deductible is very small. Even for people having the lowest costs in 2012, the chance of taking up a 500 euro deductible in 2013 is only 3.5%. As expected, this percentage drops to close to zero percent for people experiencing very high health care expenses in 2012. The small share of low-cost individuals likely to take up a 500 euro deductible indicates a high degree of choice persistence with substantial implied monetary costs. For the whole population, the weighted average reduction in the probability of choosing a 500 euro deductible for people moving from the lowest to the highest cost category is 2 percentage points.

For each individual, we can retrospectively determine whether they made the most financially optimal deductible choice in 2013. Individuals who chose a zero deductible

Figure 4. Probability of choosing a 500 euro deductible in 2013, for individuals with a zero euro deductible in 2012 in six different 2012 cost classes



Based on model 1 for individuals with the following characteristics: income=category 3, stablelowcost=0, age=[40-50], gender=male, discount=222 euro.

in 2013 did not make an optimal choice if their costs are lower than the mandatory deductible (350 euro) plus the premium discount offered by their health plan for a 500 euro deductible. Similarly, individuals who chose 500 euro deductible in 2013 and have costs on top of the mandatory deductible (350 euro) that are higher than the discount, did not make an optimal choice. In 2013, 53% of the individuals who have chosen zero deductible would be financially better off with a 500 euro deductible. And for individuals who have chosen a 500 euro deductible, 14% would be better off with a zero euro deductible. We approximated the implied monetary costs by determining the gain that those individuals would attain if they would have chosen optimally in retrospect. The

average implied monetary costs is 229 euro for individuals who have chosen a suboptimal zero deductible, and 187 euro for individuals who have chosen a suboptimal 500 euro deductible.

As discussed above, previous cost experience does not necessarily mean that people can predict future health expenses. To get an indication of the extent to which people foresee future health expenses and act upon it, we used the estimation results of model 2 to examine for each 2012 cost category the relationship between people's choice of deductible and their (at that time future) health care cost in 2013. Figure 5 illustrates this relationship.

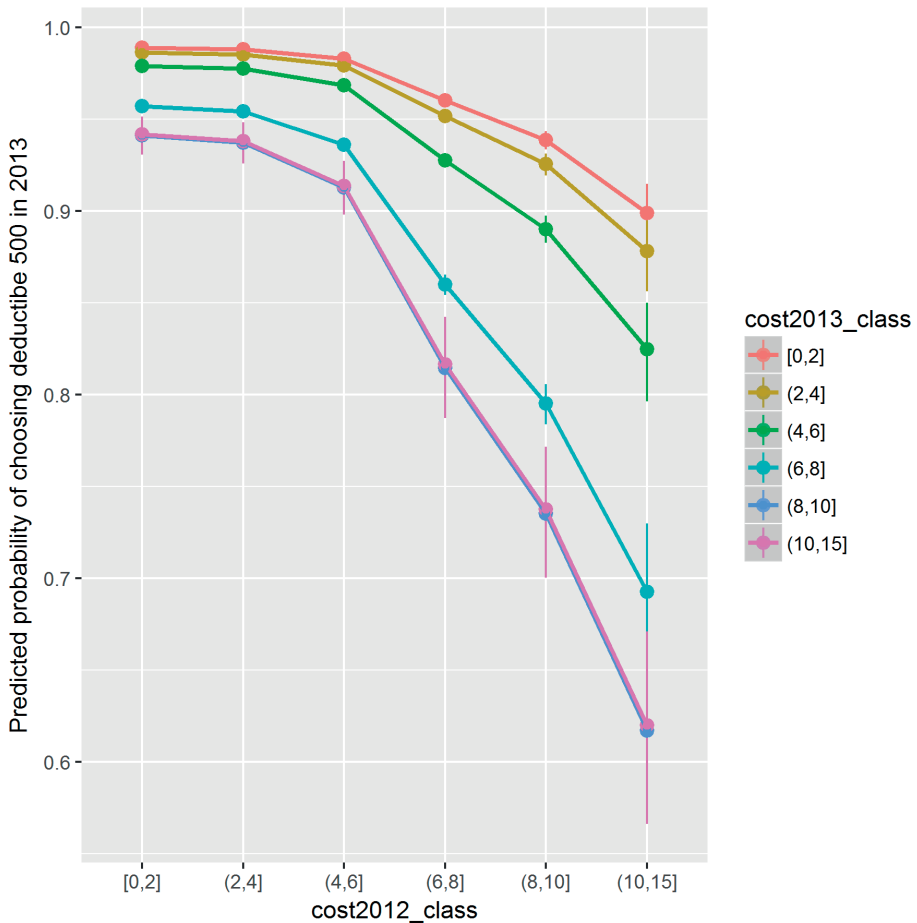
For each 2012 cost category, Figure 5 shows that the probability of people keeping a 500 euro deductible decreases if their health care expenses in 2013 increase. For instance, for people within the lowest cost category in 2012 the probability of keeping a 500 euro deductible in 2013 reduces from 99 to 94% if they shift to the highest cost category in 2013. Hence, only a small minority of about 5% of people having low cost in 2012 and high cost in 2013 seem to have anticipated a substantial cost increase in 2013 by reducing their deductible to zero. Much more pronounced is the anticipatory behavior of those who experienced high cost in 2012 followed by – on hindsight – low cost in 2013. As shown in Figure 5, for people having the highest cost in 2012 the propensity of keeping a 500 euro deductible is 62% if they have high costs in 2013, whereas this propensity increases to 90% if they have low costs in 2013. Hence, a substantial share of high cost individuals effectively anticipates future expected health care cost by (not) adjusting their choice of deductible. However, the fact that 62% of people within the highest cost groups in both years nevertheless keeps their 500 euro deductible further supports the evidence of high consumer inertia.

For individuals who had a zero deductible in 2012, Figure 6 shows a similar pattern as Figure 5. Although the probability that people switch from a zero to a 500 euro deductible is very small, we again find that at least some people effectively anticipate a change in future health care costs. For instance, the probability that people within the lowest cost class in 2012 will uptake a 500 euro deductible in 2013 reduces from 4% to less than 1% if they move from the lowest to the highest cost class in 2013. Still, the fact that 96% of those in the lowest cost class in both years sticks with a zero deductible indicates also here that consumer inertia and choice persistence are high.

7. CONCLUSION AND DISCUSSION

In the context of the Dutch health insurance market where people can annually choose a deductible level, varying from 0 to 500 euro, we find evidence of both adverse selection and consumer inertia. Using data for about 9.5 million individuals who do not

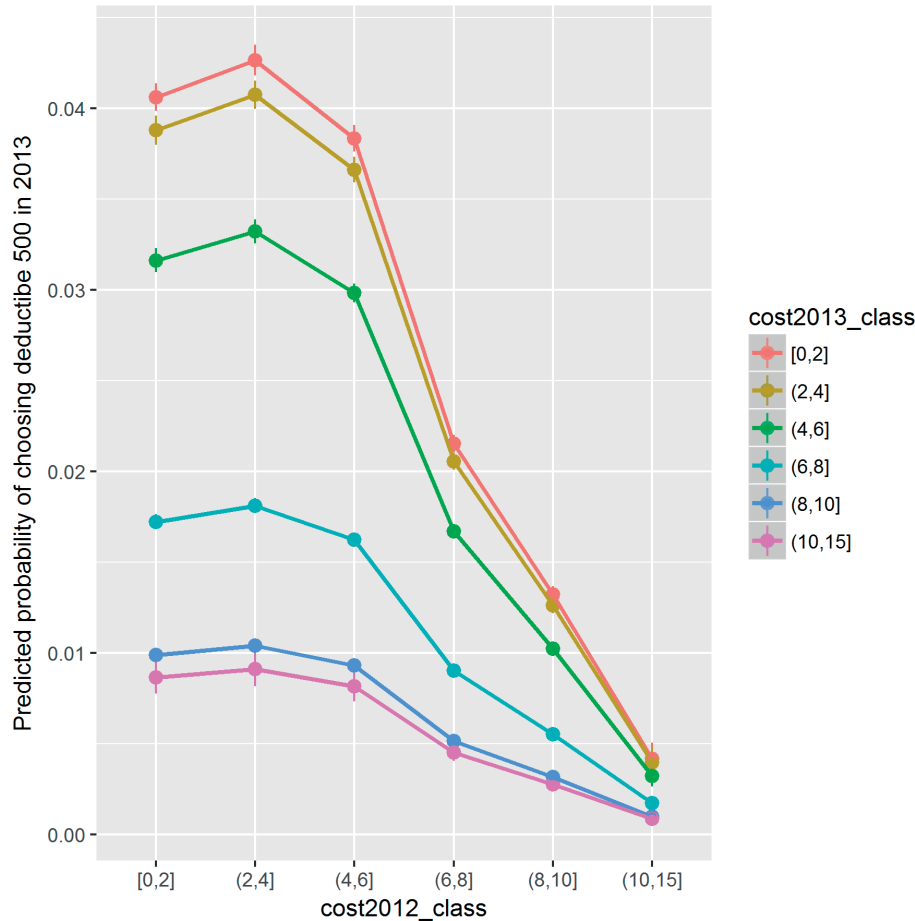
Figure 5. Relationship between the probability of sticking with a 500 euro deductible in 2013 and people's health care cost class in 2013, for individuals with a 500 euro deductible in 2012 in six different 2012 cost classes



Based on model 1 for individuals with the following characteristics: income=category 3, stablelowcost=0, age=[40-50], gender=male, discount=222 euro.

switch health plans during a four years study period (2010-2013), we first constructed 16 possible choice paths for people choosing a zero or 500 euro deductible. We find that all possible choice paths have cost patterns that are consistent with the presence of adverse selection. The patterns also suggest that adverse selection may take time to arise since it can take several years before people with low health care cost substitute a 500 euro for a zero deductible. In addition, the choice paths show that on average healthy people are able to anticipate effectively next year's health care costs. Next, using the same dataset over the period 2010-2013 (including about 37.5 million observations) we estimated a logit model to examine to what extent the individual choice of deductible

Figure 6. Relationship between the probability of changing from a zero to a 500 euro deductible in 2013 and people's health care cost class in 2013, for individuals with a 500 euro deductible in 2012 in six different 2012 cost classes



Based on model 1 for individuals with the following characteristics: income=category 3, stablelowcost=0, age=[40-50], gender=male, discount=222 euro.

in 2013 can be explained by previous and future health care cost. We find clear evidence of adverse selection, as people with higher previous and future health care cost are substantially less likely to take up or keep a 500 euro deductible. Furthermore, our results also indicate that people do anticipate future health care cost as these cost can explain part of the variation in deductible choice given their cost in the previous year.

However, we also find clear evidence of high consumer inertia as the propensity of taking up a 500 euro deductible among low-cost individuals is only 3.5%, while the propensity of keeping a 500 euro deductible among high-cost individuals is 68%. For both low-cost individuals sticking with a zero deductible and high-cost individuals

sticking with a 500 euro deductible, the implied monetary costs of choice persistence are approximately 200 euros per year. The substantial degree of consumer inertia is remarkable, given the very low transaction costs for enrollees involved in adjusting their deductible level.

By counteracting adverse selection, a certain degree of consumer inertia may be welfare increasing (Handel 2013, Handel and Kolstad 2015). Within the context of the Dutch health insurance market with regulated (or managed) competition, however, the presence of substantial consumer inertia may well be welfare decreasing. This is because the potential negative welfare effects of adverse selection are possibly substantially mitigated by a sophisticated system of risk equalization. Despite the presence of risk equalization, however, those opting for the highest deductible level appear to be profitable to insurers at the prevailing premium discount levels (Croes et al. 2017). Hence, an active policy to reduce consumer inertia may enhance adverse selection and increase insurers' profits. Nevertheless, the associated negative welfare effects are likely to be much smaller than the positive welfare effects of enhancing active consumer choice because of the large share of the population currently incurring a substantial implied monetary loss (about 200 euro per person). This suggests that within the Dutch context, activities aimed at reducing consumer inertia with regard to deductible choice are likely to be welfare improving.

REFERENCES

- Abaluck, J., Gruber, J. (2011). Choice inconsistencies among the elderly: evidence from plan choice in the medicare part d program. *American Economic Review*, 101 (4), 1180–1210.
- Bhargava, S., Loewenstein, G., Sydnor, J. (2015). Do individuals make sensible health insurance decisions? Evidence from a menu with dominated options. National Bureau of Economic Research (NBER), Cambridge (Mass.)
- Boonen, L. H., Laske-Aldershof, T., Schut, F. T. (2016). Switching health insurers: the role of price, quality and consumer information search. *European Journal of Health Economics*, 17 (3), 339–353.
- Borghans, L., Golsteyn B.H.H., Heckman, J.J., Meijers, H. (2009). Gender differences in risk aversion and ambiguity. *Journal of the European Economic Association*, 7(2-3), 649–658.
- Croes, R., Katona, K., Mikkers, M., Shestalova, V. (2017). Evidence of selection in a mandatory health insurance market with risk adjustment. Working paper, Dutch Healthcare Authority (NZa), Utrecht.
- Douven, R., Remmerswaal, M. (2016). Keuzegedrag verzekeren en risicosolidariteit bij vrijwillig eigen risico, CPB Netherlands Bureau for Economic Policy Analysis, The Hague.
- Douven, R., Katona, K., Schut, F.T., Shestalova, V. (2017), Switching gains and health plan price elasticities: 20 years of managed competition reforms in the Netherlands, *European Journal of Health Economics*, 18 (8), 1047–1064.
- Einav, L., Finkelstein A. (2011), Selection in insurance markets: Theory and empirics in pictures, *Journal of Economic Perspectives*, 25 (1), 115–138.
- Handel, B. R. (2013). Adverse selection and inertia in health insurance markets: When nudging hurts. *American Economic Review*, 103 (7), 2643–2682.
- Handel, B. R., Kolstad, J. T. (2015). Health insurance for humans: Information frictions, plan choice, and consumer welfare. *American Economic Review*, 105 (8), 2449–2500.
- Handel, B. R., Kolstad, J. T., Spinnewijn, J. (2015). Information frictions and adverse selection: policy interventions in health insurance markets. Working Paper 21759, National Bureau of Economic Research (NBER), Cambridge (Mass.).
- Heiss, F., McFadden, D., Winter, J., Wuppermann, A., Zhou, B. (2016). Inattention and switching costs as sources of inertia in Medicare Part D. Working Paper 22765, National Bureau of Economic Research (NBER), Cambridge (Mass.).
- Ho, K., Hogan, J., Morton, F. S. (2015). The impact of consumer inattention on insurer pricing in the medicare part d program. Working Paper 21028, National Bureau of Economic Research (NBER), Cambridge (Mass.)
- Marzilli Ericson, K. M. (2014). Consumer inertia and firm pricing in the medicare part d prescription drug insurance exchange. *American Economic Journal: Economic Policy*, 6 (1), 38–64.
- McGuire, T.G., Van Kleef, R. (eds) (2018), Risk Adjustment, risk sharing and premium regulation in health insurance markets. Elsevier Publishing. forthcoming.
- NZa (2013). Marktscan en beleidsbrief Zorgverzekeringsmarkt 2013, Nederlandse Zorgautoriteit, Utrecht.
- Pauly, M. V. (1984). Is cream-skimming a problem for the competitive medical market? *Journal of Health Economics*, 3 (1), 87–95.
- Polyakova, M., (2016). Regulation of insurance with adverse selection and switching costs: Evidence from Medicare Part D. *American Economic Journal: Applied Economics*, 8 (3), 165–95.
- Rothschild, M., Stiglitz, J. (1976). Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *Quarterly Journal of Economics*, 90 (4), 629–649.
- Samuelson, W., Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty*, 1(1), 7–59.

- Van Kleef, R., McGuire, T., van Vliet, R., van de Ven, W. (2017). Improving risk equalization with constrained regression. *European Journal of Health Economics*, 18 (9), 1137-1156.
- Van de Ven. W.P.M.M., Schut, F.T. (2008). Universal mandatory health insurance in the Netherlands: a model for the United States? *Health Affairs*, 27(3), 771–781.
- Van Winssen, K., van Kleef, R., van de Ven, W. (2015). How profitable is a voluntary deductible in health insurance for the consumer? *Health Policy*, 119 (5), 688–695.
- Vektis (2017), *Zorgthermometer Verzekerden in beeld 2017*. Vektis. Zeist.

Chapter 7

Conclusion

CONCLUSION

In this thesis, we analyzed aspects of some important market failures that may hamper an appropriate functioning of the Dutch health care system: market power in the provider market, and selection and consumer inertia in the insurance market.

Chapter 2 tackled the issue of market power in the Dutch hospital market. Since mergers are the most important source of increased market concentration, and many mergers took place in the Dutch hospital market, it is essential to improve our understanding of how hospital mergers may affect market performance. My first research question was: ***For a merger between two neighboring hospitals in the Netherlands, do the prices increase after a merger and do we observe differential price changes between different hospital locations, different products and different insurers?*** We examined the price effect of a hospital merger for three products: hip replacement, knee replacement and cataract surgery. We found evidence of heterogeneous price effects across health insurers, hospital products and hospital locations. After disaggregating the price effect per product and location, significant price increases related to the merger were found. These price increases, however, occurred at one of the locations of the merged hospitals and for one of the three products (i.e. hip replacement). Furthermore, the price effects varied considerably among insurers. These findings therefore suggest that when analyzing the effects of a hospital merger on market performance, it is important to take a more disaggregate approach to get a more complete picture of the relevant merger effects.

In most developed countries, a merger is submitted for approval to a competition authority, which reviews the merger and decides whether it is allowed to be consummated. The findings in Chapter 2 show that for competition authorities it may be worthwhile to investigate ex ante whether a proposed hospital merger may result in price increases in specific sub-markets or locations. This would allow competition authorities to reveal and address in greater detail the potential detrimental effects of a hospital merger.

Unfortunately, predicting the effect of a hospital merger is not an easy task. Economists have developed new methods to review mergers in hospital markets, such as the Option Demand Method. The OD-method has clear advantages over more traditional market definition approaches, because it provides antitrust agencies with direct evidence about the expected effects of the merger. Additionally, it does not require questionable assumptions associated with delineation of the relevant (geographic) market. There is, however, a lack of empirical evidence that approaches like the OD-method are able to give reliable predictions. My second research question therefore was: ***What is the predictive power of the Option Demand Method for mergers in the Dutch hospital market?*** In Chapter 3, we examined this question by comparing the ex-ante prediction of the simulation model with the ex-post observed price effects for a consummated hospital merger in the Netherlands.

Our results indicated a relationship between our measure of market power and prices for hip replacements. We were not able to establish a relationship between market power and prices for knee replacements and cataract surgeries. Therefore, only a reduced-form merger simulation for hip replacements was estimated. The comparison between the reduced-form merger simulation and ex post estimates are not conclusive. Our findings suggest that the OD-method could be a valuable addition to the antitrust agencies toolkit, but that more research remains necessary.

Although prices are of course an important market outcome, for patients the quality of hospital care may be even more important. However, in the Dutch hospital market – where most prices are not regulated by the government – the potential impact of competition on quality is ambiguous. That is, depending on the specific market conditions competition may result in lower or higher quality of care. Hence, my third research question was: ***What is the relationship between competition and quality indicators in the Dutch non-price-regulated hospital market?*** In Chapter 4, we investigated this question by combining patient-level claims data with information on quality indicators. We examined three diagnosis groups (cataract, adenoid and tonsils, and bladder tumor) delivered by Dutch hospitals in the period 2008-2011. For this study period, we evaluated the relationship between hospitals' quality scores and indicators of hospital market power. For cataract and bladder tumor, the relationship between market share and quality scores was found to be negative. For adenoid and tonsils, however, this relationship was not statistically significant. One possible explanation for this difference is that the patient group for adenoid and tonsils is less complex. Mainly children younger than 11 years are treated for adenoid and tonsils, and these patients have relatively few additional diagnoses. It could be the case that patients are less sensitive to quality when their treatment has a lower risk of complications. Thus, for two of the three diagnosis groups studied, hospitals in (more) competitive markets had better quality scores than those in concentrated markets.

In several developed (OECD) countries, markets also play an important role in the financing of health care. This is also the case in the Netherlands. Such health care markets consist of a complex interaction between consumers, insurers and providers. Generally, on the health insurance market consumers buy health plans from competing health insurers. These health plans give consumers access to a network of providers when they seek treatment. An insurer and a provider negotiate over the inclusion of the provider into the insurer's network.

The functioning of health insurance markets can be hampered by selection problems and information frictions. The interaction of both phenomena, however, may reduce welfare problems because the latter (partly) counteracts the former. Empirical identification of the exact extent of the selection problem is in general challenging because of the interaction of selection and moral hazard. The mere fact that people who choose a

voluntary deductible have lower health expenses does not, by itself, prove the presence of adverse selection. This may also be due to more cost conscious behavior induced by the deductible (i.e. less moral hazard). In Chapter 5, we tried to control for the moral hazard effect to answer my fourth research question: ***How large is the selection effect for individuals who chose voluntary deductibles in a health insurance market with risk-adjustment?*** By examining people's deductible choice, while taking account of their prior health care expenses, we were able to identify the selection effect separately from moral hazard effect. Our analysis showed that offering contracts with voluntary deductibles results in self-selection by healthier individuals, who are overcompensated by the risk-adjustment system. The expected gains on the risk-adjustment per individual with a voluntary deductible of 500 euro were estimated to be around 450 euro on average. On top of this, the corresponding individuals pay a larger share out-of-pocket, resulting in lower costs for the insurer to be reimbursed. In return, the insurer offers the individual a premium discount for taking a voluntary deductible, but this discount is typically (much) lower than the cost savings due to self-selection.

The interaction between adverse selection and consumer inertia is the subject of my fifth research question: ***Is adverse selection effectively mitigated by consumer inertia in the Dutch health insurance market?*** In Chapter 6, using panel data on all insured individuals in the Netherlands over period 2010-2013, we examined this question by constructing all possible choice paths for people choosing a zero or 500 euro voluntary deductible. For each individual, his choice path is based on the possible deductible choices that the individual can make in each year. For each individual we can distinguish 16 possible deductible choice paths. We found that all possible choice paths have cost patterns that are consistent with the presence of adverse selection. The patterns also suggested that on average healthy people are able to anticipate effectively next year's health care costs. Estimating a logit model, we found clear evidence of adverse selection. That is, people with higher previous and future health care cost are substantially less likely to take up or keep a 500 euro deductible. However, we also found clear evidence of high consumer inertia as the propensity of taking up a 500 euro voluntary deductible among low-cost individuals is only 3.5% while such a deductible would probably be (very) profitable for them.

To a certain extent consumer inertia may be welfare increasing, because it counteracts adverse selection. Within the context of the Dutch health insurance market with regulated (or managed) competition, however, the presence of substantial consumer inertia may well be welfare decreasing. This is because the potential negative welfare effects of adverse selection are possibly substantially mitigated by a sophisticated system of risk equalization. Despite the presence of risk equalization, however, we find that individuals opting for the highest deductible level appear to be overcompensated by the risk-adjustment system. Apart from this, reducing consumer inertia on the health

insurance market may lead to more adverse selection. This may have a negative effect on consumer welfare, despite the fact that moral hazard decreases when more people choose voluntary deductibles. Nevertheless, the associated negative welfare effects are likely to be much smaller than the positive welfare effects of enhancing active consumer choice because of the large share of the population currently incurring a substantial implied monetary loss (about 200 euro per person). This suggests that, within the current Dutch market for basic health insurance, activities aimed at reducing consumer inertia with regard to deductible choice are likely to be welfare improving. These potential welfare gains might even be higher if the reduction of consumer inertia is combined with an improvement in risk-adjustment system that reduces the expected gains on the risk-adjusted payments for people choosing a voluntary deductible.

Samenvatting

In dit proefschrift zijn twee belangrijke vormen van marktfalen geanalyseerd die een goede werking van het Nederlandse zorgstelsel kunnen belemmeren, te weten (i) marktmacht in de zorgverleningsmarkt en (ii) selectie en inertie in de zorgverzekeringsmarkt.

Hoofdstuk 2 behandelt het probleem van marktmacht in de Nederlandse ziekenhuismarkt. De belangrijkste oorzaak van de stijging van de marktconcentratie in de Nederlandse ziekenhuismarkt is het groot aantal fusies tussen ziekenhuizen. Daarom is het van belang om goed te begrijpen hoe fusies tussen ziekenhuizen de werking van de ziekenhuismarkt beïnvloeden. De eerste onderzoeksvraag was: ***Stijgen de prijzen van ziekenhuiszorg na een fusie tussen twee naburige ziekenhuizen in Nederland? En zien we variatie in de prijsveranderingen tussen verschillende ziekenhuislocaties, verschillende producten en verschillende zorgverzekeraars?*** We hebben het prijseffect van een ziekenhuisfusie geanalyseerd voor drie producten: heupvervanging, knievervanging en staaroperatie. Onze resultaten toonden aan dat er heterogene prijseffecten zijn tussen ziekenhuislocaties, producten en zorgverzekeraars. Wanneer we het prijseffect splitsen naar locatie en product, vonden we significantie prijsstijgingen die gerelateerd zijn aan de fusie. Deze prijsstijgingen vonden plaats bij één van de twee locaties en alleen bij één van de drie producten (heupvervanging). Daarnaast varieerde het prijseffect sterk tussen de verschillende zorgverzekeraars. Deze bevindingen geven aan dat, wanneer we het effect van een fusie op de ziekenhuismarkt analyseren, het belangrijk is om op de effecten op verschillende niveaus te onderzoeken om een completer beeld te krijgen van de prijseffecten van de fusie.

In de meeste ontwikkelde landen melden ziekenhuizen een nieuwe fusie eerst bij een mededingingsautoriteit. De mededingingsautoriteit onderzoekt de fusie en beslist of de fusie mag doorgaan. De resultaten in Hoofdstuk 2 geven aan dat het de moeite waard kan zijn voor een mededingingsautoriteit om vooraf (ex ante) te onderzoeken of de fusie mogelijke prijsstijgingen kan veroorzaken in specifieke sub-markten of locaties. Dit geeft de mededingingsautoriteit de mogelijkheid om uitvoerig de mogelijke negatieve effecten van de fusie in kaart te brengen en te adresseren.

Helaas is het voorspellen van de effecten van een ziekenhuisfusie niet eenvoudig. Economen hebben nieuwe methoden ontwikkeld om fusies te beoordelen, zoals de Option Demand Methode (OD-Methode). De OD-methode heeft duidelijke voordelen ten opzichte van de meer traditionele marktafbakening methoden, omdat de OD-methode een directe schatting doet van de effecten van een fusie. Ook vereist het geen aanvechtbare aannames, zoals die wel gebruikt worden bij methoden voor marktafbakening. Er is echter een gebrek aan empirisch bewijs dat benaderingen zoals de OD-methode in staat zijn betrouwbare voorspellingen te doen. Vandaar de tweede onderzoeksvraag: ***Wat is de voorspellende kracht van de Option Demand Methode voor fusies in de Nederlandse ziekenhuismarkt?*** In Hoofdstuk 3 hebben we deze vraag onderzocht door

de ex-ante voorspelling van een simulatiemodel te vergelijken met de ex-post waargenomen prijseffecten voor een voltooide ziekenhuisfusie in Nederland.

Onze resultaten toonden aan dat er een verband bestaat tussen onze marktmacht indicator en de prijzen voor heupvervangende. Onze bevindingen suggereren dat de OD-methode een waardevolle aanvulling zou kunnen zijn voor de mededingingsautoriteiten, maar dat er meer onderzoek nodig blijft.

Hoewel prijzen natuurlijk een belangrijk marktresultaat zijn, kan de kwaliteit van de ziekenhuiszorg voor patiënten nog belangrijker zijn. Op de Nederlandse ziekenhuismarkt - waar de meeste prijzen niet door de overheid worden gereguleerd maar vrij onderhandelbaar zijn - is de potentiële impact van concurrentie op de kwaliteit op voorhand onduidelijk. Dat wil zeggen, afhankelijk van de specifieke marktomstandigheden kan concurrentie resulteren in een lagere of hogere kwaliteit van zorg. Vandaar de derde onderzoeksvraag: ***Wat is de relatie tussen concurrentie en kwaliteitsindicatoren in de Nederlandse ziekenhuismarkt waar prijzen vrij onderhandelbaar zijn?***

In hoofdstuk 4 hebben we deze vraag onderzocht door DBC-data op patiëntniveau te combineren met informatie over kwaliteitsindicatoren. We onderzochten drie diagnosegroepen (staar, keel- en neusamandelen, en blaastumor) uitgevoerd door Nederlandse ziekenhuizen in de periode 2008-2011. Voor deze onderzoeksperiode evalueerden we de relatie tussen de kwaliteitsscores en de marktmacht indicatoren van de ziekenhuizen. Voor staar (cataract) en blaastumor vonden we een negatieve relatie tussen marktaandeel en kwaliteitsscores. Voor keel- en neusamandelen was deze relatie echter niet statistisch significant. Een mogelijke verklaring voor dit verschil is dat de patiëntengroep voor keel- en neusamandelen minder complex is. Voornamelijk kinderen jonger dan 11 jaar worden behandeld voor keel- en neusamandelen. Deze patiënten hebben relatief weinig aanvullende diagnoses. Het kan zijn dat patiënten minder gevoelig zijn voor kwaliteit bij behandelingen met een lager risico op complicaties. Voor twee van de drie onderzochte diagnosegroepen hadden ziekenhuizen in (meer) concurrerende markten betere kwaliteitsscores dan ziekenhuizen in geconcentreerde markten.

In verschillende ontwikkelde (OESO) landen spelen markten ook een belangrijke rol bij de financiering van de gezondheidszorg. Dit is ook het geval in Nederland. Dergelijke zorgmarkten bestaan uit een complexe interactie tussen consumenten, verzekeraars en aanbieders. Over het algemeen kopen consumenten polissen van concurrerende zorgverzekeraars. Deze polissen geven consumenten toegang tot een netwerk van gecontracteerde aanbieders wanneer zij medische behandeling nodig hebben. Een verzekeraar en een aanbieder onderhandelen over de contractvoorwaarden (prijs, volume en kwaliteit van zorg).

Het functioneren van de zorgverzekeringsmarkt kan worden belemmerd door selectie- en informatieproblemen. Informatieproblemen kunnen selectieproblemen echter verminderen wanneer consumenten minder of trager (dat wil zeggen inert) reageren

op premieverschillen. Dit betekent dat het beter informeren van consumenten op de zorgverzekeringsmarkt vanuit maatschappelijk perspectief per saldo ongunstig kan uitpakken. Empirische identificatie van de exacte omvang van het selectieprobleem is over het algemeen een uitdaging vanwege de interactie tussen selectie en moreel risico. Het enkele feit dat mensen die kiezen voor een vrijwillig eigen risico lagere gezondheidskosten hebben, bewijst op zichzelf niet dat er sprake is van adverse selectie. Dit kan ook te wijten zijn aan meer kostenbewust gedrag veroorzaakt door het vrijwillig eigen risico (dat wil zeggen, minder moreel risico).

In Hoofdstuk 5 probeerden we voor het moreel risico-effect te controleren om de vierde onderzoeksvraag te kunnen beantwoorden: ***Hoe groot is het selectie-effect voor individuen die vrijwillig eigen risico kiezen in een zorgverzekeringsmarkt met risicoverevening?*** We hebben op het niveau van individuele verzekerden geanalyseerd welke keuzes zij hebben gemaakt voor het vrijwillig eigen risico. Hierbij hebben we rekening gehouden met hun eerdere zorgkosten. Hierdoor konden we het selectie-effect onderscheiden van het moreel risico-effect. Uit onze analyse blijkt dat het aanbieden van contracten met een vrijwillig eigen risico resulteert in zelfselectie door gezondere personen. Bovendien blijken deze personen overgecompenseerd te worden door het risicovereveningsysteem. De verwachte winst op het risicovereveningsysteem per individu met een vrijwillig eigen risico van 500 euro wordt geschat op ongeveer 450 euro (gemiddeld). Dit voordeel komt voor de verzekeraars bovenop het voordeel van lagere kosten doordat verzekerden met een vrijwillig eigen risico een groter deel van de zorgkosten uit eigen zak betalen. In ruil voor deze voordelen biedt de verzekeraar een premiekorting aan voor het nemen van een vrijwillig eigen risico, maar deze korting is meestal (veel) lager dan de kostenbesparingen als gevolg van zelfselectie en hogere eigen betalingen door de consument.

De interactie tussen adverse selectie en inert consumentengedrag leidt tot de vijfde onderzoeksvraag: ***Wordt adverse selectie effectief verzacht door inertie van consumenten in de Nederlandse zorgverzekeringsmarkt?*** In hoofdstuk 6 hebben we deze vraag onderzocht met behulp van paneldata over alle verzekerden in Nederland gedurende de periode 2010-2013. Hiervoor hebben we alle mogelijke keuzepaden geconstrueerd voor individuen die een vrijwillig eigen risico van nul of 500 euro hebben gekozen. Voor elk individu is zijn keuzepad gebaseerd op de mogelijke vrijwillig eigen risico keuzes die het individu elk jaar kan maken. Voor elk individu kunnen we op die manier 16 mogelijke vrijwillig eigen risico keuzepaden onderscheiden. We hebben vastgesteld dat alle mogelijke keuzepaden kostenpatronen hebben die consistent zijn met de aanwezigheid van adverse selectie. De patronen suggereren ook dat gezonde mensen, gemiddeld gezien, in staat zijn om effectief te anticiperen op hun toekomstige zorgkosten. We hebben op basis van de data ook een logit-model geschat van de individuele keuze voor een vrijwillig eigen risico. We vonden duidelijk bewijs van adverse se-

lectie. Dat wil zeggen dat mensen met hoge voorgaande en toekomstige kosten voor de gezondheidszorg aanzienlijk minder geneigd zijn om een vrijwillig eigen risico bedrag van 500 euro te nemen of te houden. We vonden echter ook duidelijk bewijs van een substantiële consumenteninertie: de kans om een vrijwillig eigen risico van 500 euro te nemen bij personen met lage kosten is slechts 3,5% terwijl een dergelijk vrijwillig eigen risico voor de meeste van hen waarschijnlijk (zeer) winstgevend zou zijn.

Tot op zekere hoogte kan de consumenteninertie welvaartsverhogend zijn, omdat het adverse selectie tegengaat. In het kader van de Nederlandse zorgverzekeringsmarkt met gereguleerde concurrentie kan de aanwezigheid van substantiële consumenteninertie echter negatief uitpakken voor de consumentenwelvaart. Dit komt omdat de potentiële negatieve welvaartseffecten van adverse selectie mogelijk aanzienlijk worden beperkt door een uitgebreid systeem van risicoverevening. Ondanks de aanwezigheid van risicoverevening, vinden we dat individuen die kiezen voor het hoogste vrijwillig eigen risico overgecompenseerd worden door het risicovereveningssysteem. Los hiervan kan het verminderen van de consumenteninertie op de zorgverzekeringsmarkt mogelijk leiden tot meer adverse selectie. Dit kan, afgezien van het feit dat wanneer meer mensen een vrijwillig eigen risico kiezen de moral hazard afneemt, een negatieve effect hebben op de consumentenwelvaart. Niettemin zullen de bijbehorende negatieve welvaartseffecten waarschijnlijk veel kleiner zijn dan de positieve welvaartseffecten van het verbeteren van de actieve consumentenkeuze vanwege het grote aandeel van de bevolking dat momenteel een aanzienlijk hogere premie betaalt dan noodzakelijk (ongeveer 200 euro per persoon). Dit suggereert dat, binnen de huidige Nederlandse zorgverzekeringsmarkt, activiteiten die gericht zijn op het verminderen van de consumenteninertie met betrekking tot vrijwillig eigen risico waarschijnlijk een positief effect hebben op de consumentenwelvaart. Deze potentiële welvaartswinsten kunnen zelfs hoger zijn als de vermindering consumenteninertie wordt gecombineerd met een verbetering van het risicovereveningssysteem om de overcompensatie voor mensen die een vrijwillig eigen risico kiezen te verminderen.

