Managed competition in practice

Katona, Katalin

Document version:
Publisher's PDF, also known as Version of record

Publication date:
2019

[Link to publication](#)

Citation for published version (APA):
Katona, K. (2019). Managed competition in practice: Lessons for healthcare policy. Tilburg: CentER, Center for Economic Research.

Managed Competition in Practice

Lessons for Healthcare Policy

Katalin Katona

Managed Competition in Practice
Lessons for Healthcare Policy

Proefschrift

ter verkrijging van de graad van doctor aan de Tilburg University op gezag van de rector magnificus, prof. dr. E.H.L. Aarts, in het openbaar te verdedigen ten onderstaan van een door het college voor promoties aangewezen commissie in de Portrettenzaal van de Universiteit op

dinsdag 14 mei 2019 om 13.30 uur

door

Katalin Katona
geboren op 19 oktober 1982
te Boedapest, Hongarije

Promotiecommissie

Promotor:	prof. dr. J. Boone
Copromotor:	dr. C. Argenton
Overige leden:	prof. dr. M. Gaynor prof. dr. M. Varkevisser prof. dr. J.J. Polder dr. M.J. Bijlsma

Acknowledgments

This dissertation is the result of many years of study and work. During these years I worked together with many colleagues and I've got support from even more friends. This short acknowledgment cannot be long enough to name everyone who contributed directly or indirectly to complete my PhD. Still, I would like to thank the support of the following people explicitly.

First of all, I am grateful to my supervisors: Marcel Canoy and Wolf Sauter in the initial years, and Jan Boone and Cédric Argenton in the subsequent years. Marcel, Wolf, Jan, Cédric, thank you for your tips and tricks on doing research. Your guiding advice and critical questions helped me to improve my papers every time. Also thank you for the freedom to find my own research interest. Your style of supervision allowed me to optimally combine my work and PhD studies.

Further, I would also like to thank my co-authors for contributing so much to the chapters in this dissertation. I had the chance to learn the ins and outs of doing research and writing a paper by doing it together, which is an efficient and pleasant way. Victoria, Rein, Ramsis, Misja, Rudy, Erik and Marcel, thanks to you, my PhD didn't feel like a solitary work. Quite the opposite! I enjoyed the co-creation and I also learned a lot from you. Thank you.

Also, I enjoyed an enormous support at my work. The Dutch Healthcare Authority, and especially my unit EMB, gave the stability and supportive environment for this enormous project of writing a PhD dissertation. I would like to thank my colleagues for their genuine interest in my research. Even the small brainstorms at work helped me largely by getting a fresh look on the problems. I am especially grateful to my manager, Misja Mikkers, who not only gave me the room to combine research and work but also encouraged me to finish my PhD. Misja, thank you for your encouragement which gave me several times a fresh impetus in situations of standstill.

Finally, I would like to thank my family for their constant support. I am grateful to my parents, who always supported me and encouraged to reach even higher. Köszönöm! I am also thankful to my children, Élon, Lilith, Adin, who taught me to keep balance and put problems in perspective. And I especially thank my partner who was always there for me. Peter, thank you for your endless encouragement to finish the PhD. Your enthusiasm for a "quick" brainstorm and getting things done helped me a lot. You also believed in my PhD more than anyone else, which became a self-fulfilling prophecy in the end. Thank you.

Katalin Katona
March 2019, Utrecht

Contents

Chapter 1

Introduction

1.1 Introduction

The idea of managed competition was developed by Alain Enthoven (Enthoven, 1978, 1988, 1993). His proposed system aimed to achieve public policy goals, such as efficiency and equity, by competition forces in a regulated environment. To obtain maximum value for consumers, he relies on (price) competition that is based on the annual premium for a health plan rather than the price of individual healthcare services. Health insurers have an important role in the system, because they are supposed to organize, manage and purchase healthcare. Managed competition is based on integrated healthcare organizations that offer both financing and delivery of healthcare services.

According to the ideas of Enthoven, a sponsor (which could be a governmental entity, an employer or purchasing cooperative) takes a central role in the system by regulating the markets. The sponsor creates price-elastic demand for health plans, which is a pre-condition for price competition between insurers, manages the enrollment process, selects plans from which consumers can choose, and ensures that there is an effective price competition between insurers and no risk selection. The sponsor has to monitor the working of the system, and if necessary intervene. This means an active role; regulation in this context is not a rigid set of rules but always an answer to the emerging market failures such as inelastic demand for health plans or risk selection by insurers. In this dissertation, I will discuss questions about potential market imperfections or policy interventions in the healthcare markets in managed competition setting. I discuss the following (research) questions in the subsequent chapters of this dissertation:

1. Price elasticity in the health insurance market: How did price elasticity change in the Dutch health insurance market in the period 1995-2015?

2. Presence of risk selection in the insurance market: Did the possibility to opt for a voluntary deductible, when choosing a health plan, yield adverse selection in the Netherlands in 2013?

3. Welfare effects of increasing substitutability of health plans: What are the welfare effects of increased substitutability between health plans in the health insurance market?

4. Vertical integration and exclusive behavior of insurers and hospitals: What are the welfare effects of vertical restraints between insurers and hospitals?

5. Optimal welfare standard in hospital mergers: How should we apply the consumer welfare standard in case of hospital mergers?

Figure 1.1 depicts the relationship between the three main players in the health care sector, the insurers, consumers and healthcare providers, and places the chapters of this dissertation in this structure. The three vertices of a triangle illustrate the three players while the edges of the triangle depict the three markets, i.e. the insurance market (between consumers and insurers), the health provision market (between consumers and healthcare providers) and the healthcare purchasing market (between insurers and healthcare providers). In the health provision market, consumers do not need to deal with the financial aspect of the transaction. The direct cost in this market is the travelling cost to the provider and the waiting time for the treatment (if any). The financial aspect is detached and takes place in the health purchasing market between the insurer and the provider. Consumers pay for healthcare in the form of a health insurance premium. The premium is independent from the actual healthcare utilization (in the managed care setting) to realize solidarity between sick and healthy.

Chapter 2 and 3 of this dissertation focus exclusively on the insurance market. Chapter 2 measures price elasticity before and after the health insurance reforms in 2006, while Chapter 3 provides evidence for adverse selection due the possibility to opt for a voluntary deductible. In Chapter 4 both health purchasing and health insurance markets play an important role. The chapter explains through the case of substitutability across health plans how the bargaining between insurers and providers influences competitive outcomes on the insurance market. Chapter 5 focuses on the health purchasing market by scrutinizing the incentives for and welfare effects of exclusive behavior and vertical relationships between insurers and hospitals. Finally, Chapter 6 studies the effect of financing the healthcare expenditures through insurance (rather than directly out of packet) on the hospital merger analyses. Difference between the preferences of patients and those of consumers plays an important role in it.

The empirical studies (Chapter 2 and 3) use data from the Netherlands. However, the main conclusions and the theoretical models (Chapter 4, 5 and 6) can also be applied to other health care systems based on the principles of managed care such as the healthcare in Switzerland, Germany or Medicare Advantage, Medicaid managed care and the Marketplaces created by the Affordable Care Act in the United States.

In the subsequent section I will give a summary of the main regulatory changes in the health insurance market in the Netherlands leading to managed competition. After that, I introduce the five papers in this dissertation.

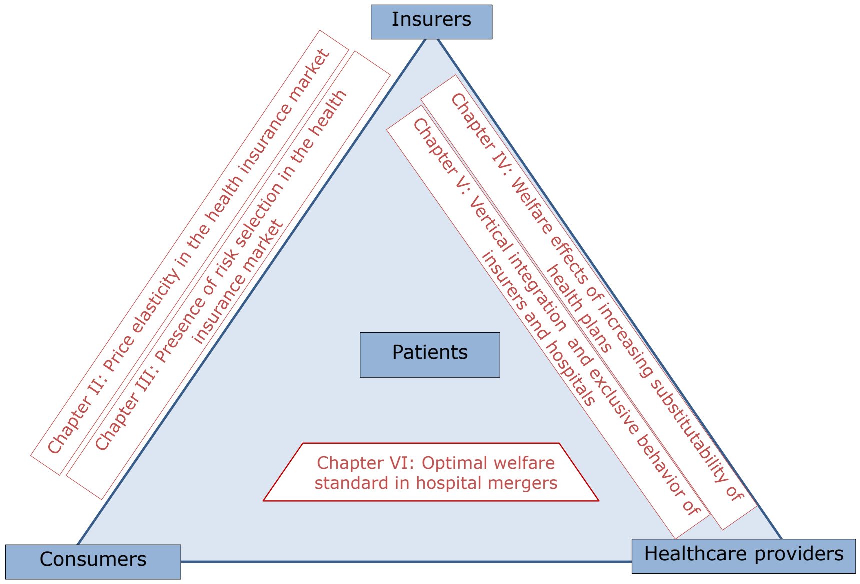


Figure 1.1: Structure of the healthcare sector and chapters of this dissertation

1.2 Managed competition in the Netherlands

During the last thirty years, the Netherlands went a long way towards completing the preconditions of managed competition in the health care markets. The top-down cost-containment policies, such as regulation of doctors, fees and hospital budgets, were replaced gradually by policies building on the principles of competition. Until 2005, reforms were implemented in the mandatory social health insurance scheme for people in lower income brackets (about two-thirds of the population). In 2006, the Health Insurance Act (HIA) was introduced that expended the mandatory basic insurance to the whole population. This basic insurance features open enrollment and community rating, i.e. the obligation to charge the same premium for all enrollees of a given health plan.

In 1986 the Dekker Committee recommended market oriented reforms in health care. From then on, several policy changes were implemented to transform the healthcare system. One of the most important preconditions of successful managed competition is free consumer choice of a health plan (Van de Ven et al., 2013). From 1993 to1995 people were allowed to switch sickness fund once every two years. In this period there were almost no premium differences between health plans. Therefore, switching between insurers was minimal. To facilitate consumer choice, fixed annual open enrollment periods were introduced in

1996. Also, financial risk for the sickness funds was substantially raised (from 3 to 13%), resulting for the first time in meaningful differences in annual premiums. Competition between health plans was however constrained by little room for differentiation. It was barely possible to negotiate different contracts with providers, because provider prices were highly regulated and selective contracting was only permitted for outpatient care. Moreover, sickness funds were not allowed to offer different health plans or to vertically integrate with providers.

In contrast, the HIA, from 2006 on, offered health insurers several options to differentiate basic health insurance contracts in order to increase consumer choice. Insurers were for example allowed to offer a voluntary deductible up to 500 euro per year in return for a premium discount or a group contract at a premium discount of at maximum 10% of a similar individual contract. The HIA also created more opportunities for health insurers to offer preferred or limited provider plans and to manage care by increasing the room for selective contracting and by allowing vertical integration with providers. From 2006 on, the government substantially increased health insurers' financial risk by gradually abolishing ex-post cost reimbursement to health insurers.

Contracting between insurers and healthcare providers also plays an important role in the working of the system. Insurers are responsible to deliver or purchase the care according their enrollees' needs. To achieve efficiency, insurers need sufficient freedom to contract selectively and to negotiate the terms of contracts like price, quality and capacity (Van de Ven et al., 2013). As already mentioned, the HIA gives this freedom to insurers and healthcare providers. For example, insurers and hospitals can freely negotiate over the prices of the most hospital services, adding up to appx. 70% of the annual hospital turnover. The other 30% consists of more complex hospital services, in which segment regulated maximum tariffs are applied.

Because hospital - insurer negotiations play an important role in two of the following chapters, I elaborate on the negotiation process here. Hospital markets, just like insurance markets, are concentrated in the Netherlands (Schut and Varkevisser, 2017). Therefore, the structure of the health purchasing market can be the best described as a two sided oligopoly, and contracting between hospitals and insurers involves bargaining. Ideally insurers negotiate on price, quality of care, capacity (coordinating healthcare delivery in the region) and other services. Their aim is to purchase care of good quality in a cost-efficient way so that they can offer attractive health plans to their enrollees. Because selective contracting is allowed, health plans may differ in the contracted hospital network besides price and additional services.

In the actual situation, there are, however, some deviations from the aimed process of hospital - insurer negotiations (Halbersma et al., 2012; Ruwaard, 2018). First, it is often difficult to find suitable quality indicators to incorporate in the contracts. Although quality is an object of negotiations, it gets a minor role in the contracts due to practical difficulties. Second, insurers and hospitals may not be able to finish the negotiations before the insurer has to announce its policy conditions and premiums. This means that the insurer has to commit itself to a network before the negotiation ends which worsens its bargaining position relative to the hospital. Finally, politics still has a role in determining the overall price increase in the hospital sector by agreeing with insurers and hospitals to keep the turnover-increase, e.g., under 1%. In this way, politics has an effect on the negotiations on the prices, which probably strengthen the position of the insurer in the bargaining.

Van de Ven et al. (2013) derive ten preconditions of effective managed competition from the theoretical model of Enthoven. Afterwards, they evaluate to what extent these preconditions are fulfilled in several countries including the Netherlands. For the most preconditions, the Netherlands get a score of higher than 8 in a scale of 10, where 0 means 'not fulfilled at all', and 10 means 'completely fulfilled'. However, new dilemmas on the most appropriate policy interventions emerge always because of developments on the market, i.e. the reaction of insurers, consumers and healthcare supliers to the changing market conditions, such as technological developments or regulation. My dissertation explores some of these dilemmas, which I outline in the following five sections.

1.3 Price elasticity in the health insurance market

Effective consumer choice, which is a precondition of a well-established price competition, is an important factor in managed competition. Pendzialek et al. (2016) reviews price elasticities in different countries. They conclude that price estimates in similar countries and settings do not differ by much. Moreover, differences between distinct settings can be explained by features such as price level and homogeneity of benefits and coverage.

The introduction of HIA in the Netherlands in 2006, meant a major change in the health insurance system. The first article of this dissertation focuses on price elasticity of health insurance in the period of social insurance (1995-2005) and in the period of the HIA (2006-2015). We expect higher price elasticity in the period of HIA than before that because of the several institutional changes aiming to enhance competition between health insurers. Although proper comparison of the price elasticity in both periods is difficult due to changes in market and institutional characteristics, we see some confirmation of this expectation in the article. At the same time, we find a high level of choice persistence in combination with high forgone potential switching gains in both periods, which indicates a suboptimal choice of consumers. Therefore an active policy to improve health plan choice may be welfare enhancing in this case.

1.4 Presence of risk selection in the health insurance market

Ideally, consumers make conscious decisions in the health insurance market, and choose the health plan that matches their needs the best. Specifically, consumers with a higher probability of costs, opt for more coverage, while consumers with a low-risk profile, prefer plans with low price and consequently less coverage. This mechanism is called adverse selection. As Handel (2013) explains, improved health plan choice goes hand-in-hand with increased adverse selection without an effective risk adjustment system. Such adverse selection decreases solidarity and induces social welfare losses (Cutler and Zeckhauser, 1998). A risk adjustment system aims to level the costs of high-risk and low-risk enrollees. In the case of perfect risk adjustment, an insurer is indifferent in the risk-profile of the enrollee and so there are no incentives for selection.

In the Netherlands there is a sophisticated risk adjustment system in place. Nevertheless, enrollees opting for a voluntary deductible make profit, i.e. they are profitable to insurers in the risk adjustment system and pay less for their insurance than the average. This profit may be the result of moral hazard or adverse selection. In the first case, the terms of the health plan (i.e. lower coverage) make the enrollee use less healthcare and so make lower costs. In the latter case, the healthcare use of the enrollees does not change, but the lower healthcare costs are due to imperfections in the risk adjustment. Exactly because of this interaction of adverse selection and moral hazard, it is challenging to identify the effects of adverse selection empirically (Cohen and Siegelman, 2010).

In the second article of this dissertation, we show that adverse selection emerges in the health insurance market due to possibility to opt for a voluntary deductible. Only enrollees who expect low health care expenditure are better off with a voluntary deductible. By choosing between a health plan with or without a voluntary deductible, enrollees reveal their private information about their expected healthcare expenditure and sort themselves according their risk level. By proving the existence of adverse selection, and due to the fact that enrollees with a high voluntary deductible pay less for their insurance (NZa, 2016), we show that equity, one of the aims of the Dutch health policy, is diminished.

1.5 Welfare effects of increasing substitutability of health plans

Besides adverse selection, increased differentiation in the health plan market has welfare effects through the shift of equilibrium outcomes. As we show in the third article of this dissertation, the negotiation between insurers and hospitals play an important role in the equilibrium level of total welfare and industry profit. If we have a closer look at it, it is logical that the outcome of hospital - insurer negotiations have an effect in the insurance market too. First, contracted hospital network is an important characteristic of health insurance policies. At the moment of buying insurance, consumers also commit themselves to the contracted provider network. Whether a health plan has a narrow or broad network of hospitals, affects the choice of consumers. Second, hospital costs count for a high percentage of health care expenditure. The share of hospital care in the national health expenditures in the US was 32.3% in 2015 (NCHS, 2017). In the Netherlands, 32,5% of expenditure on personal health care was spent on hospital care in 2014 (Bakx et al., 2016). Consequently, reimbursed hospital costs strongly affect the marginal cost of insurers. Finally, hospital prices are typically not exogenously given, but they are a result of negotiations between insurers and hospitals.

In a theoretical two-sided duopoly model, we show that increasing substitutability between health plans does not always increase economic welfare. When the market approaches perfect substitutability of health plans, total and consumer welfare may approach a suboptimal level. This result is driven by the negotiations between hospitals and insurers, and specifically by the overlapping hospital networks of the insurers. Although, each insurer and hospital negotiate independently, hospitals profit from the gains of all insurers they contract. Through the negotiated prices, which determine the marginal costs of the insurers, hospitals can help insurers to commit to a minimum premium. Because hospitals contract more competing insurers, the negotiated prices become implicitly coordinated. This commitment and co-ordination enable insurers to soften the competition in the health insurance market. In our model, premiums go up when health plans become closer substitutes if the original level of substitutability is already substantial. Higher premiums lead consequently to higher industry profit and lower consumer (and total) welfare. This welfare effect of competition is contradictory to the conventional wisdom that more intensive competition increases welfare. It also calls the attention to the special case of healthcare markets with a tight relationship between the providers and the health plans market.

In our model, there are no assumptions that are specific to healthcare, e.g. we apply a downward sloping demand curve, which assumes that consumers may stay uninsured. In the Dutch context, this modelling fits the supplementary insurance market. Due to this general approach, the same model could be applied to describe the competition in, for example, the car-glass insurance market as well. Furthermore, the idea of commitment and co-ordination through contracts appears in the theoretical literature of strategic delegation and strategic separation too (for an overview see Kopel and Pezzino, 2018). As the literature of strategic delegation describes, manufacturers (our hospitals) may be able to soften competition in their product market (health insurance market) by delegating the price setting to retailers (insurers) instead of selling the products themselves. This indicates that the essential ideas of our model have possibilities of broader application than the (health) insurance market.

1.6 Vertical integration and exclusive behavior of insurers and hospitals

Managed competition is the competition of delivery systems, not only health plans, according to the ideas of Enthoven (1993). Freedom in contracting and integration gives more room for innovative, efficiency enhancing forms of health care delivery. Vertical integration, according to Enthoven, makes total quality management and continuous quality improvement possible throughout the whole health care delivery chain. In recent years, we see indeed numerous initiatives of vertical integrations or joint ventures of a health insurer and a health care provider in the U.S. (Abelson, 2017; Zimlich, 2016). Vertical restraints and integration are also a tool to increase efficiency or gain competitive advantage.

Vertical restraints and integration may however also have anticompetitive effects. In their overview of pros and cons of vertical relations in healthcare, Bijlsma et al. (2008) argue that vertical relationships may result in anticompetitive foreclosure of competitors, but only in the presence of market power in the insurance and/or hospital market. At the same time, market power is increasingly present in the US market places and also in hospital and insurance markets in the Netherlands[1].

[1] In the US, the mean HHIs for hospitals and specialist physician organizations each increased by about 5 percent from 2010 to 2016. As a result, 90 percent of the Metropolitan Statistical Areas had an HHI for hospitals more than 2500 in 2016 (Fulton, 2017). U.S. GAO (2016) found that enrollment in private health insurance plans concentrated among a small number of issuers in most states in 2014, including in the newly established exchanges. Several exchanges established by the Patient Protection and Affordable Care Act had fewer than three issuers participating (U.S. GAO, 2016). Recently proposed (e.g., between Aetna and Humana, and Anthem and Cigna) and consummated (e.g., between Centene and Health Net) national insurer mergers make the insurance market even

In Chapter 5 of this dissertation, we model the highly concentrated insurer and hospital market as oligopolistic markets where insurers bilaterally bargain with hospitals. We focus on the bargaining model with two insurers and two hospitals and examine under which market conditions exclusive behavior and vertical integration can arise and whether this may harm competition. The applied bargaining game between insurers and hospitals is very similar to the game used in the previous chapter of this dissertation. The most important difference between the models lies in the modelling of the competition in the insurance market. While in the previous chapter, we assumed a downward sloping demand curve, here we assume the fixed demand of Hotelling competition. This choice affects the bargaining game as well, because the negotiated prices in the two-sided duopoly are undetermined in case of fixed demand. Therefore, we assume here regulated prices between the insurers and hospitals. Although these assumptions are more restrictive, they fit better the mandatory insurance environment in the Netherlands.

We show that two types of exclusive behavior can occur in this setting. First, one of the insurers may be excluded. This happens if decreasing competition (i.e., monopolization) in the insurance market leads to a substantial increase in total industry profit. Second, one hospital may not, in equilibrium, engage in a contract with one of the insurers. In such a case, a narrow-network insurer (contracting one hospital) and a broad-network insurer (contracting both hospitals) can coexist in the market, even if both insurers and both hospitals are equally efficient in their production. We show that the range of parameters under which the latter outcome can occur grows if one insurer – hospital pair integrates vertically.

The model in this chapter has some features that describe specific characteristics of a mandatory health insurance market, such as fixed demand. Nevertheless, the model and conclusions fit well in the broader literature of vertical integration and foreclosure.[2] Specifically, it contributes to the literature analyzing the potential anti-competitive effects of a vertical merger. In the beginning of the 1990s, there were several articles published showing that a vertical merger may

more concentrated. In the Netherlands, the market share of the four largest insurers is above 88% in 2017 (NZa, 2017), and an average hospital has a market share of about 50 percent in its catchment area (Schut and Varkevisser, 2017).

[2]Although we analyze exclusive dealing in equilibrium, our paper is not related to the literature of exclusive contracts analyzed in e.g. Aghion and Bolton (1987) and Rasmusen et al. (1991). These latter papers model naked exclusion, which means the outright exclusion of third parties. Our modelling of exclusivity is related to the renegotiable exclusive contracts of Segal and Whinston (2000). In that model parties sign exclusive contracts ex-ante but they may deviate from that contract (after renegotiation) whenever it is ex-post efficient to contract a third party. This idea corresponds to our assumption of mutual profitability of each signed contract. Exclusive dealing occurs in our model, when contracting is jointly unprofitable for the negotiating parties.

result in the foreclosure of rivals. Ordover et al. (1990) and Salinger (1988) show that an integrated firm may foreclose product market competition by raising rivals costs. The mechanism behind foreclosure lies in the assumptions of (i) the credible pre-commitment of the integrated supplier that it will not supply downstream rivals even if it would have an incentive to do so ex-post, and (ii) linear tariffs, which are inefficient because of double marginalization. Hart et al. (1990), and O'Brien and Shaffer (1992) describe another mechanism that may result in foreclosure, namely secret contracting.

A recent paper, Levy et al. (2018) analyzes the case of partial vertical integration. In particular, they consider the incentive to partially integrate and then foreclose rivals when the target has initially two controlling shareholders. When one of the shareholders is passive, the change in its profit is not considered in the decisions over a merger or foreclosure. Foreclosure arises for a larger parameter range, than in the case of full integration, because the profit decrease of the passive shareholder 'subsidizes' the foreclosure of rivals. Another paper, Nocke and Rey (2018), analyzes market structures with interlocking relationships, i.e. when the same upstream and downstream firms trade with each other, with differentiated products. Nocke and Rey (2018) show that firms can have incentives to engage in vertical foreclosure in order to exert market power at the expense of consumers and society.

Our paper is the closest to Nocke and Rey (2018), however, our modelling of bargaining differ (e.g. we assume that the signing a contract is common knowledge, while Nocke and Rey (2018) assumes it is private information) and we apply some healthcare specific characteristics such as fixed demand, and price competition on the downstream (i.e. insurance) market instead of quantity competition. In both papers, contracting externalities play an important role in the incentives for foreclosure. Specifically, the total industry profit can be increased by foreclosure, which means the increase of the surplus that is divided between the players.

1.7 Optimal welfare standard in hospital mergers

The last chapter in this dissertation elaborates on the ambiguous definition of consumer and the consequences of it in the merger analysis of competition authorities. The main idea is that mandatory health insurance detaches the financial and consumption side of health care markets. The group of consumers paying for a particular health insurance is always broader than the group that actually benefits from the healthcare services. The narrowest definition of consumer is 'patients directly affected by quality changes'. We expect that patients are less concerned with the potential costs of a merger because expenditures are spread across all enrollees of the insurance which is typically a larger group than the patients of the hospital. Widening the considered group, we can include potential clients of hospitals in the analyses. External effects due to insurance still can be present if insurers have enrollees who are potentially patients on (and consequently spread health care expenditures across) more hospital markets. A complete consumer welfare analysis would embrace all consumers covered by the same insurance.

Because of the ambiguity which group of consumers to include in the merger analysis, the conventional result that the consumer welfare standard is more restrictive than the total welfare standard can be reversed on the healthcare markets. In the practice of merger cases, the ambiguity takes the form of diverging value that consumers attach to given improvements in health care provision such as a quality improving investment in a hospital. Consumers' valuation depends on their current situation and expected future needs. The heterogeneity across consumers has to be taken into account in a welfare analysis of hospital mergers.

The conclusions of Chapter 6 can be generalized to other industries where significant externalities may emerge. We could take examples from environmental economics (assuming that the merger influences the magnitude of the externality, e.g. pollution), from the financial sector (thinking on the effect of a merger on the stability of the whole system) or from the insurance sector. When facing a merger case with significant external effects, competition authorities should not lean on general results but take into account the peculiarities of the market.

1.8 Outline of the dissertation

The conclusions in the five subsequent chapters of this dissertation contribute to a better understanding of the health insurance markets in managed competition setting. They also help to find the adequate regulatory response to the emerging market imperfections. The desired level of efficiency and equity in the health insurance market can only be achieved with effective competition and adequate regulation.

This dissertation is formed by the following papers:

- Chapter 2: Douven, R., Katona, K., Schut, F. T., and Shestalova, V. Switching gains and health plan price elasticities: 20 years of managed competition reforms in the netherlands. *The European Journal of Health Economics*, pages 1–18, 2017

- Chapter 3: Croes, R., Katona, K., Mikkers, M., and Shestalova, V. Evidence of selection in a mandatory health insurance market with risk adjustment. Discussion paper 2018-013, TILEC, Tilburg University, 2018
- Chapter 4: Katona, K. and Halbersma, R. Competition and bargaining in healthcare markets. *Working paper*, 2018
- Chapter V: Douven, R., Halbersma, R., Katona, K., and Shestalova, V. Vertical integration and exclusive behavior of insurers and hospitals. *Journal of Economics & Management Strategy*, 23(2):344–368, 2014
- Chapter 6: Katona, K. and Canoy, M. Welfare standards in hospital mergers. *The European Journal of Health Economics*, 14(4):573–586, 2013

Bibliography

Abelson, R. As health care changes, insurers, hospitals and drugstores team up. *The New York Times*, 26-11-2017, 2017. https://www.nytimes.com/2017/11/26/health/health-care-hospitals-insurers.html.

Aghion, P. and Bolton, P. Contracts as a barrier to entry. *The American economic review*, pages 388–401, 1987.

Bakx, P., O'Donnell, O., and Doorslaer, E. Spending on health care in the netherlands: not going so dutch. *Fiscal Studies*, 37(3-4):593–625, 2016.

Bijlsma, M., Meijer, A., and Shestalova, V. Vertical relationships between health insurers and healthcare providers. Technical report, CPB Netherlands Bureau for Economic Policy Analysis, 2008.

Cohen, A. and Siegelman, P. Testing for adverse selection in insurance markets. *Journal of Risk and Insurance*, 77(1):39–84, 2010.

Croes, R., Katona, K., Mikkers, M., and Shestalova, V. Evidence of selection in a mandatory health insurance market with risk adjustment. Discussion paper 2018-013, TILEC, Tilburg University, 2018.

Cutler, D. M. and Zeckhauser, R. J. Adverse selection in health insurance. In *Frontiers in Health Policy Research, Volume 1*, pages 1–32. MIT Press, 1998.

Douven, R., Halbersma, R., Katona, K., and Shestalova, V. Vertical integration and exclusive behavior of insurers and hospitals. *Journal of Economics & Management Strategy*, 23(2):344–368, 2014.

Douven, R., Katona, K., Schut, F. T., and Shestalova, V. Switching gains and health plan price elasticities: 20 years of managed competition reforms in the netherlands. *The European Journal of Health Economics*, pages 1–18, 2017.

Enthoven, A. Theory and practice of managed competition in health care finance, 1988.

Enthoven, A. C. Consumer-choice health plan: inflation and inequity in health care today: alternatives for cost control and an analysis of proposals for national health insurance. *New England Journal of Medicine*, 298(12):650–658, 1978.

Enthoven, A. C. The history and principles of managed competition. *Health affairs*, 12(suppl 1):24–48, 1993.

Fulton, B. D. Health care market concentration trends in the united states: evidence and policy responses. *Health Affairs*, 36(9):1530–1538, 2017.

Halbersma, R., Van Manen, J., and Sauter, W. De verzekeraars als motor van het zorgstelsel. *Zorg en financiering*, 11(10):10–23, 2012.

Handel, B. R. Adverse selection and inertia in health insurance markets: When nudging hurts. *The American Economic Review*, 103(7):2643–2682, 2013.

Hart, O., Tirole, J., Carlton, D. W., and Williamson, O. E. Vertical integration and market foreclosure. *Brookings papers on economic activity. Microeconomics*, 1990:205–286, 1990.

Katona, K. and Canoy, M. Welfare standards in hospital mergers. *The European Journal of Health Economics*, 14(4):573–586, 2013.

Katona, K. and Halbersma, R. Competition and bargaining in healthcare markets. *Working paper*, 2018.

Kopel, M. and Pezzino, M. Strategic delegation in oligopoly. In Corchón, L. C. and Marini, M. A., editors, *Handbook of Game Theory and Industial Organization: Volume II, Applications*, chapter 10, pages 248–285. Edward Elgar Publishing, Cheltenham, UK, 2018.

Levy, N., Spiegel, Y., and Gilo, D. Partial vertical integration, ownership structure, and foreclosure. *American Economic Journal: Microeconomics*, 10(1): 132–80, 2018.

NCHS, N. C. f. H. S. *Health, United States, 2016: With Chartbook on Long-term Trends in Health.* National Center for Health Statistics, Hyattsville, MD. 2017., 2017.

Nocke, V. and Rey, P. Exclusive dealing and vertical integration in interlocking relationships. *Journal of Economic Theory*, 177:183–221, 2018.

NZa. Kwantitatief onderzoek naar risicoselectie en risicosolidariteit op de zorgverzekeringsmarkt. Report, Nederlandse Zorgautoriteit, 2016.

NZa, N. Z. Marktscan zorgverzekeringsmarkt 2017. Technical report, 2017. https://puc.overheid.nl/nza/doc/PUC_3656_22/1/.

O'Brien, D. P. and Shaffer, G. Vertical control with bilateral contracts. *The RAND Journal of Economics*, pages 299–308, 1992.

Ordover, J. A., Saloner, G., and Salop, S. C. Equilibrium vertical foreclosure. *The American Economic Review*, pages 127–142, 1990.

Pendzialek, J. B., Simic, D., and Stock, S. Differences in price elasticities of demand for health insurance: a systematic review. *The European Journal of Health Economics*, 17(1):5–21, 2016.

Rasmusen, E., Ramseyer, J., and Wiley Jr, J. Naked exclusion. *The American Economic Review*, 81(5):1137–1145, 1991.

Ruwaard, S. *Purchasing healthcare: Beyond getting the financial incentives right*. PhD thesis, Tilburg University, 2018.

Salinger, M. A. Vertical mergers and market foreclosure. *The Quarterly Journal of Economics*, 103(2):345–356, 1988.

Schut, F. T. and Varkevisser, M. Competition policy for health care provision in the netherlands. *Health Policy*, 121(2):126–133, 2017.

Segal, I. and Whinston, M. Exclusive contracts and protection of investments. *RAND Journal of Economics*, pages 603–633, 2000.

U.S. GAO, U. G. A. O. Private health insurance: In most states and new exchanges, enrollees continued to be concentrated among few issuers in 2014. Technical Report GAO-16-724, 2016.

Van de Ven, W. P., Beck, K., Buchner, F., Schokkaert, E., Schut, F. E., Shmueli, A., and Wasem, J. Preconditions for efficiency and affordability in competitive healthcare markets: Are they fulfilled in belgium, germany, israel, the netherlands and switzerland? *Health policy*, 109(3):226–245, 2013.

Zimlich, R. Trend to watch: Payer-provider joint ventures. *Managed Healthcare Executive*, 3-12-2016, 2016. http://managedhealthcareexecutive.modernmedicine.com/managed-healthcare-executive/news/trend-watch-payer-provider-joint-ventures.

Chapter 2

Switching gains and health plan price elasticities: 20 years of managed competition reforms in The Netherlands

Rudy Douven, Katalin Katona, Erik T. Schut, Victoria Shestalova

This chapter was published as:
Douven, R., Katona, K., Schut, F. T., and Shestalova, V. Switching gains and health plan price elasticities: 20 years of managed competition reforms in the netherlands. *The European Journal of Health Economics*, pages 1–18, 2017

There are three changes made in the text with respect to the published version:

- In the second sentence of footnote 20, the word "downward" is deleted. Also from the sentence that refers to footnote 20, the word "downward" is deleted.
- In the last sentence of the second paragraph under equation (5), it is added that GMM controls for the endogeneity of $log(s_{i,t-1})$ as well.
- Footnote 16 is added.

Abstract

In this paper we estimate health plan price elasticities and financial switching gains for consumers over a 20 years period in which managed competition was introduced in the Dutch health insurance market. The period is characterised by a major health insurance reform in 2006 to provide health insurers with more incentives and tools to compete, and to provide consumers with a more differentiated choice of products. Prior to the reform, in the period 1995-2005, we find a low number of switchers, between 2 and 4% a year, modest average total switching gains of 2 million euros per year and short-term health plan price elasticities ranging from -0.1 to -0.4. The major reform in 2006 resulted in an all-time high switching rate of 18%, total switching gains of 130 million euro, and a high short-term price elasticity of -5.7. During 2007-2015 switching rates returned to lower levels between 4 and 8% per year, with total switching gains in the order of 40 million euros per year on average. Total switching gains could have been 10 times higher if all consumers had switched to one of the cheapest plans. We find short-term price elasticties ranging between -0.9 and -2.2. Our estimations suggest substantial consumer inertia throughout the entire period, as we find degrees of choice persistence ranging from about 0.8 to 0.9.

Keywords: managed competition, health insurance, health plan price elasticity, switching gains. JEL-code: I18, C33.

2.1 Introduction

In health care systems with a competitive health insurance market sufficiently price-elastic demand is important for motivating health insurers to act as cost-conscious purchasing agents on behalf of their customers. A recent systematic review of empirical studies on price elasticity of health plan choice identified clear-cut price elasticity ranges for different country settings but substantial variation in price elasticities across various countries (Pendzialek et al., 2016). For the Netherlands, where competition among health insurers was introduced within the social health insurance (SHI) scheme in 1996, the review study found short-term price elasticities smaller than -0.5, which were well below most of those found in other countries.[1] As noticed by Pendzialek et al. (2016), however, evidence about the Netherlands is dated, since the empirical studies only relate to the situation before a major health insurance reform in 2006, and almost no information could be found on price elasticities in the years following the reform. This limitation is particularly troublesome because the primary goal of the reform was to enhance consumer choice and competition in order to reinforce insurers' incentives to improve the efficiency of care.

The main contribution of this paper is to fill this gap in the empirical literature by estimating the price elasticity of health plan choice in the Netherlands after the major reform in 2006. Using data on prices and market shares of all health plans over the period 2005-2015, we examine whether price elasticities of health plan choice increased relative to the low price elasticities prior to the reform. For a good comparison between the two periods, we re-estimated the price elasticities for the entire pre-reform period 1995-2005. This is because previous empirical studies use different methodologies and typically cover only part of the pre-reform period. As noticed by Pendzialek et al. (2016), health plan price elasticities are difficult to compare because of the differences in methodologies and data sources of the included studies. Therefore, a second important contribution is that we provide consistent estimates of health plan price elasticities using the same methodology and data over a 20-year period. We are not aware of any other study that consistently estimated annual health plan choice over such a long period.[2] Third, we contribute to the literature by also calculating the annual net financial switching gains for consumers over a 20-years period, uncovering

[1]Short-term health plan price elasticites in the SHI market have been estimated before in several empirical studies, covering the years before 2000 (Schut and Hassink, 2002; Schut et al., 2003), the period until 2002 (Douven et al., 2007) and the same SHI period as in this paper (van Dijk et al., 2008).

[2]In the systematic review by Pendzialek et al. (2016) most of the 41 included empirical studies cover only a few years. Only three studies cover more that 10 years but include various institutional settings over the years.

also the sources of these gains. This provides a unique indication about the extent to which consumers financially benefited from switching and how these benefits changed over time. Therefore, our findings may offer important insights for health policy on how to influence consumer choice and price competition in health insurance markets.

Our paper is organised as follows. In "Overview of the Dutch hgealth insurance market 1995-2015" we describe the main differences between the pre-reform and the post-reform health insurance market in the Netherlands. Section "Financial switching gains for premium payers" discusses the financial switching gains for premium payers. Section "Model and estimation methods" explains the estimation methods and empirical strategy. In "Data" we describe the data and in "Estimation results" the estimation results. Section "Conclusion" concludes.

2.2 Overview of the Dutch health insurance market 1995-2015

2.2.1 SHI-market 1995-2005

In the past 20 years the Dutch health insurance system gradually moved towards a system of managed competition. Until 2005, health insurance for basic health care services consisted of a mandatory social health insurance (SHI) scheme for people in lower income brackets (about two-thirds of the population) and a voluntary private health insurance system for people with a higher income.[3] The SHI scheme was administered by sickness funds (not-for profit health insurers). Health care expenses were largely covered by income-related contributions that were collected in a central fund and then redistributed to sickness funds. The share of income-related contributions as a percentage of total expenses was about 90% until 2002, and was reduced to about 80% in 2003. As a result in 2003 the annual community-rated premium increased from about 10 to 20% of total expenses (see row “Out of pocket premiums / total cost (%)” in Table 2.1).[4] To cover the

[3]The SHI scheme can be regarded as the precursor of the HIA scheme introduced in 2006 because both schemes have many features in common (i.e. both are mandatory insurance schemes with a comprehensive standardized benefits package, partly income-related and partly community-rated premiums, and carried out by competing health insurers). Therefore, we compare the SHI with HIA here and do not consider the voluntary private health insurance scheme. Private indemnity health insurance covered about a third of the Dutch population with earnings above a legally determined income threshold. Benefit packages were similar to that of SHI, although there was substential variation in both the scope of benefits and cost-sharing arrangements. Enrollment was voluntary, premiums were risk-rated and medical underwriting was allowed. For an extensive description of the private insurance market, see Tapay and Colombo (2004), and for the SHI market, see Douven and Schut (2011).

[4]Income-related contibutions were annually set by the government as a precentage of gross income up to an annually adjusted threshold (about 32,600 euros in 2005).

residual costs, sickness funds were allowed to charge an annual community-rated premium (Table 2.2). Since 1993 sickness funds were increasingly put at risk for the medical expenses of their enrollees, by gradually replacing retrospective reimbursement by risk-adjusted capitation payments. In addition, the former legally protected regional monopolies were abolished and sickness funds were allowed to compete for customers all over the country. Eligible people were allowed to change sickness funds, and sickness funds were obliged to accept all applicants.[5]

As shown in Table 2.1, in 1996 the financial risk for the sickness funds was substantially raised (from 3 to 13%), resulting for the first time in meaningful differences in annual premiums.[6] For this reason we chose 1995 as the starting year for estimating health plan price elasticities and switching gains. Incentives for price competition among sickness funds were gradually reinforced by stepwise increasing sickness funds' risk on medical expenses. This resulted in an increasing premium variation across sickness funds (Table 2.2).[7]

Next to premiums, sickness funds had only limited room to distinguish themselves. The room for negotiating different contracts with providers was almost nonexistent, since provider prices were highly regulated and selective contracting was only permitted for outpatient care. Moreover, sickness funds were not allowed to offer different health plans or to vertically integrate with providers. Five small sickness funds entered the market in the early years, but after 1998 only mergers took place and the number of sickness funds decreased from 26 to 21 in 2005 (Table 2.1 and Appendix 2.B).

Sickness funds also provided supplementary health insurance, comprising about 5% of total revenues. Supplementary coverage typically includes dental care for adults, physiotherapy and medical appliances, such as spectacles and hearing aids. Supplementary and basic health insurance are often sold together so consumers might base their decisions to switch also on the combined premium (Schut and Hassink, 2002).[8]

[5]From 1993 to 1995 people were allowed to switch once every 2 years. To facilitate consumer chice, since 1996 fixed annual open enrollment periods were introduced.

[6]From 1993 to 1995, except for one small sickness fund, all sickness funds charged the same annual premium.

[7]An increase in price competition may also result in lower premiuim variation. However, since premium competition was absent before 1996 we interpret the increase in premium variation as a sign of increasing price competition.

[8]Supplementary benefit packages were quite similar across health insurers. Schut and Hassink (2002) found a somewhat higher price elasticity for combined health insurance (-0.4) than basic health insurance (-0.3).

Year	Social Health Insurance (about two-third of population)											Health Insurance Act (total population)									
	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Population size of total market (million) [a]	9.7	9.8	9.9	9.9	9.9	10.3	10.3	10.2	10.1	10.2	10.1	16.3	16.4	16.3	16.4	16.5	16.6	16.7	16.7	16.8	16.8
Population of premium payers (million) [b]	7.7	7.7	7.9	8.0	8.1	8.3	8.3	8.2	8.2	8.2	8.2	12.5	12.6	12.8	12.9	13.0	13.1	13.2	13.3	13.3	13.4
Total number of health insurers[c]	26	27	29	29	29	26	24	21	21	21	21	33	32	32	30	28	27	26	26	26	25
Number of health insurers leaving/merging	1	0	0	2	0	3	2	3	0	0	0	n.a.	1	0	2	2	1	1	0	0	1
Number of health insurers entering	0	1	2	2	0	0	0	0	0	0	0	n.a.	0	0	0	0	0	0	0	0	0
Annual premiums / total cost (%)[d]	8	10	6	5	10	10	9	10	22	19	22	50	50	50	50	50	50	50	50	50	50
Insurers risk on medical expenses (%) [e]	3	13	27	28	35	36	38	41	52	53	53	53	53	52	60	73	70	88	90	92	99
Number of different insurance policies	26	27	29	29	29	26	24	21	21	21	21	46	53	60	56	57	56	60	65	67	71
Number of limited provider plans	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	5	5	8	10	12	17
Population share limited provider plans (%)	0	0	0	0	0	0	0	0	0	0	0	0	0	0.1	0.2	0.5	1.1	1.7	3.3	4.4	7.5
Population share with group contracts (%) [f]	0	0	0	0	0	0	0	0	0	0	0	53	57	60	60	64	66	67	69	70	69

If not otherwise indicated, the data was obtained from the Dutch Healthcare Authority (NZa, 2007, 2009, 2011, 2012, 2014, 2015) and the Dutch National Healthcare Institute (ZIN).
[a] The private insured are excluded from the population for 1995-2005.
[b] The share of actual premium payers is about 80 percent of the total number of enrolees because insurance for children under 18 was free (financed by taxation).
[c] From 1995-2005 they were called sickness funds (not-for-profit health insurers). After 2006, also profit insurers were on the market.
[d] Authors' own calculations. As of 2006 the government mandated that 50% of the total cost should be paid in the form of annual premiums charged by the health insurer to the individual consumers (of 18 years and older).
[e] Insurance risk on medical expenses from Van Kleef et al. (2013) and for 2013-2015 from personal communication with René van Vliet.
[f] Before 2006 there were many group contracts in the private insurance but not in the SHI market.

Table 2.1: Characteristics of the population and health insurers in the Dutch SHI and HIA markets

Year	Social Health Insurance (about two-third of population)											Health Insurance Act (total population)									
	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Average annual premium (euro) [a]	89.8	155.7	98.2	97.9	178.8	189.7	163.6	182.6	344.7	304.6	378.1	1037	1115	1077	1080	1118	1214	1243	1232	1111	1168
Weighted premium, before switching (euro) [b]	89.9	155.4	97.7	97.7	178.7	187.7	157.0	181.5	355.7	307.5	384.8	1036	1105	1050	1059	1099	1201	1229	1217	1102	1163
Weighted premium, after switching (euro) [c]	89.9	155.4	97.7	97.7	178.6	187.5	156.8	180.4	355.2	307.0	384.0	1025	1103	1049	1059	1095	1199	1226	1213	1098	1158
Weighted premium individual contracts (euro) [d]	89.8	155.7	98.2	97.9	178.8	189.7	163.6	182.6	344.7	304.6	378.1	1056	1144	1091	1094	1137	1245	1275	1262	1137	1195
Weighted premium group contracts (euro) [d]	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	998	1070	1019	1036	1070	1174	1201	1190	1080	1142
Standard deviation annual premiums (euro)	0	4	8	8	12	17	26	25	33	31	41	46	51	56	60	60	60	60	61	70	82
Number of switchers across insurers (%)	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	2.6	n.a.	2.8	2.4	4.2	17.8	4.4	3.6	3.6	4.3	5.5	6.0	8.2	7.0	7.3
Total switching gains (million euro) [e]	0.0	0.1	0.2	0.4	0.8	1.6	1.3	1.2	4.3	3.7	6.7	130.0	31.8	26.6	2.6	51.1	45.1	44.6	49.2	53.0	53.8
Average switching gains premium payer (euro) [f]	0.0	0.0	0.0	0.1	0.1	0.2	0.2	0.1	0.5	0.5	0.8	10.4	2.5	2.1	0.2	3.9	3.5	3.4	3.7	4.0	4.0
Average gain per switcher (euro) [g]	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	5	n.a	15	15	16	45	44	45	4	72	50	45	36	45	44
Potential total switching gains (million euro) [h]	5	52	117	119	201	230	251	321	547	416	523	400	498	508	431	483	552	692	376	873	636
Potential average gain per premium payer (euro) [i]	1	5	12	12	20	22	24	31	54	41	51	32	39	40	33	37	42	52	28	65	47

If not otherwise indicated, the data was obtained from the Dutch Healthcare Authority [10, 33-38] and the Dutch National Healthcare Institute (ZIN).
[a] Calculated as unweighted average annual premiums that each consumer pays to its health insurer (group and individual contracts taken together).
[b] Annual premiums weighted by insurers' market shares in year t-1 (before switching) (group and individual contracts taken together, see also Appendix 2.A).
[c] Annual premiums weighted by insurers' market shares in year t (after switching) (group and individual contracts taken together, see also Appendix 2.A).
[d] Calculated as weighted average premiums (weighted with health insurers' market shares).
[e] Switching gains are calculated as the difference between the weighted premiums (after and before switching)
multiplied by the number of premium payers in year t (see also Appendix 2.A).
[f] Calculated as the total switching gains divided by the number of premium payers in the population (see table 1).
[g] Calculated as the total switching gains divided by the number of switchers across insurers.
[h] Calculated as the potential average gain per premium payer (last row) times the total population of premium payers.
[i] SHI: Calculated as the difference between the (weighted) average premium in the market in a year and a (weighted) average premium of the five cheapest insurers in the same year.
HIA: The same calculation, however, first we calculated for each insurer the (weighted) average premium over all individual and group contracts.

Table 2.2: Annual premiums and switching characteristics in the Dutch SHI and HIA markets

From 2001 to 2005 about 2-4% of the enrollees annually switched to another sickness fund (Laske-Aldershof and Schut, 2005). Although the number of switchers from earlier years is lacking, the percentage of switchers from 1996-2000 is likely to be lower since the premium differences across sickness funds were small (Table 2.2) and many consumers might not even have been aware of the possibility of switching.[9]

2.2.2 Health Insurance Act 2006-2015

In 2006 the scope of the managed competition model was broadened to the entire population by the introduction of a new Health Insurance Act (HIA). Former sickness funds and former private indemnity insurers were allowed to compete for providing basic health insurance to all Dutch citizens. The basic idea behind this reform was to increase efficiency by promoting more competition among health insurers and among health care providers. To preserve universal access and maintain equity the government followed a setup along the lines of the SHI, including mandatory insurance for a standardized basic benefit package, a partly community-rated and partly income-related premium, open enrollment and a risk adjustment system.

The government substantially increased health insurers' financial risk by further abolishing ex-post cost reimbursement to health insurers. As shown in Table 2.1, for all health insurers the financial risk on medical expenses was gradually raised from 53% in 2005 to 99% in 2015.

In addition in 2006, the share of income-related premiums in total health care expenditure was reduced from 80% (SHI) to 50% (HIA), and this latter share is fixed by law (see row "Annual premiums / total cost (%)" in Table 2.1). This implied a significant increase of the annual premium for people previously enrolled in SHI from about 380 euro in 2005 to about 1050 euro in 2006 (see Table 2.2). The idea of policymakers behind this change was that a higher annual premium would make people more aware and cost-conscious of the high health care costs. To maintain equity, households with earnings below a certain threshold were compensated by monthly income-dependent subsidies.

Both sickness funds and private insurers were allowed to offer health insurance under the HIA. In 2006 basic health insurance was offered by 33 health insurers, but due to mergers and consolidation this number decreased to 25 in 2015, while no new insurers entered the market during this period (Table 2.1 and Appendix 2.B).

In contrast to the former SHI scheme, the HIA offered health insurers several options to differentiate basic health insurance contracts to increase consumer

[9] In 2001, 83% of the sickness fund enrollees responded in a survey that they had not even considered switching. This share declined to 77% in 2005 (Laske-Aldershof and Schut, 2005).

choice. First, insurers were allowed to offer a voluntary deductible up to 500 euro per year in return for a premium discount. Next, health insurers were also allowed to offer group contracts at a premium discount of at maximum 10% of a similar individual contract. Third, health insurers were allowed to provide coverage in terms of service benefits, indemnity payments and a combination of both. Fourth, the HIA created more opportunities for health insurers to offer preferred or limited provider plans and to manage care by increasing the room for selective contracting and by allowing vertical integration with providers (Van de Ven and Schut, 2008).

The introduction of the HIA had a large impact on the health insurance market. In the first year of the reform health insurers engaged in a premium war.[10] Although people were not forced to switch health insurers since basic health plans were offered by both former sickness funds and private health insurers, for all people the choice setting and choice options radically changed. The massive media coverage around the reform, combined with a large increase in choice for different benefit packages and large premium differences, made many people aware of potential switching benefits. Hence, many people were triggered to reconsider their previous choice of health insurer. The threat of many customers making a cost-conscious choice forced insurers to offer contracts at annual premiums below the break-even price, resulting in substantial losses by insurers in 2006 (Douven and Schut, 2006). This effectuated an all-time-high switching rate of 17.8% in 2006 (Table 2.2).[11] Such a high switching rate was far above what was experienced before in the Dutch health insurance markets (including the former private health insurance market). As shown in Table 2.2, during the first four years after 2006, switching rates between health insurers dropped from 18 to about 4%, but then increased to about 6-8% during the next five years.

Most health insurers offer both individual and group contracts. Group contracts can be concluded with any legal entity, and in total more than 50,000 group contracts are concluded annually with a huge variety of groups (NZa, 2015).[12] Table 2.2 shows that group contracts are on average 50 to 70 euro per

[10]Health insurers have to announce new health plan premiums each year before November 20 and enrollees that are willing to switch have to notify their insurer before the end of the year that they want to terminate the contract. Every year since 2006 the same small regional insurer is the first to announce its premium early in October, attracting a lot of free publicity. Until 2015 most health insurers announced new health plan premiums 10-20 days before the deadline, but in 2015 about half of the insurers waited until the last week and the variation in announcement dates decreased (NZa, 2015). The health insurer with the lowest health plan premium typically waits until all other health plan premiums are known, in order to be sure of being the cheapest health plan.

[11]In addition to the 18% of the population switching between health insurers, 5-10% changed health plans within their insurer, so including these intra-insurer switchers raises the total number of switchers in 2006 to 23-28% (NZa, 2006).

[12]In 2015, 56% of group contracts were employer-based, while other group contracts were concluded with a large variety of entities, such as labor unions, sport federations, cooperative banks, and interest

year lower priced than individual policies (implying a premium discount of about 5%). Most Dutch people have several options to join a group contract and the share of the population opting for a group contract increased from 53% in 2006 to 69% in 2015 (Table 2.1).

In addition to mandatory basic insurance, most people (about 85%) also bought voluntary supplementary insurance, just as in the former SHI market (NZa, 2015). Although people can buy basic and supplementary coverage from different insurers, almost none did (only 0.19% of those buying supplementary insurance) (NZa, 2015). As in the SHI market, the most important supplementary benefits still are dental care for adults and physiotherapy, but the variation in coverage substantially increased. Premiums for supplementary insurance plans are on average about 20-25% of those of basic health insurance plans (Vektis, 2016). Medical underwriting is allowed, but in practice only required for 5% of all supplementary policies (typically the most extensive ones) and for 24% of dental insurance policies (NZa, 2015). Since almost all consumers buy supplementary and basic health insurance together, high-risk individuals may be restricted in choosing basic health plans by the underwriting practices of health insurers with respect to supplementary insurance. Indeed, several studies found that a substantial number of elderly and high-risk individuals do not switch to another insurer because of they believe that they will not be accepted for supplementary insurance by another insurer (Roos and Schut, 2012; Duijmelinck and van de Ven, 2014)]. Boonen et al. (2016) found that having supplementary insurance significantly reduces older people's switching propensity.

Health insurers also compete with the premium discounts for people opting for a voluntary deductible on top of the mandatory deductible.[13] The number of consumers choosing a voluntary deductible has increased from about 3% in 2006 to 12% in 2015 (Vektis, 2016).

Since 2010 an increasing number of health insurers introduced lower-priced contracts with restricted provider networks and substantial co-payments for accessing outside network providers (Table 2.1). In 2015 about 7.5 percent of the population (1.25 million people) was enrolled in such a limited provider plan (NZa, 2015).

associations for elderly and patients (e.g. for diabetes and rheumatioid arthritis)(NZa, 2015). Group contracts with elderly and patient organizations are feasible because health insurers are compensated for predictably high expenditures by the risk adjustment system.

[13]Initially, in 2006 and 2007, there was a mandatory no-claim rebate of 255 euros per year. In 2008, this no-claim rebate was replaced by a mandatory deductible of 150 euros per year, which was gradually raised to 375 euros per year in 2015.

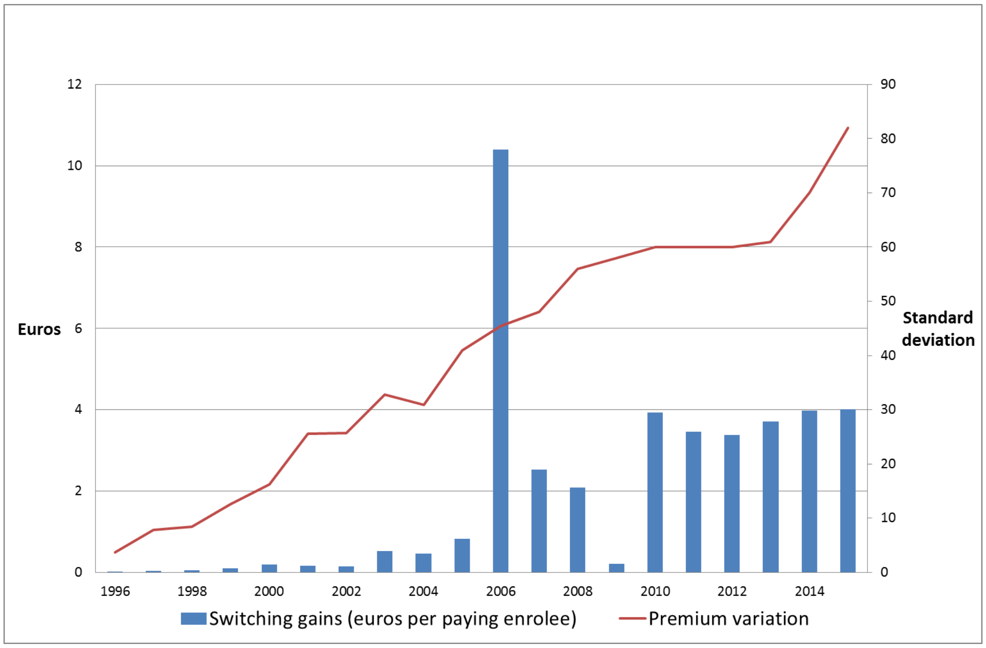


Figure 2.1: Average annual switching gains per premium payer and annual premium variation, 1996-2015. Switching gains per premium payer (see Table 2.2) are displayed on the left axis and premium variation (i.e. standard deviation annual premiums of group and individual contracts, see Table 2.2) on the right axis.

2.3 Financial switching gains for premium payers

Figure 1 exhibits switching gains and premium variation for the total 20-year period (1996-2015) making clear that switching gains for premium payers substantially increased due to the reform, with a peak in the reform year itself. We also observe an increasing trend in premium variation, corresponding with an increasing variety in health plan products after the reform year and the growing insurers' risk on medical expenses. In the next subsections we discuss these switching gains.

2.3.1 Switching gains in the SHI-market (1996-2005)

To examine whether consumers respond to premium differences across health plans we calculated total annual switching gains. To that end we compare the total average annual premiums (weighted by insurers' market shares) before and after switching (see Appendix 2.A for a more detailed explanation). In Table 2.2 we show that total switching gains increased from zero in 1995 to 6.7 million euro in 2005. Thus average total switching gains over the period are about 2 million euro per year. These switching gains are very modest. For example, in the year of the highest switching rate, 2005, average gains were 0.8 euro per premium payer and about 16 euro per switcher (i.e. 4% of the average annual premium). They are also very modest compared to potential total switching gains in the SHI market, which in any year could have been 100 to 200 times higher if all consumers would have switched that year to one of the cheapest health insurers (see Table 2.2).[14]

There are several potential explanations for the observed increase in switching gains. First, only since 1996 consumers could switch sickness funds once every year, and all sickness funds slowly started to compete on price and to attract customers from other sickness funds. Hence, it is likely that consumer awareness of switching opportunities has increased over time. Second, switching gains are likely to be larger when premium variation increases. Third, switching gains may also depend on institutional changes that affect insurers' price setting behaviour. For instance in 2003, several sickness funds had to raise their annual premiums because the government reduced the income-related contribution from 90 to 78% of total expenses. This change may have induced several sickness funds to adopt another pricing strategy. For example, large sickness funds were becoming relatively more expensive, which is reflected in Table 2.1 by the fact that for the first time weighted premiums before switching substantially exceeded the average premium. Notice, however, that in 2003 the weighted premiums after switching were also substantially higher than the average annual premiums, suggesting substantial consumer inertia since many enrolees apparently decided to stick with the relatively expensive large sickness funds.

2.3.2 Switching gains in the HIA-market (2006-2015)

Calculating switching gains in the HIA market is more complicated than in the SHI market because consumers do not only switch between health plans but also between individual and group contracts of these health plans (for a detailed explanation, see Appendix 2.A). The last two rows in Table 2.2 report the financial switching gains in this period. Switching gains were particularly high in the reform year 2006 with total switching gains of 130 million euro. The average gain per switcher remained fairly stable around 45 euro during 2006-2015 (see Table 2.2) but compared to 2006, the number of switchers were substantially

[14] Notice that (potential) switching gains and premium setting are interrelated. It is likely that an increase in the number of switchers would result in more premium competition, which would reduce premium varitation and potential switching gains.

lower after the reform year.[15] Still, with an average of 40 million per year during 2007-2016 total switching gains are quite modest, although much higher than prior to the reform. If in any year since 2007 all consumers would have switched to one of the cheapest health insurers total switching rates that year could have been about 10 times higher (see Table 2.2).

Nevertheless, consumers substantially benefited from switching since the introduction of the HIA. Table 2.3 shows a decomposition of the total switching gains into gains from switching within and between individual and group contracts. Initially most switching gains came from switching from individual to group contracts, but as of 2011 this changed and most gains came from switching within individual contracts. In 2015, we observe for the first time a reverse trend and that more consumers switch from a group contract to an individual contract. This is likely to be the result of the introduction of cheaper individual contracts for health plans with limited provider networks in recent years that are targeted at young people, which are much more inclined to switch (Duijmelinck and van de Ven, 2016).

	2008	2009	2010	2011	2012	2013	2014	2015
Total switching gains[a]	26.6	2.6	51.1	45.1	44.6	49.2	53.0	53.8
Within individual contracts	0.6	3.1	14.7	25.5	25.1	24.0	23.1	37.6
Within group contracts	1.3	3.2	2.2	0.9	9.2	11.6	22.2	24.2
Shift from individual to group contract	27.3	2.7	34.2	18.7	10.3	13.6	7.7	8.0

[a] Authors' own calculations, see Appendix 2.A.

Table 2.3: Decomposition of switching gains in HIA market, 2006-2015 (in millions euro)

2.4 Model and estimation methods

We estimate health plan price elasticities for three different periods: 1) prior to the reform 1995-2005, 2) the reform year 2006 and 3) the post-reform period 2007-2015[16].

[15]The low switching gains per premium payer in 2009 (see Table 2.2) are for a large part explained by one insurer who raised its premium on a large group contract to above the average premium in the market, while only a few consumers in this group contract switched to a lower priced contract (Vektis, 2016). In general, participating in a group contract is associated with a lower switching propensity (Boonen et al., 2016).

[16]We split the dataset in a pre- and post-reform period because of the different number of observations. In 2006, many new health plans were introduced, and insurers were also allowed to offer

For periods 1) and 3) we estimate an advanced dynamic model that follows from a standard discrete choice model, in which a consumer chooses an option out of all possible insurance policies in the market that maximizes his/her utility (Tamm et al., 2007; Train, 1986). The market share s_{it} of each insurance policy i in year t is represented by the multinomial logit equation:

$$s_{it} = \frac{exp(\beta p_{it} + \gamma_i + \varepsilon_{i,t})}{\sum_j exp(\beta p_{jt} + \gamma_j + \varepsilon_{jt}}, \tag{2.1}$$

where p_{it} denotes the community-rated annual premiums. The health plan fixed effect γ_i captures unobservable attributes that may differ across health plans, such as differences in the basic benefit package, health insurer quality, amount of advertising and the provision of supplementary insurance. Since data on these health plan attributes are not available we have to make the rather restrictive assumption that the impact of these attributes on market share do not change over time. We discuss the potential impact of this assumption on the estimation results in the Discussion Section. In addition, we assume that the stochastic term ε_{it} in the individual utility function is independent and has identically distributed extreme values (Train, 1986; McFadden, 1973). Taking logarithms and transforming this equation, we obtain:

$$log(s_{it}) = \beta p_{it} + \gamma_i + \delta_t + \varepsilon_{it}, \tag{2.2}$$

in which the term δ_t represents the denominator in Eq. (2.1). This model assumes that all consumers deliberately instantaneously choose a utility maximizing health insurance policy. Many researchers have already shown that this assumption does not hold for health insurance markets, which are characterized by a strong degree of persistence in health plan choice due to status quo bias, switching costs and information frictions (Samuelson and Zeckhauser, 1988; Strombom et al., 2002; Handel and Kolstad, 2015). To account for persistence in insurers' market shares we follow Tamm et al. (2007) and modify the equation by including a lagged market share term:

$$log(s_{it}) = \alpha log(s_{i,t-1}) + \beta p_{it} + \gamma_i + \delta_t + \varepsilon_{it}, \tag{2.3}$$

where $0 \leq \alpha \leq 1$ captures the average degree of persistence in the market. If $\alpha = 0$ the model is static and in that case the model in Eq. (2.3) is similar to

different types of a plan such as group contracts. As shown in Table 2.4 and 2.6, the number of observations, which we could use in our estimation, was 243 in the period 1996-2005, while it was 577 in the period 2007-2015. Merging the two dataset, would have resulted in a very unbalanced dataset. Furthermore, there was much public attention to the choice of a health plan and potential switching gains in the reform year, which probably influenced the switching behavior of consumers, and resulted in an extraordinary switching rate in 2006.

the instantaneous choice model in Eq. (2.2). If $0 < \alpha < 1$ there is some degree of persistence in the market that becomes larger when α is closer to one. From the specifications (2.1) and (2.3) we can derive the individual short-term and long-term premium elasticities, which we denote ϵ_{it} and τ_{it}, and subsequently annual average price elasticities ϵt and τ_{it} that we will report in this study.[17]

$$\begin{aligned} \epsilon_{it} &= \frac{\partial s_{it}}{\partial p_{it}}\frac{p_{it}}{s_{it}} = \beta p_{it}(1 - s_{it}), && (2.4) \\ \epsilon_t &\approx \beta \bar{p}_t \text{ in case } s_{it} \text{ is sufficiently small} \\ \tau_{it} &= \frac{1}{1-\alpha}\epsilon_{it}, \tau_t \approx \frac{1}{1-\alpha}\epsilon_t && (2.5) \end{aligned}$$

A property of the discrete choice model is that the elasticity in (2.5) is linearly related to the premium level p_{it}, implying that health plans face a convex demand curve with regard to the level of the annual premium prevailing in the market. All else equal, if there is a linear relationship between price and elasticity, with the same coefficient β, the price elasticity is about 3 times higher in a market with a premium level of about 1000 euro (after the reform) than about 350 euro (prior to the reform, since 2003).

In "Estimation results" we will estimate specification (2.3) with an OLS-estimation and subsequently with generalized methods of moment (GMM) estimation. It is well known that estimating the dynamic specification (2.3) with standard fixed or random effect models is complicated since the lagged term $log(s_{i,t-1})$ is likely to be correlated with the error term, the sum of γ_i and εit. Under the assumption of serially uncorrelated errors of ε_{it} we can use a GMM estimator to obtain consistent estimates (Arellano and Bond, 1991). Premiums in (2.3) may also be endogenous. For example, setting a lower premium to attract new consumers may be less profitable for a large insurer because its loss on the incumbent enrollees is predictably higher than for a smaller insurer. GMM controls for this possible endogeneity of p_{it} and $log(s_{i,t-1})$ by using lagged market shares and lagged premiums as instruments.

In the SHI market each insurer offers a single health plan with a similar benefit package. This is indicated in (1) by subscript i. However, in the HIA market, each insurer offers several health plans and often both individual and group contracts. We do have access to all individual contract prices in the market but for group contracts we have only information about the total number of enrollees of all group contracts per health plan and the corresponding (weighted) average

[17] For a more extensive discussion of the derivation and properties of the elasticitites, see Tamm et al. (2007) and Train (1986). Average annual elasticities are calculated by a weighted average of the insurer specific elasticities. In our study we have many health insurance policies (or health plans) allowing us to make the simplifying assumption $\epsilon \approx \beta \bar{p}_t$.

premium of these group contracts.[18] In the HIA market the subscript i in (2.1) therefore refers to all individual contracts and a group contract with a weighted premium per health plan.

Finally, due to the integration of the former SHI and private insurance schemes into the HIA, we performed a separate estimation of health plan price elasticities during the year of the reform (period 2). The separate dataset for this transition covers 2 years, before (2005) and after (2006) the introduction of the HIA.

2.5 Data

We obtained our data of health plan premiums and market shares from three different sources corresponding with the three periods, SHI, 1995-2005, the reform year 2006, and the HIA, 2007-2015. The first dataset was obtained from the Dutch National Health Care Institute (ZIN) and constitutes an unbalanced panel of 37 health plans (sickness funds) for 1995-2005 in the SHI (241 observations). In Appendix 2.B, Table 2.B.1, we describe all sickness funds in the market.[19]

The second dataset was constructed by the Dutch Healthcare Authority (NZa) including 30 health insurers that were active in the years just before (2005) and just after the reform (2006) (see Appendix 2.B, Table 2.B.2). For 2005 market shares in the voluntary private health insurance market were combined with market shares in the SHI market in order to construct a dataset that was comparable with HIA-data on market shares (of both individual contracts and group contracts) and premiums in 2006 (in total 54 observations).

The third source is an unbalanced panel dataset of 26-32 health insurers for 2007-2015 in the HIA (in total 694 observations) that was also obtained from the NZa. Since health insurers were allowed to offer various health plans we collected information on all "legally different" health plans, that is plans differing in terms of reimbursement method (in cash, in kind, or a combination of both) and contracted provider network. Next, we collected for each health insurer market shares for all individual contracts and an aggregated market share for all group contracts. Furthermore, we collected the corresponding annual premiums, and an average premium for all group contracts per insurer. For a description of the data, see Table 2.B.3 of Appendix 2.B. Since many group contracts are not accessible for the entire population, aggregating all group contracts and using an average group premium per insurer is a simplification that may bias our estimates.[20] However, aggregating all group contracts has the advantage that it

[18]This information is collected through insurer survays by the Dutch Healthcare Authority.

[19]The panel is unbalanced because of mergers. After a merger the merging insurers were removed from the dataset and a new merged insurer was added to our dataset in the year before the merger.

[20]Price elasticities are estimated under the assumption that consumers have free choice. Since not all group contracts are equally accessible this will bias our estimate for the price elasticity.

suits our discrete choice model better, since a very large part of the population has the option to choose at least one group contract at the average premium, which would certainly not be the case if we would consider each group contract separately in our estimations. This is because group contracts only differ in the price discounts offered by the insurer, and per insurer a group contract with an average discount rate is available to most individuals.[21]

2.6 Estimation results

2.6.1 Estimated health plan price elasticities for the SHI-market (1996-2005)

As we explained in the introduction, health plan price elasticities in the Dutch SHI market have been estimated before in several empirical papers. These studies found that estimated annual price elasticities were small and often below -0.5. However, these studies typically cover only part of the pre-reform period, use different estimation methods and did not include a lagged market share to control for persistence in health plan choice. Therefore the price elasticities reported by these studies can be seen as short-term elasticities. By contrast, our study covers the entire period, and we estimate a dynamic model taking into account choice persistence which allows us to estimate long term price elasticities as well. Especially in a longitudinal study over many years it is important to control for changing market shares because these dynamic effects are not captured by fixed insurer effects.

Table 2.4 summarizes both OLS and GMM estimates of health plan price elasticities for the entire SHI period. We included the OLS estimates for a better comparison with the results for the reform year in which we could not use GMM because of the small dataset. In this particular case, the results of both methods appear to be close to each other.

The OLS estimates correspond to short-term price elasticities ranging between -0.1 and -0.4 depending on the premium level. This range is consistent with the result of previous studies. For the GMM estimations we included time dummies, individual effects and reported robust standard errors. We report the GMM system estimator with endogenous premiums (Blundell and Bond, 1998), which we prefer for the following reasons. First, according to the economet-

[21] A distinction can be made between employer-based group contracts and other group contracts. Employer-based group contracts are typically only accessible for employees and dependents of the specific employers. Most other group contracts are open group contracts, meaning that they are acessible for all people joining the legal entity concluding the contract. In 2015 the average premium discount of employer-based group contracts relative to the same individual contract was 8.5%, whereas the average discount on the other group contracts was 6.4% (NZa, 2015).

$log(s_{it}) = \alpha log(s_{i,t-1} + \beta p_{it} + \gamma_i + \delta_t + \varepsilon_{it}$, $\bar{p}_t$ between 100 and 400 euros (see also Table 2.1)		
(i) OLS estimation, number of observations: 243		
$\hat{\alpha} = 0.86 *** (0.03)$	$\hat{\beta} = -0.0011$ (0.0006) $\hat{\varepsilon}_t \approx \hat{\beta}\bar{p}_t$ between -0.1 and -0.4 $\hat{\tau}_t$ between -0.8 and -3.1	R^2=0.97
(ii) System GMM estimation, number of observations used (including levels): 449		
$\hat{\alpha} = 0.91 *** (0.02)$	$\hat{\beta} = -0.0011 ***$ (0.0003) $\hat{\varepsilon}_t \approx \hat{\beta}\bar{p}_t$ between -0.1 and -0.4 $\hat{\tau}_t$ between -1.2 and -4.8	R^2=0.98

The estimations are performed with the plm-package in R (Croissant and Millo, 2008), total number of insurance policies used is 37 (because a merged policy is treated as a new ID). Estimation (ii) includes individual effects. Sargan test: 36.1 (D.f.=106, P-value=1), Wald test for coefficients (D.f.=2) has a P-value $<$ 0.2 e-16. statistics is not a part of standard GMM output. It is added for the sake of comparison with the first regression in this table, defined as $corr(s - \hat{s})^2$. Additional estimations results are available by the authors upon request.
Note *:P-value <0.05 ; **: P-value <0.01; ***: P-value <0.001.

Table 2.4: Estimation results for the health plan price elasticity in the SHI market 1996-2005

ric literature, the system GMM estimator has a better performance in terms of bias and efficiency than the first-difference GMM estimator. Second, premiums are likely to be endogenous both because market shares may be associated with market power and because large health plans may be less willing to reduce premiums (e.g. because of solvency regulations). Based on Sargan statistics (Sargan, 1958; Hansen, 1982), we cannot reject the hypothesis that the over identifying restrictions of the system GMM estimator are valid.

The short-term price elasticity resulting from the GMM estimator in estimation (ii) ranges from -0.1 (at a base premium of 100 euro) and -0.4 (at a base premium level of 400 euro). This estimate implies that a health insurer increasing its annual premium by 1% (about 1-4 euros) would cause an insurer's market share to decline by about 0.1-0.4%, depending on the premium level. The size of these price elasticities is similar to the OLS-estimates and those found in previous studies.

We found a high degree of persistence in the SHI market of around 90%, implying that most enrolees were sticking with a once chosen sickness fund. Strong persistence implies that long-term price elasticities are much higher than short-term price elasticities. According to our discrete choice model, Eq. 2.5, this strong persistence implies that long-term price elasticities range from -0.8 to -4.8%. This means that if an insurer would increase its premium by 1% each year, for an infinite number of years, then its market share would decline by 0.8-4.8%. The long-term elasticities are extremely sensitive to the precise estimation of the degree of persistence.

2.6.2 Estimated health plan price elasticities for the reform year 2006

For estimating health plan price elasticities in the reform year a specific dataset was constructed, comprising only two years (2005 and 2006). Given the small dataset we can only use OLS to estimate Eq. (2.3) without fixed effects and time dummies. Table 2.5 summarizes the estimations results. The results indicate a high degree of choice persistence of 84% in the market.

$log(s_{it}) = \alpha log(s_{i,t-1} + \beta p_{it} + \varepsilon_{it}$, $\bar{p}_t = 1025$ euros		
OLS estimation, reform year $t = 2006$, number of observations: 54		
$\hat{\alpha} = 0.84 *** $ (0.05)	$\hat{\beta} = -0.0055 ***$ (0.0021), $\hat{\epsilon}_t \approx \hat{\beta}\bar{p}_t$ =-5.7 (2.1)	R^2=0.84
	$\hat{\tau}_t = -35.6$	

We have fewer observations in our estimations than insurance policies in the data because new insurers entering the market in 2005 have a market share of zero and drop out of the sample.
Note *:P-value <0.05 ; **: P-value <0.01; ***: P-value <0.001.

Table 2.5: Estimation results for the health plan price elasticity in the reform year 2006

This result is in line with the observation that a large part of the population did not switch from health insurer. We found a high health plan price elasticity of about -5.7, which corresponds to the all-time high number of switchers of 18% in 2006. The estimated price elasticity implies that an average health insurer increasing its annual premium by 1% (about 10 euro) would, in 2006, experience a decline in market share of about 5.7%. Nevertheless, even in the reform year most people did not switch health plans, despite health plan annual premiums for all people changed dramatically relative to the preceding year. As a matter of fact, we still found a high degree of persistence of 84%, implying a corresponding extremely high long-term price elasticity of -35.6 for the reform year 2006.

A limitation of the estimated price elasticity is that because of the short period we could not include fixed effects to account for insurer specific characteristics (e.g. differences in supplementary insurance, service quality and rebates for voluntary deductibles). However, surveys among consumers indicate that, especially in 2006, price was the most important determinant of health insurer choice (NZa, 2006).

2.6.3 Estimated health plan price elasticities for the HIA market (2007-2015)

Table 2.6 summarizes the estimation results for the health plan price elasticity during the post-reform period. As for the SHI market we present the results of both the OLS estimation (including only time dummies) and GMM estimation (including time dummies and individual effects), and we report the GMM system estimator with endogenous premiums (Blundell and Bond, 1998).

$log(s_{it}) = \alpha log(s_{i,t-1} + \beta p_{it} + \gamma_i + \delta_t + \varepsilon_{it}$, $\bar{p}_t$ between 1100 and 1300 euros (see also Table 2.2)

(iv) OLS estimation, number of observations: 577

$\hat{\alpha} = 0.95 *** (0.01)$	$\hat{\beta} = -0.0017 ** (0.0008)$ $\hat{\varepsilon}_t \approx \hat{\beta}\bar{p}_t$ between -1.9 and -2.2	R^2=0.97
	$\hat{\tau}_t$ between -34 and -40	

(v) GMM estimation, number of observations used (including levels): 1013

$\hat{\alpha} = 0.81 *** (0.03)$	$\hat{\beta} = -0.0008 *** (0.0001)$ $\hat{\varepsilon}_t \approx \hat{\beta}\bar{p}_t$ between -0.9 and -1.0	R^2=0.99
	$\hat{\tau}_t$ between -4.7 and -5.5	

The estimations are performed with the plm-package in R (Croissant and Millo, 2008), total number of policies used is 155, making distinction between collective and individual policies and using only policies with a minimum of 10.000 enrolees. We use a system GMM estimator with endogenous premiums, including time dummies and fixed effects. Sargan statistics (85.83, D.f.=68, P value=0.071), Wald test for coefficients of this model (D.f.=2) has a P-value < 0.2 e^{-16}. Additional estimations results are available by the authors upon request.

Note *:P-value <0.05 ; **: P-value <0.01; ***: P-value <0.001.

Table 2.6: Estimation results for the health plan price elasticity in the HIA market 2007-2015

As shown in Table 2.6, estimated short-term price elasticities range between -0.9 and -2.2 depending on the estimation method. This is higher than in the pre-reform SHI-market, but substantially lower than in the reform year. Compared to the previous period, we find a larger discrepancy between the OLS and GMM estimates. This can arise due to both a wider variety and greater fluctuation in the number of health plans offered in the market in this period (as shown in Table 2.2), which increase premium endogeneity affecting OLS estimates. Yet, we report the OLS results as an upper bound, since robustness checks using alternative GMM specifications resulted in price elasticities higher than 0.8 in absolute value.[22]

[22] Not shown here, but available form the authors upon request.

The results in Table 2.6 show that consumer inertia in the HIA market is almost as high as in the SHI market, with a degree of persistence of 80-90%. The degree of switching persistence does not substantially differ between the three estimation periods, but is significantly different from one. Long-term price elasticities range from -5 to -40 and are again very sensitive to the estimation of the degree of persistence.

It is possible that our long-term price elasticities are overstated because the calculations assume that the short-term price elasticity remains constant over the years. However, it is likely that some consumers are more persistent in their choice of health plan than others, and that in practice we only observe a limited group of potential switchers that are price sensitive. Comparing our findings with Tamm et al. (2007) for the German social health insurance market, the only other study that measured the degree of persistence in the same way, we find lower long-term price elasticities. Tamm et al. (2007) cannot reject a degree of persistence of 100% ($\alpha = 1$) implying that long-term price elasticties are infinite, indicating that market shares of German health insurers could follow a random walk.

A conceptual problem with the interpretation of elasticities is that a high degree of persistence may indicate the presence of perfect competition or high consumer inertia. In the case of perfect competition there would be no consumer mobility since consumers would have chosen the optimal insurance product and insurers would fully adjust their prices to changes in marginal costs over the years. As shown in many studies, however, status quo bias and consumer inertia play a large role in health insurance markets (Samuelson and Zeckhauser, 1988; Handel, 2013). Our findings show that the Dutch health insurance market is no exception. From 2006 to 2014, 69% of the Dutch population never switched to another health insurer despite a growing premium variation (Vektis, 2015).[23]

2.6.4 Limitations

Although we estimated the health plan price elasticities using the same methodology between the SHI and HIA period is still complicated.

A limitation of our study is that by using fixed effects we can control for constant differences of unobserved market and institutional characteristics, but we cannot control for possible changes in these characteristics. A first market characteristic that may change over time and for which we lack sufficient data is supplementary insurance.[24] Since about 85% of the population buys supplemen-

[23]Individual level data from almost all health insurers show that 20% of the population switched once, 7% twice, 2% three times and 1% swithced four times between 2006 and 2015 (Vektis, 2015).

[24]There is some fragmented information available on supplementary benefits packages and premiums for a few recent years but consistent time series are lacking. Schut and Hassink (2002) tackled the

tary insurance and almost all from the same insurer from which they obtain basic health insurance, changes in supplementary insurance may have affected basic health plan choice. In contrast to supplementary insurance, basic health insurance offers are very transparent and easy to compare, because benefit packages are the same across all insurers, premiums are community-rated and cost-sharing arrangements are standardized (deductible levels are the same across all basic health plans). Therefore, most health insurers primarily use the basic health plan premiums in their marketing activities during the open enrolment period at the end of each year. Comparing prices and benefit packages of supplementary insurance is much more complicated for consumers because products and prices are difficult to compare and because of the vast number of supplementary insurance packages that are offered.[25] Since having supplementary insurance is found to be negatively related to people's propensity to switch basic health plans (Boonen et al., 2016), the increasing differentiation of supplementary insurance products may have resulted in downward-biased price elasticities of basic health plan choice in later years.

A second market characteristic that since 2006 is changing over the years is the proportion of people opting for a voluntary deductible for basic health insurance.[26] People of 18 years and older can choose a voluntary deductible of 100, 200, 300, 400 or 500 euros on top of the mandatory deductible in return for a premium discount. The number of people choosing for a voluntary deductible has increased from about 6% in 2006 to about 12% in 2015 (Vektis, 2016). If there is no strong correlation across insurers between health plan premiums with or without a voluntary deductible, our price elasticities are likely to be biased. However, we find that premiums of both types of health plans are highly correlated, and therefore this bias may not be substantial.[27]

A third changing characteristic of the HIA-market is the increasing role of limited-provider plans, making health plans more heterogeneous over time. As shown in Table 2.1 the share of the population enrolled in limited provider plans gradually increased from 0.1% in 2008 to 7.5% in 2015. Limited provider plans

problem of differentiated supplementary insurance products by using the price of a 'most common' supplementary benefits package and by re-estimating their equations with the total price (i.e. the sum of basic and supplementary insurance premiums). However, their study covered only the period 1996-1998 in the SHI market when product differentiation was limited.

[25] In 2015 people could choose among 276 supplementary health plans, of which 94 were specific dental health plans, resulting in more than 1300 possible choice combinations (NZa, 2015).

[26] Taking into account voluntary deductibles complicates the estimation precedure because only a few consumers opt for such a deductible.

[27] The Pearson correlation between the full premium and the discounted premium ranges from 0.81 to 0.98 in 2008-2015, whereas the Spearman rand correlation ranges from 0.63 to 0.88 over the same period. In 2007, both correlations were somewhat lower (about 0.5) but in that year the share of the enrollees choosing a voluntary deductible was extremely low.

are typically lower priced than health plans with unrestricted provider choice. If limited provider plans are offering lower (perceived) quality than unrestricted health plans, our price elasticities may be biased downwards because it is likely that more people would have chosen lower-priced plans if quality would have been the same.

A comparison of price elasticities between SHI and HIA is also complicated by the changes in institutional characteristics of the choice setting. An important difference between the SHI and HIA is the way in which premium subsidies are structured. As a result premium levels in the HIA-market are on average 3 to 6 times higher than in the SHI-market. Adjusting for the different premium levels would reduce the difference between the estimated price elasticities between both markets, but it is not clear to what extent because the impact of the premium level on price elasticity is difficult to assess. For instance, different premium levels may induce a different behavioural response, since consumers may not only respond to *absolute* but also to *relative* premium differences (Tversky and Kahneman, 1981). All other things equal, this would result in higher price elasticities in the SHI-market than in the HIA-market.[28] Price elasticities may also be somewhat higher in SHI than HIA because the SHI did not cover high income people, who may be less sensitive to price than lower income people (because of a diminishing marginal utility of money).

Another difference that may complicate a good comparison is the much higher number of choice options in the HIA-market than in the SHI-market. This may have had a downward effect on the price elasticity in the HIA-market because of information overload (Frank and Lamiraud, 2009).

Since it is not possible to disentangle all these possible effects, we cannot determine to what extent the increase in estimated health plan price elasticities was driven by the reform. For this reason the quantitative changes in health plan price elasticities we estimated should be interpreted only as a rough indication of the impact of 20 years of managed competition reforms in the Netherlands.

2.7 Conclusion

In 1996 managed competition was introduced in the Dutch social health insurance (SHI) scheme. From 1996 to 2005 health insurers had few tools and limited incentives to compete, and consumers had little incentives to switch. In 2006 a major reform was implemented to provide insurers with more incentives and tools to compete and to provide consumers with a more differentiated health plan choice. Using data on prices and market shares of all health plans over a 20 years period (1995-2015) we provide a long-term overview with respect to

[28] This is an interesting topic for future research.

the number of switchers, switching gains and health plan price elasticities in the Dutch insurance market. The Dutch setting is especially interesting because it describes the first and subsequent steps of introducing managed competition into a social health insurance market. This information is not only useful for Dutch policymakers but also for other countries following a similar path.

Prior to the reform (1995-2005) we find modest increasing total switching gains increasing from about 0 in 1995 to 7 million euro per year in 2005. The reasons are small premium variations and a low number of switchers, between 2 to 4% a year. If all consumers would have switched to one of lowest priced health plans, switching gains could have been 100 to 200 times larger in any year (of course this holds only for a single year and not for the entire period). We find modest short-term health plan price elasticities ranging from -0.1 to -0.4, depending on the annual premium level.

The introduction of the reforms (2006) resulted in an all-time high switching rate of 18% and a health price elasticity of -5.7. Moreover, switching gains for consumers peaked with total gains of 130 million in the first year of the reform (2006). The main reason is that the reform had a large impact on consumer awareness of switching possibilities.

In the post-reform decade (2007-2015), the number of switchers returned to lower levels with the number of switchers increasing from about 4 to 8%. Consumers financially benefited much more from switching health plans than in the SHI (on average about 45 euro per switcher per year since 2006), although total switching gains in any year still could have been about 10 times higher if all people would have switched to one of the lowest priced health plans. We find health plan price elasticities that range from -0.9 to -2.2.

A good comparison of short- and long-term health plan price elasticities between the SHI and HIA period is complicated because in our estimations we may not perfectly control for unobserved changing market and institutional characteristics. We do find strong evidence of substantial consumer inertia as the degree of switching persistence varied from about 0.8 to 0.9 during the 20-year period. Strong persistence also implies that long-term price elasticities could be much higher than short-term price elasticities, because people only slowly respond to changing prices.

The high forgone potential switching gains and high level of persistence suggest that many people make suboptimal choices, particularly because quality differences between health plans appear to be small (NZa, 2015). Therefore, an active policy to improve health plan choice may be welfare enhancing in this case. One option is to increase transparency in the insurance market by facilitating better informed consumer choices.[29] Although comparative health plan infor-

[29] As shown by Handel (2013), the welfare effects of improving health plan choice are theoretically ambiguous because the positive welfare effects may be offset by increasing adverse selection. In the

mation is readily available on the internet, this information is often incomplete (e.g. lacking information on available group contracts) and somtimes biased by commercial interests (e.g. brokers' fee are paid when the consumer enrolls into a health plan via a comparison website). It is important to ensure that choice sites offer independent, complete and compehensive information on health plans (and if possible also on group contracts). Also, consumer education campaigns on how to choose a suitable health plan and how to recognize a good quality choice sit eis a way of improving consumer choice. Another option for lowering consumer search costs is to improve the choice structure for the type of health plans offered by insurers (perhaps 'bronze', 'silver', 'gold' and 'platinum' health plans, which would be easy to distinguish for consumers). In the Netherlands, some steps have been taken in this direction by requiring insurers to publish prices of "similar" health plans that are sold under different labels or through different channels. It is expected that this will make it easier for consumers to select the cheapest plan. Similar rules could also be imposed with respect to information on premium discounts on group contracts, many of which are open to all consumers. Lastly, insurers could be obliged to inform consumers actively, regarding changes in the contracted provider network relevant to the consumers' residential area.

Dutch case, however, adverse selection is effectively constrained by sophisticated risk adjustment. Therefore, it is likely that in the Dutch setting improving individual-level plan choies will enhance welfare.

Appendix

2.A Calculating switching gains

In this appendix we explain how we calculated switching gains. In the SHI market, 1996-2005, this is straightforward and switching gains SG_t in year t are defined as:

$SG_t = N_t \sum_j P_{j,t}(s_{j,t-1} - s_{j,t})$

where $P_{j,t}$ is the premium of insurance policy of insurer j in year t, $s_{j,t}$ is the market share of insurer j in year t, and N_t is the total number of premium payers in year t.

In the HIA market it is more complicated because we have individual and group contracts. Most of these policies were offered both as individual contracts and as group contracts. Our dataset is complete for all years, except for 2006 and 2007.[30] When an insurance firm withdraws some policy from the market, it usually reallocates the enrolees to the closest alternative policy in its portfolio, and this alternative policy becomes in our calculations the default option of these enrolees for the next year.[31] In the computation of switching gains we make the following two assumptions. First, if an insurer offers a policy in year $(t-1)$ but does not offer the same policy in year t then we assume that in year t these enrolees would be offered the closest available policy of the same insurer as a 'default option'. We use this assumption for individual and group contracts. Next, we allocate the individual and group market shares of policies that exited the market in year $(t-1)$ to the respective default options. Secondly, if a new policy enters the market in year t (which is not a default option of a policy leaving the market) then we assume the market share of this policy in year $(t-1)$ was zero. The total switching gains represent the weighted average price change that arises because of reallocation of enrolees among contracts, multiplied by the number of premium payers N_t:

[30] A few missing values were imputed using the information from neighbouring years.

[31] It would be incorrect to see the enrollees who remain on this policy as 'switchers'.

$$SG_t = N_t \sum_j P^i_{j,t}(s^i_{j,t-1} - s^i_{j,t}) + P^c_{j,t}(s^c_{j,t-1} - s^c_{j,t}),$$

where $P^i_{j,t}$ is the premium of all individual insurance policies of insurer j in year t, $P^c_{j,t}$ is the premium of all group (or collective) insurance policies of insurer j in year t, $s^i_{j,t}$ is the market share of all individual insurance policies of insurer j in year t and $s^c_{j,t}$ is the market share of all group insurance policies of insurer j in year t.

These gains can be attributed to three sources:

1. Switching gains within the individual insurance segment: SG^i_t
2. Switching gains within the group insurance segment: SG^c_t
3. Switching gains from individual to group (or versa): SG^{ic}_t

We can calculate these three types of switching gains. First we define

$$\begin{aligned}
\Delta P^i_t &= \sum_j P^i_{j,t}(s^i_{j,t-1}/s^i_{t-1} - s^i_{j,t}/s^i_t) \qquad (2.A.1)\\
\Delta P^c_t &= \sum_j P^c_{j,t}(s^c_{j,t-1}/s^c_{t-1} - s^c_{j,t}/s^c_t)\\
\Delta P^{ic}_t &= \sum_j P^i_{j,t}s^i_{j,t-1}/s^i_{t-1} - P^c_{j,t}(s^c_{j,t-1}/s^c_{t-1}
\end{aligned}$$

where s^i_t and s^c_t are market shares of the respective segment at time t. Note that the market shares sum up to one at any time t: $\sum_j s^i_{j,t} + s^c_{j,t} = s^i_t + s^c_t = 1$, and $P^i_{j,t}$ and $P^c_{j,t}$ denote individual and collective premiums of policy j at time t. Using these notations, the total switching gains are decomposed into the three sources as follows:

$$SG_t = SG^i_t + SG^c_t + SG^{ic}_t = \Delta P^i_t s^i_{t-1}N_{t-1} + \Delta P^c_t s^c_{t-1}N_{t-1} + \Delta P^{ic}_t(s^c_t - s^c_{t-1})N_{t-1}$$

2.B Tables on insurers

	Health insurer	Operating years	No. of Annual obs.	Additional information[b]
1	AGIS	2001-2005	5	New large insurer, merger of Anova, ZAO, ANOZ
2	Amicon	1995-2005	11	Large insurer
3	Anderzorg	1995-2005	11	Small insurer
4	Anova	1995-2001	7	Large regional insurer, merged in 2002 into AGIS
5	ANOZ	1996-2001	6	Large regional insurer, merged in 2002 into AGIS
6	Azivo	1995-2005	11	Medium sized regional insurer.
7	CZ Groep	1995-2005	11	Large insurer
8	De Friesland	1995-2005	11	Medium sized regional insurer
9	DSW	1995-2005	11	Medium sized regional insurer.
10	Geove	1995-2005	11	Medium sized insurer
11	Groene Land	1995-2005	11	Large insurer
12	Nederzorg	1998-2005	8	New small insurer
13	Nuts	1995-2005	11	Medium sized insurer
14	NZC	1997-1999	3	New small insurer, left market in 2000.
15	OHRA	1995-1999	5	Small insurer, merged with Nuts in 2000
16	ONVZ	1997-2005	9	New small insurer
17	OZ	1996-2005	11	Large insurer.
18	OZB	1998-2005	8	New small insurer
19	Pro Life	1996-2000	4	New small insurer, merged in 2001 with ANOVA
20	PWZ	1995-2001	7	Medium sized insurer, merged in 2002 with Groene Land
21	Salland	1995-2005	11	Small insurer
22	SR Rotterdam	1995-2005	11	Small regional insurer
23	Topzorg	1995-1999	4	Small regional insurer, merged with Geove in 2000
24	Trias	1995-2005	11	Medium sized regional insurer
25	Univé	1995-2005	11	Large insurer
26	VGZ	1995-2005	11	Large insurer
27	ZAO	1995-2001	7	Large regional insurer, merged in 2002 into AGIS
28	ZK	1995-2005	11	Large insurer
29	ZK Noordwijk	1995-1997	3	Medium sized regional insurer, merged in 1998 with ZK
30	ZK Spaarneland	1995-1997	3	Medium sized regional insurer, merged in 1998 with ZK
31	ZON	1995-1999	5	Medium sized insurer, merged in 2000 with Amicon
32	Zorg & Zekerheid	1995-2005	11	Medium sized regional insurer

[a] Our observation series are unbroken and cover 100% of the market. In total we obtained 270 observations. Note that in our estimations we need at least three consecutive years of data to perform GMM estimations. An insurer is denoted “small” if in the last year of the sample the market share was less than 1%, “medium sized” if the market share was between 1 and 5%, and “large” if the market share was larger than 5%. We also indicated whether an insurer operated mainly regional.

Table 2.B.1: Health insurers that were active in the SHI market during 1995- 2005

	Health insurer	Operating years	No. of Annual obs.[a]	Additional information[b]
1	AGIS	2005-2006	2	Large insurer, operated in 2005 on SHI and PHI market
2	Anderzorg	2005-2006	2	Small Insurer, operated in 2005 on SHI and PHI market
3	Avéro	2005-2006	2	Medium sized insurer, operated in 2005 on PHI market
4	Azivo	2005-2006	2	Small insurer, operated in 2005 on SHI market
5	AZVZ	2005-2006	2	Small insurer, operated in 2005 on PHI market
6	Confior	2005-2006	2	Small insurer, operated in 2005 on PHI market
7	CZ-Group	2005-2006	2	Large insurer, operated in 2005 on SHI and PHI market
8	De Friesland	2005-2006	2	Medium sized insurer,operated in 2005 on SHI and PHI market
9	De Goudse	2005-2006	2	Small insurer, operated in 2005 on PHI market
10	Delta Loyd	2005-2006	2	Medium sized insurer, operated in 2005 on PHI market
11	DSW	2005-2006	2	Medium sized insurer, operated in 2005 on SHI market
12	FBTO	2005-2006	2	Medium sized insurer, operated in 2005 on PHI market
13	FORTIS	2005-2006	2	Medium sized insurer, operated in 2005 on PHI market
14	Groene Land	2005-2006	2	Large insurer, operated in 2005 on SHI market
15	Interpolis	2005-2006	2	Small insurer, operated in 2005 on PHI market
16	IZA	2005-2006	2	Medium sized insurer, operated in 2005 on PHI market
17	IZZ	2005-2006	2	Medium sized insurer, operated in 2005 on PHI market
18	Menzis	2005-2006	2	Large insurer, operated in 2005 on SHI and PHI market
19	OHRA	2005-2006	2	Medium sized insurer, operated in 2005 on PHI market
20	ONVZ	2005-2006	2	Medium sized insurer, operated in 2005 on SHI market
21	OZ	2005-2006	2	Medium sized insurer, operated in 2005 on SHI market
22	OZB	2005-2006	2	Small insurer, operated in 2005 on SHI and PHI market
23	PNO	2005-2006	2	Small insurer, operated in 2005 on PHI market
24	Salland	2005-2006	2	Small insurer, operated in 2005 on SHI market
25	SR	2005-2006	2	Small insurer, operated in 2005 on SHI market
26	Trias	2005-2006	2	Medium sized insurer, operated in 2005 on SHI and PHI market
27	UMC	2005-2006	2	Small insurer, operated in 2005 on PHI market
28	Univé	2005-2006	2	Medium sized insurer, operated in 2005 on SHI and PHI market
29	VGZ	2005-2006	2	Large insurer, operated in 2005 on SHI and PHI market
30	Zorg & Zekerheid	2005-2006	2	Medium sized insurer, operated in 2005 on SHI and PHI market
31	ZK	2005-2006	2	Large insurer, operated in 2005 on SHI and PHI market

[a] Prior to 2006 some health insurers were only active on the social health insurance (SHI) market, some only on the private health insurance market (PHI) and some on both markets. Since 2006 all insurers are active in the same market, market shares of all insurers had to be collected from both markets in 2005. For each health insurer we obtained market shares of all individual contracts and (the sum) of all group contracts in 2005. For 2006 we obtained individual and group market shares and corresponding nominal premiums, where the premium for the group contracts for each insurer is calculated by taking the average (with market share) weighted premiums of all individual group contracts.
[b] We denoted whether an insurer was active in 2005 as a former sickness fund on the SHI-market and/or as a private indemnity insurer on the PHI market. An insurer is denoted "small" if in the year 2006 insurers' market share was less than 1%, "medium sized" if the market share was between 1 and 5% , and "large" if the market share was larger than 5%.

Table 2.B.2: Health insurers that were active during the reform years 2005 - 2006

	Health insurer	Holding 2016 c	Operating years	No. of Annual obs.	No. of Policies[a]	Additional information[b]
1	AGIS	Achmea	2007-2014	8	2 - 5	large insurer, since 2008 part of holding Achmea
2	Anderzorg	Menzis	2007-2015	8	1	medium sized insurer
3	ASR	ASR	2007-2015	8	2 - 4	medium sized insurer
4	Avéro	Achmea	2007-2011	8	2 - 4	large insurer
5	Azivo	Menzis	2007-2015	8	1	medium sized insurer, since 2008 part of holding Menzis
6	Confior	Menzis	2007-2008	2	2	small insurer, merged in 2009 into Menzis
7	De Friesland	Achmea	2007-2015	8	2 - 5	large insurer, since 2012 part of holding Achmea
8	Delta Lloyd	CZ	2007-2015	8	1 - 2	medium sized insurer, since 2008 part of holding CZ
9	Eno	Eno	2007-2015	8	1 - 2	medium sized insurer
10	FBTO	Achmea	2007-2015	8	1 - 2	medium sized insurer
11	Groene land	Achmea	2007-2009	3	1 - 2	large insurer, merged in 2010 into Zilveren Kruis
12	Interpolis	Achmea	2007-2015	8	1	medium sized insurer
13	IZA	VGZ	2007-2015	8	1 - 3	large insurer
14	IZZ	VGZ	2007-2015	8	1 - 2	medium sized insurer
15	Menzis	Menzis	2007-2015	8	2 - 4	large insurer
16	Univé	VGZ	2007-2015	8	3 - 5	large insurer
17	Cares	VGZ	2007-2015	8	2 - 3	small insurer
18	UMC	VGZ	2007-2015	8	1	medium sized insurer
19	OHRA	CZ	2007-2015	8	1 - 3	medium sized insurer, since 2008 part of holding CZ
20	OHRA Zorg	CZ	2007-2015	8	1 - 4	medium sized insurer, since 2008 part of holding CZ
21	ONVZ	ONVZ	2007-2015	8	1	medium sized insurer
22	AZVZ	Z&Z	2007-2010	4	1	small insurer, exit in 2011, taken over by holding Z&Z
23	CZ	CZ	2007-2015	8	2 - 4	large insurer
24	DSW	DSW-SH	2007-2015	8	1	medium sized insurer
25	Z&Z	Z&Z	2007-2015	8	2 - 3	medium sized insurer
26	OZF	Achmea	2007-2015	8	1	small insurer
27	PNO	ONVZ	2007-2009	3	1	small insurer, exit in 2010, taken over by holding ONVZ
28	Stad Holland	DSW-SH	2007-2015	8	1	small insurer
29	Trias	VGZ	2007-2011	5	2	medium sized insurer,since 2012 part of holding VGZ
30	Univé Zorg	VGZ	2007-2008	2	2	medium sized insurer
31	VGZ	VGZ	2007-2015	8	1 - 4	large insurer
32	Zilveren Kruis	VGZ	2007-2015	8	2 - 11	large insurer

[a] The Table reports the number of different health plans existing within one year. Number of different health plans offered by the same insurer may fluctuate over the years. These health plans could be sold either via an individual contracts or group contracts or both. In the period studied, most health plans were sold via both individual and group contracts.
[b] An insurer is denoted "small" if insurers' market share was smaller than 100 000, "medium sized" if the market share was between 100 000 and 500 000, and "large" if the market share was larger than 500 000 enrolees.
[c] Many health insurers are operating within a larger holding company as separate legal entities. Often the name of the holding company is the same as the name of the largest health insurer within the holding.

Table 2.B.3: Health insurers that operated in the HIA market during 2007 - 2015

Bibliography

Arellano, M. and Bond, S. Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *The review of economic studies*, 58(2):277–297, 1991.

Blundell, R. and Bond, S. Initial conditions and moment restrictions in dynamic panel data models. *Journal of econometrics*, 87(1):115–143, 1998.

Boonen, L. H., Laske-Aldershof, T., and Schut, F. T. Switching health insurers: the role of price, quality and consumer information search. *The European Journal of Health Economics*, 17(3):339–353, 2016.

Croissant, Y. and Millo, G. Panel data econometrics in r: the plm package. *Journal of Statistical Software*, 27(2):1–43, 2008.

Douven, R. and Schut, E. Premieconcurrentie tussen zorgverzekeraars. *Economisch Statistische Berichten ESB*, (91):272–275, 2006.

Douven, R., Ligthart, M., Lieverdink, H., and Vermeulen, I. Measuring annual price elasticities in dutch health insurance; a new method. Technical report, CPB Netherlands Bureau for Economic Policy Analysis, 2007.

Douven, R., Katona, K., Schut, F. T., and Shestalova, V. Switching gains and health plan price elasticities: 20 years of managed competition reforms in the netherlands. *The European Journal of Health Economics*, pages 1–18, 2017.

Douven, R. C. and Schut, F. T. Pricing behaviour of nonprofit insurers in a weakly competitive social health insurance market. *Journal of health economics*, 30(2):439–449, 2011.

Duijmelinck, D. M. and van de Ven, W. P. Choice of insurer for basic health insurance restricted by supplementary insurance. *The European Journal of Health Economics*, 15(7):737–746, 2014.

Duijmelinck, D. M. and van de Ven, W. P. Switching rates in health insurance markets decrease with age: empirical evidence and policy implications from the netherlands. *Health Economics, Policy and Law*, 11(2):141–159, 2016.

Frank, R. G. and Lamiraud, K. Choice, price competition and complexity in markets for health insurance. *Journal of Economic Behavior & Organization*, 71(2):550–562, 2009.

Handel, B. R. Adverse selection and inertia in health insurance markets: When nudging hurts. *American Economic Review*, 103(7):2643–82, 2013.

Handel, B. R. and Kolstad, J. T. Health insurance for" humans": Information frictions, plan choice, and consumer welfare. *American Economic Review*, 105 (8):2449–2500, 2015.

Hansen, L. P. Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, pages 1029–1054, 1982.

Laske-Aldershof, T. and Schut, F. Monitor verzekerdenmobiliteit. onderzoek in opdracht van het ministerie van vws. *Rotterdam: iBMG, Erasmus Universiteit*, 2005.

McFadden, D. Conditional logit analysis of qualitative choice behavior. In Zarembka, P., editor, *Frontiers in econometrics*, pages 105 – 142. Academic Press, New York, 1973.

NZa. Monitor zorgverzekeringsmarkt juni 2006. Technical report, Dutch Healthcare Authority, 2006.

NZa. Monitor zorgverzekeringsmarkt. de balans 2007. Technical report, Dutch Healthcare Authority, 2007.

NZa. Monitor zorgverzekeringsmarkt 2009. Technical report, Dutch Healthcare Authority, 2009.

NZa. Marktscan zorgverzekeringsmarkt. weergave van de markt 2007–2011. Technical report, Dutch Healthcare Authority, 2011.

NZa. Marktscan zorgverzekeringsmarkt. weergave van de markt 2008–2012. Technical report, Dutch Healthcare Authority, 2012.

NZa. Marktscan en beleidsbrief zorgverzekeringsmarkt. weergave van de markt 2010–2014. Technical report, Dutch Healthcare Authority, 2014.

NZa. Marktscan van de zorgverzekeringsmarkt 2015. Technical report, Dutch Healthcare Authority, 2015.

Pendzialek, J. B., Simic, D., and Stock, S. Differences in price elasticities of demand for health insurance: a systematic review. *The European Journal of Health Economics*, 17(1):5–21, 2016.

Roos, A.-F. and Schut, F. T. Spillover effects of supplementary on basic health insurance: evidence from the netherlands. *The European Journal of Health Economics*, 13(1):51–62, 2012.

Samuelson, W. and Zeckhauser, R. Status quo bias in decision making. *Journal of risk and uncertainty*, 1(1):7–59, 1988.

Sargan, J. D. The estimation of economic relationships using instrumental variables. *Econometrica: Journal of the Econometric Society*, pages 393–415, 1958.

Schut, F. T. and Hassink, W. H. Managed competition and consumer price sensitivity in social health insurance. *Journal of Health Economics*, 21(6): 1009–1029, 2002.

Schut, F. T., Gress, S., and Wasem, J. Consumer price sensitivity and social health insurer choice in germany and the netherlands. *International Journal of Health Care Finance and Economics*, 3(2):117–138, 2003.

Strombom, B. A., Buchmueller, T. C., and Feldstein, P. J. Switching costs, price sensitivity and health plan choice. *Journal of Health economics*, 21(1):89–116, 2002.

Tamm, M., Tauchmann, H., Wasem, J., and Greß, S. Elasticities of market shares and social health insurance choice in germany: a dynamic panel data approach. *Health Economics*, 16(3):243–256, 2007.

Tapay, N. and Colombo, F. Private health insurance in the netherlands: A case study. Technical report, OECD Publishing, 2004.

Train, K. *Qualitative choice analysis: Theory, econometrics, and an application to automobile demand*, volume 10. MIT press, 1986.

Tversky, A. and Kahneman, D. The framing of decisions and the psychology of choice. *science*, 211(4481):453–458, 1981.

Van de Ven, W. P. and Schut, F. T. Universal mandatory health insurance in the netherlands: a model for the united states? *Health Affairs*, 27(3):771–781, 2008.

van Dijk, M., Pomp, M., Douven, R., Laske-Aldershof, T., Schut, E., de Boer, W., and de Boo, A. Consumer price sensitivity in dutch health insurance. *International Journal of Health Care Finance and Economics*, 8(4):225–244, 2008.

Van Kleef, R. C., Van Vliet, R. C., and Van de Ven, W. P. Risk equalization in the netherlands: an empirical evaluation. *Expert review of pharmacoeconomics & outcomes research*, 13(6):829–839, 2013.

Vektis. Verzekerden in beeld 2015, zorgthermometer. Technical report, Vektis, 2015.

Vektis. Verzekerden in beeld 2016, zorgthermometer. Technical report, Vektis, 2016.

Chapter 3

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

Ramsis Croes, Katalin Katona, Misja Mikkers, Victoria Shestalova

This chapter is based on:
Croes, R., Katona, K., Mikkers, M., and Shestalova, V. Evidence of selection in a mandatory health insurance market with risk adjustment. Discussion paper 2018-013, TILEC, Tilburg University, 2018

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. One 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.

JEL Classification Numbers: D82, G22, I18

Keywords: risk-adjustment,selection, adverse selection, favorable selection, deductibles.

3.1 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 insurance[1].

Regulation of the health insurance market serves also the public policy goals of equity and efficiency. In a competitive market, insurers price their product according to the risk of the enrollee. In case of community rating, which prevents insurers from risk-rating, insurers may try to rely on selection to obtain homogenous risk groups of enrollees. This may mean trying to attract only healthy enrollees but also to separate the enrollees into different health plan options designed for the different risk classes of enrollees (i.e. second degree price discrimination). Both types of selection, however, decrease solidarity.

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 this way, the risk-adjustment system helps to achieve solidarity because it removes the incentives for selection or price discrimination. Although, a recent theoretical insight from Bijlsma et al. (2014) suggests that selection could exists even in a market with optimal risk-adjustment, the selection is then based on other consumer characteristics than healthcare costs (i.e. risk)[2].

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 et al., 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 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.

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

[2]Bijlsma et al. (2014) shows that 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. The sorting is then based on price sensitivity, rather then on healthcare costs, because insurers cannot profit from the lower costs of healthy enrollees due to the risk-adjustment system. The paid premium is correlated with the health of the enrollee but the price difference is justified by the demand elasticities rather than differences in healthcare costs.

adverse selection) and (*ii*) the incentives for health insurers to select low-risk individuals by tailoring insurance options to attract them (Newhouse et al., 2015, 2012; Einav and Finkelstein, 2011; Rothschild and Stiglitz, 1976).

In this paper, we study whether the possibility to opt for a voluntary deductible induces selection of enrollees in the Dutch insurance market by separating them in different health plan options. We also control for heterogeneity in risk preference among the enrollees so that we can focus on the effects of private information on expected healthcare costs. Because of the detailed risk-adjustment system, one might expect that this selection effect is low. However, van Kleef et al. (2008b) shows on Swiss data and additional simulations that the risk-adjustment system (as in 2006) does not fully compensate for the adverse selection effect of voluntary deductibles. We will confirm this finding in the more refined risk-adjustment system of 2013 and by using actual data.

Empirical identification of selection in actual data is challenging because of the interaction of selection with moral hazard (Cohen and Siegelman, 2010). Using a unique dataset on the Dutch health insurance market, we can take account of the past health costs and we are able to identify the selection effect separately from moral hazard effect. First, we assess the difference in healthcare costs between enrollees with and without voluntary deductible. We include potential confounding factors step-by-step in the analysis to control for heterogeneity in risk aversion among enrollees. Second, we show that the current risk-adjustment system cannot eliminate the whole cost difference. We repeat the estimations of the first step but we use the risk-adjusted result as dependent variable. The risk-adjusted result is the difference between the compensation from the risk-adjustment fund and the actual health care cost, which can be interpreted as the actual healthcare costs that an insurer faces. Also here, we see differences between enrollees with and without voluntary deductible, which suggests that there are incentives for the insurers to separate these groups of enrollees.

This paper contributes to the existing literature by showing empirically that the possibility for consumers to opt for voluntary deductible can results in selection, even in a health insurance market with a risk-adjustment system in place. 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 3.2 explain the organization of the Dutch health insurance sector in section 3.3. This is followed by the exposition of our empirical approach in section 3.4, after which we describe the data and our empirical findings in sections 3.5 and 3.6. Section 4.5 ends with concluding remarks.

3.2 Literature

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

3.2.1 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 et al., 2013). Most studies attribute this cost reduction to lower health care utilization (Keeler, 1992; Gerfin and 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 cost[3].

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 et al., 2008a; Tollen et al., 2004).

More specific to our context, van Winssen et al. (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.

[3] Note that enrollees who have to pay deductibles (or a co-insurance rate) may reduce or postpone necessary health care treatment (Brot-Goldberg et al., 2015; Fronstin and Collins, 2008; Lohr et al., 1986; Davis et al., 2005; Galbraith et al., 2011).

3.2.2 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 et al. (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 et al. (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).

3.3 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 plan[4]. Insurers can offer a voluntary deductible to its enrollees. Deductible options of €100,

[4] 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.

€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.

3.3.1 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, based on the past health care consumption of the individual.[5].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

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

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 deductible[6].

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.

3.3.2 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

[6] 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.

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 years[7]. Figure 3.1 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.

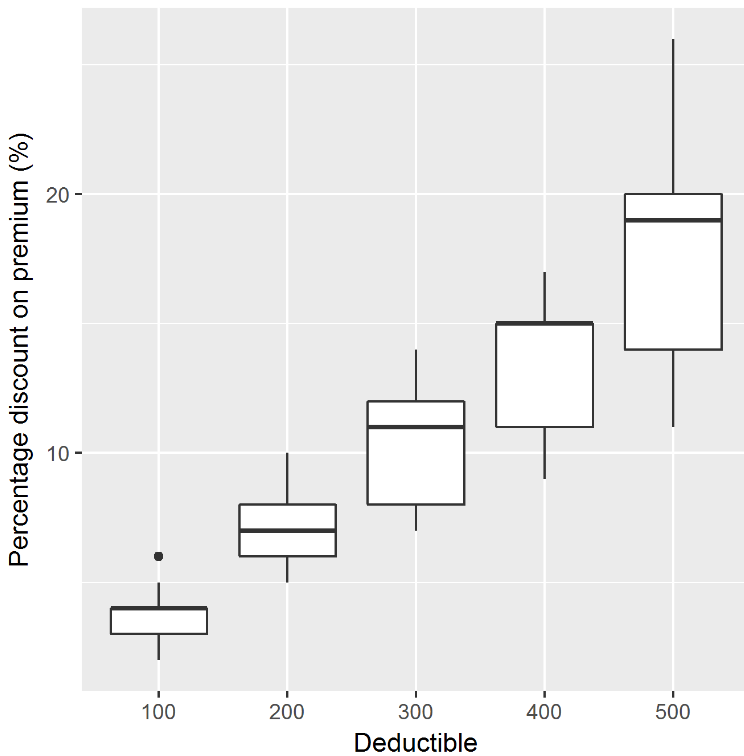


Figure 3.1: Boxplot of discounts offered at different deductible levels in 2013
Insurers can offer for each deductible level a discount on the health plan's premium. Insurers may not differentiate the discount among enrollees of the same health plan. In this graph, for each deductible level the average discount is calculated by taking the average of the health plans' discounts.

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

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

in 2013 is shown in Figure 3.2. 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%.

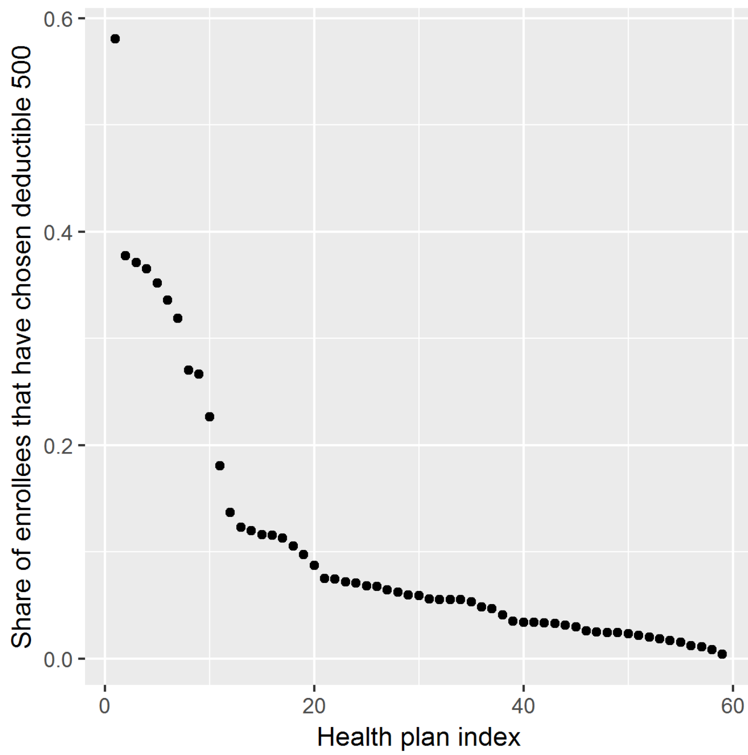


Figure 3.2: Share of enrollees with deductible level €500 in 2013
For each health plan, we calculated the share of enrollees that have chosen the maximum deductible level in 2013. We excluded health plans that have less than 1000 enrollees

3.4 Empirical strategy

In this paper, we are interested to determine (*i*) how strong the relationship is between healthcare expenditure and the choice for deductibles (i.e. the degree of adverse selection) and (*ii*) to what extent is this effect mitigated by the risk-adjustment system, i.e. what are the cost differences between enrollees with and without a voluntary deductible that an insurer faces. The latter one is important because the cost differences between enrollees give an incentive for insurers to attract low-risk enrollees.

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 et al. (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 effect that is mitigated by the risk-adjustment (the incentives for insurers), 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.[8]

What follows is a more formal exposition of the conditional correlation test for the examination of adverse selection and the mitigating effect of the risk-adjustment system.

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

$$cost_i = g(X_i, \mu_i) \tag{3.1}$$

[8]An other possibility would be to use only the expenses for care by the general practitioner (GP). These costs are exempted from deductibles, therefore it should not be affected by moral hazard. Differences in GP costs between enrollees with a high voluntary deductible and without voluntary deductible reflect pure the selection effect in theory. Although enrollees may decrease their GP visits in practice when having a deductible (e.g. because they are not aware of the exemption of GP costs from deductible or because they want to avoid referral to a specialist, the cost of which does fall under the deductible), it is an interesting possibility to study in future research.

$$deductible_i = h(X_i, \nu_i) \tag{3.2}$$

where μ_i and ν_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 relevant characteristics, then the error terms μ and ν 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 3.2, 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 ν. A significant negative correlation between the error terms μ and ν 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:

$$cost_i = \alpha \ deductible_i + \beta X_i' + \gamma Z_i' + \epsilon_i \tag{3.3}$$

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.

An important question is what we should include in vector X, i.e. what the relevant consumer characteristics are. In general, researchers are interested whether consumers have private information on their health or risks that an insurer cannot observe and thus cannot use when setting the price of the insurance. In this case, the observables in the test are the consumer characteristics that the insurer can observe. In our case, insurers are not allowed to differentiate the premium or the premium discount for the voluntary deductible. Insurers do not use any consumer characteristics in their pricing; consequently, our vector X can be empty in principle. However, we will include a number of consumer characteristics step-by-step so that we can control for risk aversion.

Heterogeneity in risk aversion may bias 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 and McGarry, 2006; Aarbu, 2017). As Finkelstein and McGarry (2006) demonstrate in the long-term care insurance market, consumers may have private information on their risk aversion (i.e. taste for insurance), and also on their risk (i.e. chance to need care in a nursing home). Consumers that know they have a high risk to be admitted to a nursing home, buy more long-term care insurance than consumers with a low risk. At the same time, consumers that have a preference for insurance buy insurance more often than consumers who are less risk averse. Also consumers with a preference for insurance, i.e. risk averse consumers, have lower risk of admission perhaps because they do more on prevention. In aggregate, those with more insurance are not higher risk consumers, thus the positive correlation test fails to identify the asymmetric information over risk-type in the aggregated data. The correlation can be found if one can control for the heterogeneity in risk aversion among the consumers.

In our application, we do not have an explicit measure of risk aversion of the enrollees. However, we can include explanatory variables in X that are good proxies for the risk aversion (Outreville, 2014; Carson et al., 2018; Li et al., 2007; Handel, 2013). In this way, we measure asymmetric information in the subgroups of consumers formed as the intersect of the consumer characteristics included in X.

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, (*i*) 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; and (*ii*) we replace the health care costs incurred in that year with the health care costs of the previous year, year $(t-1)$. By looking back three years, we deselect enrollees who try to plan their healthcare expenditures and adjust their level of voluntary deductible accordingly.[9] The expenditures of these enrollees are probably more volatile in time and more prone to moral hazard then the expenditures of less shrewd enrollees. By replacing the costs of year t by the costs of year $t-1$, we try to filter out the rest of moral hazard, e.g. skipping to visit a specialist for a general check-up.

This selection and replacement result 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:

$$cost_{i,t-1} = \alpha\ deductible_{it} + \beta X'_{it} + \gamma Z'_{it} + \epsilon_{it} \tag{3.4}$$

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

[9]For example a healthy enrollee may generally opt for a high voluntary deductible. When he needs a non-acute operation, e.g. knee replacement, he can try to postpone it until next year when he opts for no deductible. The year after that, he can switch back again to the high level of deductible.

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. When we choose risk-adjusted result as dependent variable, we can measure to what extent the risk-adjustment system can mitigate the selection effect that the insurer faces. This is interesting to measure, because it can be interpreted as incentives for the insurer for risk selection. For these reasons, we also estimate the following model:

$$result_{i,t-1} = \tilde{\alpha}\ deductible_{it} + \tilde{\beta} X'_{it} + \tilde{\gamma} Z'_{it} + \tilde{\epsilon}_{it} \tag{3.5}$$

The conditions to avoid endogeneity in the estimated equations, i.e. to gain unbiased coefficients, are that there are no omitted variables, and explanatory variables are not determined simultaneously with the dependent variable. An omitted variable may be risk aversion. We will control for this by using consumer characteristics, e.g. age, gender and income that are found to have correlation with risk aversion (Outreville, 2014). Simultaneous determination of the variable deductible with the dependent variable cannot be a problem, because we use the cost (result) of year $t-1$. Although costs in year t and $t-1$ are correlated, we know that there is no reversed causality between deductible and costs $t-1$ because costs in year $t-1$ cannot be effected by the choice of deductible in year t.

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, what could introduce selection bias in our estimate. If the characteristics, in which our sample differs from the whole population 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 et al. (2015) showed that in the Netherlands for a large share of enrollees who has not chosen any voluntary deductible would have been better of if they would have chosen a voluntary deductible. No voluntary deductible was the default option for almost all enrollees at the introduction of the actual health insurance system in 2006; furthermore, a large share of enrollees is inert, they stay long in the default option. Because we exclude the enrollees that chose a high voluntary deductible earlier in the period, our subsample may include an over-representation of inert enrollees, i.e. enrollees that are less interested and/or less shrewd when choosing their deductible. If this is the case, and less inert enrollees have also less costs (i.e. they can save more by choosing a voluntary deductible), then we may underestimate the adverse selection effect. This means, that the difference in healthcare costs between enrollees choosing no deductible and enrollees choosing a high deductible can be even more than it seems in our analysis.

3.5 Data description

3.5.1 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 3.1 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 health care costs are the major costs of health care[10], the cost variable that we use in our analysis are defined as the sum of these two cost components[11].

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

Table 3.1: Size of the dataset

[10]Some fixed cost components are not included in risk-adjustment.

[11]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.

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 3.4 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.

3.5.2 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 3.2 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 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 3.3 compares the mean values of different variables for 2012 (before choosing a deductible) and 2013 (after choosing a deductible) for each deductible category in 2013. The first two columns provide insight in the health care use by the individuals; the next two columns show the profitability of each subgroup in the risk-adjustment system; and the last two columns characterize the attractiveness of a higher deductible option from the enrollees' perspective.

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

Table 3.2: For each deductible category the number of enrollees and the share of healthy enrollees in 2012. Only enrollees who did not choose a voluntary deductible in the period 2010-2012 are included in the calculation.

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)[12]. 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 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 3.3 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

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

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). The calculations cannot take into account moral hazard, i.e. the behaioral change of enrollees due to the higher level of deductible. Since moral hazard is likely to decrease the expenses, our estimations of the counterfactual are probably upward-biased.

Focusing on the column headed 'mean expenses 2012' in Table 3.3, 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).

	mean cost		mean result		mean expenses	
deductible 2013	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

Table 3.3: 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 did not choose a voluntary deductible in the period 2010-2012 are included in the calculation.

Combining the findings from Tables 3.2 and 3.3, 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.

3.6 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 3.4. 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 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:

$$\begin{aligned} cost_{i,t-1} = \alpha + \sum_{k=1}^{n=6-1} \beta_k deductible_{kit} + \sum_{k=1}^{n=7-1} \theta_k age_{kit} + \\ + \delta male_{it} + \sum_{k=1}^{n=3-1} \mu k income\_src_{kit} + \sum_{k=1}^{n=4-1} \lambda_k income_{kit} + \\ + \sigma oneperson\_hh_{it} + + \sum_{k=1}^{n=59-1} \gamma_h healthplan_{kit} + \epsilon_{it} \end{aligned} \quad (3.6)$$

$$\begin{aligned} result_{i,t-1} = \tilde{\alpha} + \sum_{k=1}^{n=6-1} \tilde{\beta}_k deductible_{kit} + \sum_{k=1}^{n=7-1} \tilde{\theta}_k age_{kit} + \\ + \tilde{\delta} male_{it} + \sum_{k=1}^{n=3-1} \tilde{\mu}_k income\_src_{kit} + \sum_{k=1}^{n=4-1} \tilde{\lambda}_k income_{kit} + \\ + \tilde{\sigma} oneperson\_hh_{it} + \sum_{h=1}^{n=59-1} \tilde{\gamma}_k healthplan_{kit} + \tilde{\epsilon}_{it} \end{aligned} \quad (3.7)$$

In model (3.6) $cost_{i,t-1}$ is the cost of enrollee i in 2012 and in model (3.7) $result_{i,t-1}$ is the risk-adjusted result for enrollee i in 2012. In both models, $deductible_{kit}$ 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_{kit} (7 categories), gender dummy variables $gender_{kit}$ (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 $income_{kit}$ (4 categories), one-person household dummy variables $oneperson\_hh_{kit}$ (2 categories). These variables are taken from the risk classes from the risk-adjustment system. See section 3.3 for a description of these variables.

In both models dummy variable $healthplan_{kit}$ 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_{it} and $\tilde{\epsilon_{it}}$ are the error terms.

In our estimation, we did not include all risk classes 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 3.4, 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.

In the literature on measuring risk aversion, the most often studied variables are gender, age, and income. Several studies showed strong relationship between these variables and the level of risk aversion (Outreville, 2014; Carson et al., 2018; Li et al., 2007; Handel, 2013). There are also a few studies that looked at the effect of occupation/education on risk aversion and found a significant effect (Outreville, 2014). We also know that these variables affect the healthcare costs directly (Meerding et al., 1998; Ettner, 1996; van Vliet et al., 2012). Therefore, in the model on costs (equation 3.6), the coefficients describe the sum of direct and indirect effect of these explanatory variables on costs. The coefficient of deductible (β_k) can be interpreted as the a measure for adverse selection in the consumer groups formed by the intersection of the consumer characteristics used in the equation.

In the model on risk-adjusted result (equation 3.7), the dependent variable is already adjusted for consumer characteristics. However, consumer characteristics are still important to control for heterogeneity in risk preference among consumers. The coefficient of the variables reflects only residual effects that stem from the imperfection of the risk-adjustment system applied to our subsample. The results may not hold for the whole population; therefore, cannot be interpreted as results on the goodness of the risk-adjustment system.

We estimated models (3.6) and (3.7) with the ordinary least squares (OLS). Table 3.4 shows the results of the OLS estimation of the cost model (3.6) and Table 3.5 shows the results of the OLS estimation of the risk-adjusted result model 3.7. 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 3.5.2, we estimated the model using the dataset of enrollees who chose no voluntary deductible in the period 2010-2012[13]

Table 3.4 econometrically confirms our conclusions from the descriptive analysis in section 3.5.2: 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 other independent variables 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. The added variables can explain a part of the cost variation indeed. The coefficients are feasible and consistent with results on the effect of these characteristics on healthcare costs (van Vliet et al., 2012). It is important to note, however, that the results are specific to our sample and cannot be generalized to the whole population.

However, these coefficients are not directly important to our analysis. We included them in the model to control for heterogeneity in risk preferences among the enrollees. The consumer groups defined by the intersection of the categories of the consumer characteristics, are homogeneous in risk aversion. The coefficient of the deductible categories in the last model (3) are still significant and negative, which is a signal for adverse selection. 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 (when controlling for several consumer characteristics and risk aversion, as described).

Similarly, Table 3.5 confirms our conclusions from the descriptive analysis in section 3.5.2: 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. The risk-adjustment system cannot compensate for the whole cost difference between the enrollees. 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 deductible

[13] As a check of robustness, we also made the same calculation with enrollees who not only changed the level of their voluntary deductible in 2013 but also switched health plan. The results are similar to the results presented here, and they do not alter our conclusions. These additional calculations are not shown here but they are available from the authors upon request.

	Dependent variable: cost 2012		
	(1)	(2)	(3)
Constant	2,261.756***	3,578.001***	3,391.542***
	(1.696)	(13.733)	(20.448)
deductible_2013 100	−1,490.718***	−1,270.968***	−669.262***
	(35.081)	(35.103)	(34.263)
deductible_2013 200	−1,618.965***	−1,482.874***	−852.448***
	(27.475)	(27.486)	(26.834)
deductible_2013 300	−1,608.693***	−1,491.396***	−871.760***
	(36.662)	(36.646)	(35.768)
deductible_2013 400	−1,634.726***	−1,514.096***	−957.012***
	(62.731)	(62.673)	(61.157)
deductible_2013 500	−1,661.473***	−1,412.512***	−891.700***
	(10.969)	(11.371)	(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 3.4: OLS estimation of cost 2012 on deductibles 2013.

	Dependent variable: result 2012		
	(1)	(2)	(3)
Constant	-26.029^{***}	-105.554^{***}	-582.410^{***}
	(1.447)	(11.730)	(17.893)
deductible_2013 100	312.270^{***}	300.549^{***}	345.379^{***}
	(29.916)	(29.984)	(29.982)
deductible_2013 200	409.890^{***}	403.887^{***}	445.349^{***}
	(23.429)	(23.478)	(23.481)
deductible_2013 300	431.057^{***}	430.744^{***}	469.269^{***}
	(31.264)	(31.302)	(31.298)
deductible_2013 400	471.235^{***}	471.247^{***}	501.459^{***}
	(53.494)	(53.534)	(53.516)
deductible_2013 500	445.857^{***}	446.350^{***}	475.240^{***}
	(9.354)	(9.713)	(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

Table 3.5: OLS estimation of risk adjusted result 2012 on deductibles 2013.

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

The coefficients of consumer characteristics in model (3) reflect the imperfection of the risk-adjustment system when applying it to our sample. Our purpose to include them in the model is to control for heterogeneity in risk aversion and not to evaluate the risk-adjustment system. Therefore, the coefficients are not relevant in our analysis. They also cannot be interpreted as results on the goodness of the risk-adjustment system because they cannot be generalized to the whole population.

3.7 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. The risk-adjustment system cannot fully compensate the cost difference between enrollees with high voluntary deductible and without voluntary deductible, which leaves insurers incentives for selection.

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.

Given the substantial expected gains on the risk-adjustment for enrollees with a voluntary deductible, it would be to study in greater detail how the gains from the risk-adjustment system 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 stipulates that insurers are not allowed to differentiate the price discounts associated with voluntary deductible options between enrollees of the same health plans.

It should also be noted that the literature on the 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).

Our paper provides convincing evidence that individuals have substantial private information that is 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. If solidarity (i.e. equity between healthy and unhealthy enrollees) has a high priority in the regulation of the insurance market, it is worth to consider abolishing differentiation in coverage, such as the level of deductible. The disadvantage of this policy option would be the less possibility for enrollees to choose the health plan that fits their needs.

An other option that the Dutch government could consider to fully eliminate the risk-adjustment gain of the deductibles without abolishing voluntary 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[14]. Also, as van Kleef et al. (2008b) explain, insurers could probably offer only lower premium discounts for choosing high voluntary deductible because there were no gains from the risk-adjustment system. Lower premium discount may then lead to a lower number of enrollees opting for a voluntary deductible. Consequently, the potential moral hazard effect would be lower too, which would be a disadvantage of this policy option. Because including the choice of voluntary deductible in the risk-adjustment system 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.

[14]Introducing a new risk-adjuster has further effects too. For example, if it is based on the actual costs, it may discourage efficiency, or can be manipulated by the insurer. The actual form of the risk-adjuster has to be determined carefully considering the different incentives it may give to the insurers.

Acknolegements

We would like to thank Erik Schut, Marco Varkevisser, Jan Boone, Cédric Argenton, Michiel Bijlsma, Johan Polder and Martin Gaynor for their valuable comments.

Appendix

3.A Summary statistics

Table 3.A.1 shows the summary statistics of the dependent variables of the analysis and the counterfactual expenses, also discussed in Table 3.3. In addition, Table 3.A.2 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 3.A.2 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.

Tables 8 to 12 present the descriptive statistics of the independent variables used in the analysis. Furthermore, Table 3.A.8 presents the strength of association between the independent variables in our analysis measured by Cramer's V. The range of Cramer's V is 0 to 1, where 0 means statistical independence and 1 full dependence.

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

Table 3.A.1: Summary statistics for cost, counterfactual (cf) out-of-pocket expenses and risk-adjusted result in 2012 and 2013. Only enrollees who did not choose any voluntary deductible in the period 2010-2012 are included in the calculation.

			share of deductible in 2013					
percentile	result mean	result SD	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

Table 3.A.2: For each deductible category the distribution of population percentiles. Only enrollees who did not choose a voluntary deductible in the period 2010-2012 are included in the calculation.

age	enrollees	share_enrollees_in_cathegory
[18,30)	1411925	0.13
[30,40)	1681241	0.15
[40,50)	2202266	0.20
[50,60)	2113302	0.19
[60,70)	1881881	0.17
[70,70+)	1899101	0.17

Table 3.A.3: Descriptive statistics: age

gender	enrollees	share_enrollees_in_cathegory
F	5831121	0.52
M	5358595	0.48

Table 3.A.4: Descriptive statistics: gender

income	enrollees	share_enrollees_in_cathegory
group1 (Low)	2588304	0.23
group2 (Middle)	4620463	0.41
group3 (High)	3838548	0.34
group0 (rest)	142401	0.01

Table 3.A.5: Descriptive statistics: income

income_src	enrollees	share_enrollees_in_cathegory
benefitsreceiver	1104470	0.10
rest	9485427	0.85
selfemployed	599819	0.05

Table 3.A.6: Descriptive statistics: income source

oneperson_hh	enrollees	share_enrollees_in_cathegory
0 (No)	9156830	0.82
1 (Yes)	2032886	0.18

Table 3.A.7: Descriptive statistics: oneperson household

	age	income	gender	income_source	oneperson_hh	healthplan
age	1.00	0.16	0.05	0.16	0.20	0.11
income	0.16	1.00	0.04	0.14	0.17	0.13
gender	0.05	0.04	1.00	0.08	0.04	0.05
income_source	0.16	0.14	0.08	1.00	0.10	0.13
oneperson_household	0.20	0.17	0.04	0.10	1.00	0.08
healthplan	0.11	0.13	0.05	0.13	0.08	1.00

Table 3.A.8: Association matrix (Cramer's V)

Bibliography

Aarbu, K. O. Asymmetric information in the home insurance market. *Journal of Risk and Insurance*, 84(1):35–72, 2017.

Abbring, J. H., Heckman, J. J., Chiappori, P.-A., and Pinquet, J. Adverse selection and moral hazard in insurance: Can dynamic data help to distinguish? *Journal of the European Economic Association*, 1(2-3):512–521, 2003.

Aron-Dine, A., Einav, L., and Finkelstein, A. The rand health insurance experiment, three decades later. *The Journal of Economic Perspectives*, 27(1): 197–222, 2013.

Bijlsma, M., Boone, J., and Zwart, G. Competition leverage: how the demand side affects optimal risk adjustment. *The RAND Journal of Economics*, 45(4): 792–815, 2014.

Bolhaar, J., Lindeboom, M., and van der Klaauw, B. A dynamic analysis of the demand for health insurance and health care. *European Economic Review*, 56 (4):669–690, 2012.

Brot-Goldberg, Z., Chandra, A., Handel, B., and Kolstad, J. What does a deductible do? the impact of cost-sharing on health care prices, quantities, and spending dynamics. Technical report, National Bureau of Economic Research, 2015.

Carson, J., Dumm, R., Halek, M., and Liebenberg, A. What factors portend changes in household relative risk aversion? *Journal of Insurance Issues*, 41 (1):1–21, 2018.

Chiappori, P.-A. and Salanie, B. Testing for asymmetric information in insurance markets. *Journal of political Economy*, 108(1):56–78, 2000.

Chiappori, P.-A., Jullien, B., Salanié, B., and Salanie, F. Asymmetric information in insurance: General testable implications. *The RAND Journal of Economics*, 37(4):783–798, 2006.

Cohen, A. and Siegelman, P. Testing for adverse selection in insurance markets. *Journal of Risk and Insurance*, 77(1):39–84, 2010.

Croes, R., Katona, K., Mikkers, M., and Shestalova, V. Evidence of selection in a mandatory health insurance market with risk adjustment. Discussion paper 2018-013, TILEC, Tilburg University, 2018.

Cutler, D. M., Zeckhauser, R. J., et al. Adverse selection in health insurance. Technical report, National Bureau of Economic Research Cambridge, Mass, 1997.

Dardanoni, V. and Donni, P. L. Incentive and selection effects of medigap insurance on inpatient care. *Journal of Health Economics*, 31(3):457–470, 2012.

Davis, K., Doty, M. M., and Ho, A. How high is too high? implications of high-deductible health plans. publication, Commonwealth Fund, 2005.

Einav, L. and Finkelstein, A. Selection in insurance markets: Theory and empirics in pictures. *The Journal of Economic Perspectives*, 25(1):115–138, 2011.

Ettner, S. L. New evidence on the relationship between income and health. *Journal of health economics*, 15(1):67–85, 1996.

Finkelstein, A. and McGarry, K. Multiple dimensions of private information: Evidence from the long-term care insurance market. *The American Economic Review*, pages 938–958, 2006.

Fronstin, P. and Collins, S. R. Findings from the 2007 ebri/commonwealth fund consumerism in health survey. *EBRI Issue Brief*, (315), 2008.

Galbraith, A. A., Ross-Degnan, D., Soumerai, S. B., Rosenthal, M. B., Gay, C., and Lieu, T. A. Nearly half of families in high-deductible health plans whose members have chronic conditions face substantial financial burden. *Health Affairs*, 30(2):322–331, 2011.

Gerfin, M. and Schellhorn, M. Nonparametric bounds on the effect of deductibles in health care insurance on doctor visits–swiss evidence. *Health economics*, 15 (9):1011–1020, 2006.

Handel, B. R. Adverse selection and inertia in health insurance markets: When nudging hurts. *The American Economic Review*, 103(7):2643–2682, 2013.

Keeler, E. B. Effects of cost sharing on use of medical services and health. Technical Report RP-1114, Greenbranch Publishing, 1992.

Li, D., Moshirian, F., Nguyen, P., and Wee, T. The demand for life insurance in oecd countries. *Journal of Risk and Insurance*, 74(3):637–652, 2007.

Lohr, K., Brook, R., Kamberg, C., Goldberg, G., Leibowitz, A., Keesey, J., Reboussin, D., and Newhouse, J. Use of medical care in the rand health insurance experiment: diagnosis-and service-specific analyses in a randomized controlled trial. *Medical care*, 24(9):S1–S87, 1986.

Meerding, W. J., Bonneux, L., Polder, J. J., Koopmanschap, M. A., and van der Maas, P. J. Demographic and epidemiological determinants of healthcare costs in netherlands: cost of illness study. *Bmj*, 317(7151):111–115, 1998.

Newhouse, J., Price, M., Hsu, J., McWilliams, J., and McGuire, T. How much favorable selection is left in medicare advantage? *American journal of health economics*, 2015.

Newhouse, J. Consumer-directed health plans and the rand health insurance experiment. *Health Affairs*, 23(6):107–113, 2004.

Newhouse, J., Price, M., Huang, J., McWilliams, J., and Hsu, J. Steps to reduce favorable risk selection in medicare advantage largely succeeded, boding well for health insurance exchanges. *Health Affairs*, 31(12):2618–2628, 2012.

NZa. Zorgverzekeringsmarkt 2014: Weergave van de markt 2010-2014. Technical report, 2014.

NZa. Marktscan zorgverzekeringsmarkt 2015 deel a. Report, Nederlandse Zorgautoriteit, 2015.

NZa. Kwantitatief onderzoek naar risicoselectie en risicosolidariteit op de zorgverzekeringsmarkt. Report, Nederlandse Zorgautoriteit, 2016.

Olivella, P. and Vera-Hernández, M. Testing for asymmetric information in private health insurance. *The Economic Journal*, 123(567):96–130, 2013.

Outreville, J. F. Risk aversion, risk behavior, and demand for insurance: A survey. *Journal of Insurance Issues*, pages 158–186, 2014.

Panthöfer, S. Risk selection under public health insurance with opt-out. *Health economics*, 25(9):1163–1181, 2016.

Rothschild, M. and Stiglitz, J. Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *The Quarterly Journal of Economics*, 90(4):629–649, 1976.

Tollen, L. A., Ross, M. N., and Poor, S. 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, 2004.

van de Ven, W. P. and Ellis, R. P. Risk adjustment in competitive health plan markets. *Handbook of Health Economics*, 1(Chapter 14):755–845, 2000.

van Kleef, R., Beck, K., van de Ven, W., and van Vliet, R. Risk equalization and voluntary deductibles: a complex interaction. *Journal of Health Economics*, 27(2):427–443, 2008a.

van Kleef, R., Beck, K., van de Ven, W., and van Vliet, R. Risk equalization and voluntary deductibles: A complex interaction. *Journal of Health Economics*, 27(2):427–443, 2008b.

van Kleef, R., McGuire, T., van Vliet, R., and van de Ven, W. Improving risk equalization with constrained regression. *The European Journal of Health Economics*, 18(9):1137–1156, 2017.

van Vliet, R., Asselt, M., Mazzola, G., Goudriaan, R., and Everhardt, T. Berekening van normbedragen risicovereveningsmodel 2013. Report, APE, Aarts de Jong Wilms Goudriaan Public Economics bv, 2012.

van Winssen, K., van Kleef, R., and van de Ven, W. How profitable is a voluntary deductible in health insurance for the consumer? *Health Policy*, 119(5):688–695, 2015.

van Winssen, K., van Kleef, R., and van de Ven, W. A voluntary deductible in health insurance: the more years you opt for it, the lower your premium? *The European Journal of Health Economics*, pages 1–18, 2016.

Vektis. Zorgthermometer verzekerden in beeld 2017. Report, Vektis, 2017.

Wharam, J. F., Landon, B. E., Galbraith, A. A., Kleinman, K. P., Soumerai, S. B., and Ross-Degnan, D. Emergency department use and subsequent hospitalizations among members of a high-deductible health plan. *Jama*, 297(10): 1093–1102, 2007.

Chapter 4

Substitutability of health plans and bargaining in the health purchasing market

Katalin Katona and Rein Halbersma

Abstract

We study the welfare effects of changes in the degree of substitutability in the health insurance market in a theoretical model. We model negotiations between insurers and hospitals in the health purchasing market explicitly. In our bilateral duopoly setting, insurers and hospitals bargain on two-part prices, and insurers compete on premiums and contracted network of hospitals. We show that insurers can alleviate competition in the insurance market by contracting with the same hospitals, which principle also appears in models of common agency and strategic delegation. In our model, this 'common agency effect' can be strong enough to compensate or even to overwhelm the competitive pressure on premiums. Consequently, consumer and total welfare can decrease due to increased substitutability of health plans in our model, which is in contrast to results in standard models of horizontal product differentiation.

JEL Classification Numbers: L51, D43, I13

Keywords: health insurance market, substitutability, insurer-hospital bargaining

4.1 Introduction

In jurisdictions where the working of health insurance market is based on the principles of competition, such as the Netherlands and Switzerland, consumers are supposed to compare all available insurance policies and choose the one that best fits their needs. However, health insurance policy is a complex product which is not easy to understand and compare (Loewenstein et al., 2013). Besides the difficulty for consumers to optimize over uncertain future health care needs, insurance companies may generate further complexity among health plans by brand building or other marketing techniques. High complexity and overwhelming differentiation often result in choice overload of consumers (Chernev et al., 2015) that subsequently leads to suboptimal market outcomes.

Regulation on the health insurance market can support informed choice of consumers by, for example, standardizing the supply to a certain extent. The idea is to reduce the number of attributes along which health plans can be differentiated, which promotes transparency. For example, as Ericson and Starc (2016) describes, regulators of Massachusetts Health Insurance Exchange standardized the financial cost-sharing characteristics (deductible and co-payment levels) of the insurance plans in 2010. Before the policy intervention, consumers were supposed to compare complex financial contingencies for each plan. After standardization, only seven distinct cost-sharing plan types were allowed.[1] In the Netherlands, health insurance companies are obliged to report which plans of them are identical in terms of contracted network and reimbursement policy. Due to this regulation consumers can focus on the remaining differences in, for example, services. The aim of these regulations is to improve market outcomes, i.e. increase total and consumer welfare. Standardization can be a tool to increase substitutability, which should lead to higher levels of competition in the market and enhance welfare.

We study the welfare effects of increasing substitutability by means of a theoretical bilateral duopoly model of two insurers and two hospitals. Our results contrast with the tenet in industrial organization according to which more homogeneity between competing products yield higher level of consumer surplus and lower level of industry profit. In our theoretical model, increasing degree of homogeneity among health plans at high levels of substitutability can decrease welfare and increase industry profits.

There are a two other theoretical papers stating that increasing substitutability might increase the industry profit. Zanchettin (2006) studies how industry profit changes with the degree of horizontal product differentiation. His result

[1] However, insurers were required to offer all seven designs. The policy intervention also simplified the choice interface on the market places' website by structuring the choice: consumers first had to choose a standardized financial package and after that they opt for an insurer.

of decreasing industry profit follows from the asymmetry in the efficiency of the firms. When the degree of product differentiation decreases, two opposite effects emerge. First, competition increases between the two firms, leading to decreased prices and profits (price effect). Second, there is a shift of production from the inefficient firm to the efficient firm (the selection effect), which increases the total industry profit. Consider, for example, two insurers in the market. Both sell a health insurance but the two health plans differ in the services they offer, such as the possibility of cost reimbursement on-line or support when choosing a hospital. Furthermore, assume that one of the insurers is more efficient, i.e. has lower cost per enrollee, than the other insurer. In a differentiated Bertrand competition both insurers can be active, regardless of the difference in efficiency, because consumers are willing to pay for the services of their preferred insurer. When the two health plans become closer substitutes, e.g. because the insurers introduce similar new services, consumers experience less difference in services and will pay more attention to price differences. Fiercer price competition will then drive prices down (price effect), which may force the inefficient insurer to exit (selection effect). The shift in production from the inefficient to the efficient insurer increases, ceteris paribus, the total industry profit. Zanchettin (2006) shows that when the original level of product differentiation is low, the selection effect is stronger than the price effect, and an increase in substitutability yields higher industry profit.

Fanti (2013) also investigates the relationship between industry profit and horizontal product differentiation. He incorporates wage bargaining with firm-specific unions and cross-ownership in his model.[2] The asymmetry in his duopoly model lies in the latter: shareholder A has shares in both firms, while shareholder B has shares only in one firm. Because of the asymmetric cross-ownership, the two firms do not react similarly strongly to changes in the level of product differentiation and their production levels will be asymmetric. When wages are not exogenously given but bargained with a union, the different production levels yield cost-asymmetry. Just like in Zanchettin (2006), cost-asymmetry may reverse the results of the standard product differentiation models.

The negative price effect of decreasing product differentiation, shown in the previous models, is present in our model too. The upward price pressure that reverse the results is caused, however, by bargaining between insurers and hospitals. There is a positive effect on prices when substitutability increases because insurers bargain with the same hospitals in the health purchasing market. Insurers can alleviate competition in the insurance market by indirectly coordinating wholesale prices through bargaining with the same hosptials. The new feature in

[2]In the Netherlands, wages are indeed bargained with a union, but not at the level of individual firms.

our model is that it yields unconventional results even if hospitals and insurers are symmetric.

The mechanism of coordination through bargaining is similar to the one in common agency models (Bernheim and Whinston, 1986; Gal-Or, 1991). A similar effect is also described in the literature on strategic delegation and strategic separation (for an overview see Kopel and Pezzino, 2018). The main idea is that a principal (e.g. manufacturer) can commit itself to softening competition by delegating the decision on price setting to an agent (e.g. a retailer) who has just different incentives. When this delegation and the contract between the manufacturer and retailer are observable, it can influence the expectations and actions of the competitors. If the manufacturer sets the wholesale prices higher than marginal cost, the retailer is forced to set higher retail prices than the competitive level. This induces the competitors to raise the prices too if the products are substitutes. Corts and Neher (2003) show that a strategic effect of delegation is also possible in the case of unobservable contracts between manufacturers and retailers if (and only if) there are multiple retailers and they receive an ownership stake. These conditions are applicable to our model because a hospital (which corresponds to the manufacturer) contract multiple insurers (which correspond to retailers), and they agree on the wholesale price by maximizing their joint profit (which corresponds to an ownership stake). Our model adds to this literature by modelling bargaining between the hospitals and insurers and looking at the welfare effects of changing degree of product differentiation.

We apply a version of the bargaining model in de Fontenay and Gans (2014) that is also applied in Douven et al. (2014). In our model, two insurers and two hospitals bargain bilaterally on two-part prices. They set the wholesale unit price to maximize their joint bilateral profit. Through the fixed fee, they can split the gain from the contract. Although, the bargaining pairs negotiate independently from each other, hospitals profit from the gains of both insurers on the insurance market. Through the wholesale prices, hospitals can help insurers to commit to a minimum premium. We see that it is not profitable to cut wholesale prices under the competitor's prices even at high levels of substitutability in our model. Therefore, wholesale prices are monotonously increasing in substitutability, which gives an increasing upward pressure on the premium. Just like higher degree of substitutability makes insurers compete more fiercely, this upward pressure increases with substitutability because the negotiating pair is more sensitive to the price of the other insurer when the products are more comparable. The upward pressure compensates the downward pressure of price effect, and at a certain point, exceeds it. When product differentiation further decreases, premiums and industry profit go up, while consumer and total welfare go down. When the market approaches perfect substitutability of health plans, consumer and total welfare converge to a suboptimal level in our model.

Common agency effect through contracting the same hospitals emerges in the model of Gal-Or (1997) as well. Gal-Or (1997) studies in a bargaining model of two insurers and two hospitals whether a hospital or insurer will be excluded from the market in equilibrium. In her bargaining model, insurer - hospital pairs negotiate over a unit wholesale price in simultaneous bilateral Nash bargaining games. Even though hospitals and insurers are treated symmetrically in each bargaining, only insurers can optimize over hospital networks that they contract and premiums for the associated health plans. Hospitals, in contrast, take the insurers' choice of network and premium as given. Gal-Or finds that selective contracting, i.e. contracting a single hospital, can arise in equilibrium. She also concludes that insurers are better off contracting the same hospital than each contracting a different hospital, which she describes as a similar effect as in models of common agency. In her model however, this effect is never stronger than the price effect due to more intensive competition.

The structure of our paper is as follows: in section 4.2, we describe our model. Besides our main focus, the model of bilateral duopoly, we also discuss the case of two competing vertical chains (when the insurers contract a single, not common hospital), and the case of upstream monopoly (when the insurers contract a single, common hospital). In this way, we can illustrate how contracting the same hospital(s) affects the equilibrium outcomes of the model. Section 4.3 presents the results focusing on the welfare effects of changes in substitutability between health plans. Finally, section 4.4 discusses our findings, while section 4.5 concludes.

4.2 The model

Hospital markets, just like insurance markets, are typically concentrated (Gaynor et al., 2015; Schut and Varkevisser, 2017) hence the structure of the health purchasing market can be the best described as a two sided oligopoly. Contracting between hospitals and insurers involves bargaining.[3] To properly analyze wel-

[3]Our reason to include bargaining in our theoretical model is the following. First, contracted hospital network is an important characteristic of health insurance policies. At the moment of buying an insurance, consumers also commit themselves to the contracted provider network. Whether a health plan has a narrow or broad network of hospitals, affects the choice of consumers. Second, hospital costs count for a high percentage of health care expenditure. The share of hospital care in the national health expenditures in the US was 32.3% in 2015 (NCHS, 2017). In the Netherlands, 32,5% of expenditure on personal health care was spent on hospital care in 2014 (Bakx et al., 2016). Consequently, reimbursed hospital costs strongly affect the marginal cost of insurers. Finally, hospital prices are typically not exogenously given, but they are a result of negotiations between insurers and hospitals.

fare effects of decreasing product differentiation in the insurance market, we treat negotiations between hospitals and insurers as endogenous actions in our model.

Our model consists of the following multi-stage game

0. Number of hospitals and insurers, and the negotiable contracts between them exogenously defined.
1. Multilateral bargaining among hospitals and insurers along the lines of de Fontenay and Gans (2014) resulting in contracted hospital networks with corresponding lump sum transfers and wholesale unit prices.
2. Simultaneous price setting by the insurers in the insurance market.
3. Realization of consumer demand for insurance and hospital visits when needed.

In stage 0, we define two alternative market structures besides our main interest of bilateral duopoly. Our goal with presenting these additional models is to give a better insight in the mechanism yielding the unexpected results of our main model. Furthermore, the outcome in the third stage is fully determined by the consumer preferences and the strategic choices of hospitals and insurers in the preceding stages. We solve the first two stages, multilateral bargaining followed by simultaneous downstream price setting, through backward induction and present the results of the perfect Bayesian equilibrium.

4.2.1 Insurance and hospital market

We consider differentiated price competition between two health insurers i and $-i$ in the health insurance market. In the hospital market, we model different situations that are defined in stage 0. We focus on the game that defines two hospitals in stage 0 and contracting between both insurers and hospitals is possible. Besides this, we also analyze two alternative games. First, we assume in stage 0 that only exclusive contracting is possible between the two hospitals and insurers. That would result in two competing vertical chains. Second, we discuss the game when there is a single hospital but it may contract both insurers, which yields upstream monopoly as market structure if both contracts are signed in equilibrium. We choose these two alternative market structures to analyze because the insurers are symmetric regarding their contracted hospital network just like in our main focus, and there is also downstream competition, which is crucial to be able to analyze the effect of product differentiation.

We allow for horizontal product differentiation in the insurance market parametrized through the degree of substitutability μ.[4] This parameter can be interpreted as

[4]We use the terms 'degree of substitutability' and 'degree of differentiation' both to describe to which degree consumers consider the two health plans in the market as alternatives. When the degree

differences in the insurers' services such as the possibility of cost reimbursement only or as a difference in the perception of consumers by branding. This horizontal differentiation is exogenously given, however, it can eventually be affected by policy interventions in the insurance market, like standardization of the health plans.

Because we include negotiations between hospitals and insurers in our model, it is endogenously determined whether an insurer and hospital signs a contract. In our main model, this means that network size, i.e. contracting a single or both hospitals, is endogenously determined. We assume that consumers have a preference for a broad hospital network over a narrow network. Therefore, variation between the insurers in network size is an endogenous vertical differentiation of the health plans. We denote the set of contracted hospitals of insurer i as G_i where we use the notation $G_i = \{j : i \text{ and } j \text{ have a contract}\}$. We introduce $v(G_i)$ as the value of having access to a hospital network G_i.

To model the demand of consumers on the health insurance market, we employ the parsimonious yet flexible linear demand system of Shubik and Levitan (1980) considering substitutability and vertical differentiation between the health plans:[5]

$$q_i(p_i, p_{-i}, G_i, G_{-i}) = \frac{1}{2}\left[v(G_i) - p_i + \frac{\mu}{2}\left(\left(v\left(G_i\right) - p_i\right) - \left(v\left(G_{-i}\right) - p_{-i}\right)\right)\right] \quad (4.1)$$

where q_i is the volume and p_i is the price of product i and μ is the level of substitutability.

Equation 4.1 gives the total demand that insurer i faces. It is a downward sloping function of the own price and it is increasing in the value of own hospital network. Intensity of competition, i.e. cross price elasticity and elasticity with regard to the value of the competitor's hospital network, depends on the degree of substitutability (μ) between the two plans. For $\mu = 0$, we obtain a monopolistic insurance market, whereas $\mu \to \infty$ describes a perfectly homogeneous insurance market.

In our model, μ has a similar role as the transportation parameter in Hotelling competition, often used to model product differentiation in the insurance market (e.g. Gal-Or, 1997; Douven et al., 2014; Bijlsma et al., 2014). Our demand model

of differentiation is low, consumers have no strong preferences to choose one of the alternatives, i.e. the degree of substitutability of the health plans is high.

[5]We consider this demand as a function that describes the observed volumes demanded in the insurance market. Although, we do not introduce this demand function as an aggregation of utility maximizing choices of individuals, it can be derived by maximizing the aggregated utility function 4.2 with regard to q_i and subject to a budget constraint. Under the assumption that utility maximizing choice of consumers maximizes aggregated utility too, i.e. consumers do not expose other consumers to externalities, we can argue that this demand represents the aggregation of individual demands.

has similar features as a Hotelling line but instead of a fixed demand, it uses a general downward sloping demand curve allowing for uninsured consumers.[6] As shown in Halbersma and Katona (2011) for the case of symmetric insurers, this demand function can be transformed in the limit into the fixed demand function of the Hotelling model. Furthermore, our demand function has the same functional form when approximated with the first order Taylor polynomial in prices as a logit demand function often used in empirical work (Halbersma and Katona, 2011).

Note that the market demand function ($q_i + q_{-i} = \frac{v(G_i)+v(G_{-i})-p_i-p_{-i}}{2}$) does not depend on the degree of substitutability (μ). As we will show, equilibrium demand does depend indirectly on substitutability through the equilibrium premiums. Also, demand of an individual insurer (q_i) does depend on μ unless the insurers are symmetric both in price and network size. When one of the insurers contracts a narrow network and the other one a broad network, the relative value of both (together with the value of μ) determines the maximally feasible premium difference, which still allows the insurer with narrow network to remain active in the market. When the two health plans are horizontally not differentiated, i.e. $\mu \to \infty$, and a broad network is strongly preferred to a narrow network, such asymmetric situation cannot exist in equilibrium.

Consumers derive utility from having health insurance, which we express by the aggregated utility function (4.2). As mentioned in footnote 5, this function can be seen as the underlying utility function of the demand function (4.1) under the assumption that consumers do not expose other consumers to externalities.

$$U(q_i, q_{-i}, G_i, G_{-i}) = v(G_i)q_i + v(G_{-i})q_{-i} - \frac{1}{2}q_i^2 - \frac{1}{2}q_{-i}^2 - \frac{q_i^2 + q_{-i}^2 + 2\mu q_i q_{-i}}{2(1+\mu)} \quad (4.2)$$

Terms with $v(G_i)$ in equation (4.2) reflect that consumers derive utility from health insurance. Terms with q_i^2 state that there is a satiation point, when no more consumers want a health insurance. Finally, the cross term $2\mu q_i q_{-i}$ describes the general taste for variety among consumers. Utility is increasing in the degree of substitutability (μ) as long as the two insurers do not sell the same volume of health plans, i.e. $q_i \neq q_{-i}$. When the two insurers have the same contracted network, they are symmetric and the equilibrium quantities of health

[6] This is for example applicable for the supplementary insurance in the Netherlands, which is not mandatory. In case of mandatory insurance, we could interpret the assumption of downward sloping demand curve as consumers seeking insurance in an unrelated market. For example, they buy a health plan from an insurer that contracts hospitals in a different region. However, in this case our analysis is a partial analysis because it does not consider the complete insurance market. Extending the model to capture the whole insurance market, would be an interesting direction for further research.

plans equal ($q_i = q_{-i}$). In this case, utility does not depend on the degree of substitutability (μ).

In the hospital market, we consider competition between two hospitals, H_j and H_{-j} with equal marginal costs which we normalize to zero, $c_j = c_{-j} = 0$.[7] Insurers do not use provider dependent co-payments. The hospitals are symmetric in all ex ante observable aspects (such as quality of treatment) but differ in some characteristics which are unobservable ex ante. One may think of the average availability of single rooms, on which there is no readily available comparison information ex ante. When the consumer needs to be hospitalized (ex-post), he gathers information about room types in, and distances to the hospitals relevant for him and chooses according to his preferences. We model this choice process in reduced form and assume that consumers visit hospitals in a fixed proportion.[8]

We denote the set of insurers with which hospital j has a contract as G_j, where we use the notation $G_j = \{i : j \text{ and } i \text{ have a contract}\}$. Because hospitals are symmetric, there are only three inequivalent insurance networks in our model: with no hospital (the insurer is unable to contract and leaves the market), one contracted hospital ('narrow network') or two contracted hospitals ('broad network').

$$v(G_i) = \begin{cases} 0, & \text{empty network.} \\ v_{narrow}, & \text{narrow network, i.e. single hospital contracted.} \\ v_{broad}, & \text{broad network, i.e. both hospitals contracted.} \end{cases} \tag{4.3}$$

We assume $0 < v_{narrow} \leq v_{broad}$.

Hospitals' market shares among the consumers of a specific insurer are fully determined by consumer preferences conditional on the insurer's network: $s_{ij} = s_j(G_i)$. Because we have only three possible networks, $s_j(G_i)$ has only three different values too.

$$s_j(G_i) = \begin{cases} 0, & \text{if } G_i = \emptyset\text{, i.e. an empty network.} \\ 1, & \text{if } G_i = \{j\}\text{, i.e. a narrow network.} \\ \frac{1}{2}, & \text{if } G_i = \{j, -j\}\text{, i.e. a broad network.} \end{cases} \tag{4.4}$$

Note that the specific value of s_j in case of a broad network does not affect the results. For simplicity of notation, we set it to 1/2.

[7] In the alternative market structure with a single hospital, we have a monopoly in the hospital market obviously.

[8] This modelling choice means that we do not catch the dynamic of the hospital competition, and only model the equilibrium. We assume that there are no changes in the characteristics of hospitals that would change the equilibrium outcome in the hospital market.

Hospital j's demand coming from insurer i, which we denote q_{ij}, is determined by the demand for insurance i (q_i) and the share of consumers that choses hospital j given the network of insurer i, i.e. $q_{ij} = q_i s_j(G_i)$. The total demand for hospital j is thus: $q_j = q_i s_j(G_i) + q_{-i} s_j(G_{-i})$. To save on notation, we assume implicitly that all enrollees of the insurer need hospital care. This assumption does not affect our qualitative results.

Hospital j's reimbursement from insurer i is given by a two-part price $T_{ij} = t_{ij} + w_{ij} q_{ij}$. Here, the fixed fee t_{ij} is a lump-sum transfer which can either be negative or positive. The wholesale price w_{ij} is paid over every unit of supplied quantity q_{ij}. The profit function of hospital j is the sum of reimbursement from contracted insurers, i.e. G_j.

$$\Pi_j = \sum_{i \in G_j} q_i s_j(G_i) w_{ij} + \sum_{i \in G_j} t_{ij} \tag{4.5}$$

Insurer i's profit function is then given by

$$\begin{aligned} \Pi_i &= q_i \left(p_i - \sum_{j \in G_i} w_{ij} s_j(G_i) \right) - \sum_{j \in G_i} t_{ij} \\ &= q_i \left(p_i - w_i \right) - \sum_{j \in G_i} t_{ij} \end{aligned} \tag{4.6}$$

where we have defined the insurer's weighted average wholesale price as $w_i = \sum_{j \in G_i} w_{ij} s_j(G_i)$.

Note that a model of mandatory insurance, such as the model of Douven et al. (2014), does not allow for endogenous determination of the wholesale prices when both insurers contract both hospitals. The reason is that the total market demand is not downward sloping in premiums, because it is fixed. Hospitals will get their share of the insured consumers no matter from which insurer they come. Only when wholesale prices influence the total market demand, as in our model, is the simultaneous optimization program for the wholesale prices well-defined. Otherwise, exogenously regulated wholesale prices need to be used.

4.2.2 The bargaining game

To model the bargaining between hospitals and insurers, we first introduce some notation. The union of all sets of contracts G_i and G_j is denoted as $G = \{(i, j) : i \in G_j, j \in G_i\}$. This means that G contains all pairs of firms that have a contract. In mathematical terms, G is an *undirected graph*.[9] Each firm

[9] G is undirected since if insurer i is negotiating with hospital j, then the converse is also true.

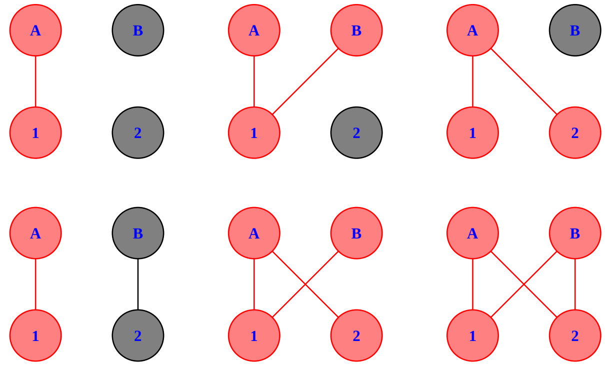


Figure 4.1: The six inequivalent graphs for a bilateral hospital-insurer duopoly. On the top row from left to right: a bilateral monopoly (BM), a downstream monopoly (DM) and an upstream monopoly (UM); on the bottom row from left to right: a disconnected duopoly (DD), an asymmetric duopoly (AD) and a connected duopoly (CD).

represents a vertex of the graph, and each contract can be drawn as an edge between the corresponding vertices of the firms. In our bilateral duopoly, there are six non-equivalent (sub)graphs: bilateral monopoly (BM, there is only one hospital and one insurer on the market), downstream monopoly (DM, one insurer contracting both hospitals), upstream monopoly (UM, both insurers contracting a single and the same hospital), disconnected duopoly (DD, both insurers contracting a single hospital that is not the same), asymmetric duopoly (AD, one insurer contracts a single hospital, the other insurer contracts both hospitals) and connected duopoly (CD, both insurers contract both hospitals). These six industry structures are depicted in Figure 4.1.

We use a modified version of the bargaining game of de Fontenay and Gans, which they describe as an application of their main model (de Fontenay and Gans, 2014, section V.). The bargaining model of de Fontenay and Gans (2014) is a non-cooperative game that gives the same equilibrium payoffs as a cooperative bargaining game dividing a (reduced) surplus. The surplus that the players split is reduced, i.e. non-cooperative, because of externalities between the bargaining pairs. In the following, we shortly present the non-cooperative bargaining game and its specific equilibrium properties that link its solution to the coalitional bargaining solution provided by de Fontenay and Gans (2014). Later on, we derive our results by applying these equilibrium properties instead of solving explicitly the bargaining game.

We present the sequential form incomplete information bargaining game described in (de Fontenay and Gans, 2014, Appendix, viii.). The game is sequential because hospital - insurer pairs negotiate in sequence. The information of players is incomplete because specific terms of the agreement (if reached) are assumed to be known exclusively by the bargaining pair. First, we describe the information structure of the game. Second, we present the sequential game in which the hospital networks of insurers are formed. Finally, we describe the bargaining game that each hospital-insurer pair plays.

Market parties have perfect information on the actions of the other players, i.e. they can observe if a negotiation breaks down. However, they have incomplete information on the agreed terms in the contract of other pairs because they can only observe their own agreement. Parties assume that all pairs signing a contract, agree on the equilibrium values given the industry structure. During the negotiations, parties also believe that all other bargaining pairs that negotiate later in the sequence will reach agreement. If a breakdown occurs, this belief is updated to take the breakdown into account. Furthermore, the belief structure of the game is defined as passive beliefs, i.e. beliefs on the specific terms in other pairs' contracts are not updated when an out-of-equilibrium action is encountered.[10]

Hospital - insurer pairs negotiate in sequence.[11] If a hospital - insurer pair reaches agreement, the next pair starts the negotiations. When a hospital - insurer pair does not reach an agreement, i.e. the negotiation breaks down, all contracts that have already been signed will be renegotiated. Pairs that failed to reach agreement are not allowed to reopen the negotiations any more. The renegotiation of the other contracts is practically a new subgame where all pairs, except the pairs that already broke down their negotiations, bargain again in sequence. The game ends when after reaching agreement(s) no more breakdown occurs or when every pair broke down its negotiation. The order in which the pairs negotiate is fixed, i.e. it is the same in all subgames. Because the equilibrium is independent of this order,[12] it can be chosen arbitrarily.

[10]This is often described as a negotiation between delegates of the market players where every negotiation is led by a separate delegate of the market player.

[11]This assumption is crucial to have a unique the equilibrium solutions. If we alter it to the potentially more realistic assumption of simultaneous bargaining of hospital - insurer pairs, we still get the same equilibrium result, although it may not be unique. Furthermore, the contracts must be binding and network contingent (i.e. an insurer and a hospital agree on a payment in case of each possible network status) in case of simultaneous negotiations (de Fontenay and Gans, 2014).

[12]The equilibrium outcomes of contracts are determined in a subgame where all possible contracts are signed. To give an intuition, assume that one negotiation breaks down. Then a new subgame would begin and all contracts signed before would be renegotiated. The game only concludes if there are no more breakdowns after signing any contract. Because parties assume that all contracts would

Now we turn to the bilateral bargaining game between hospitals and insurers. Each pairwise bargaining follows the procedure of Binmore et al. (1986). Specifically, bargaining takes place over time and parties make offers and reject or accept the offer in turns. If an offer is accepted, the bargaining concludes, if an offer is rejected, the process of alternating offers continues. There is an exogenously given probability that negotiations break down in case of the rejection of an offer.

In our context, the negotiating pairs bargain over the wholesale price w_{ij} and the lump-sum transfer t_{ij}.[13] The perfect Bayesian equilibrium strategy, which we are looking for, has to meet the condition of sequential rationality. Therefore, in the equilibrium strategy, parties sign the contract only if they are better off with the agreement than without it. This is an individual participation constraint that can be written as

$$\Pi_i \geq \widetilde{\Pi}_i, \quad \Pi_j \geq \widetilde{\Pi}_j \text{ for } i \in \{i, -i\} \text{ and } j \in \{j, -j\} \tag{4.7}$$

where $\widetilde{\Pi}_i$ and $\widetilde{\Pi}_j$ are the outside option of insurer i and hospital j, respectively, in case they cannot reach an agreement. The outside option of every pair is the situation without that specific contract, taking earlier breakdowns into account and assuming that all other pairs reach agreement.

However, the individual participation constraints have to be met in the outside option as well so that rationality holds in all subgames. This means that feasibility should be checked recursively to find the equilibrium network and the equilibrium outside options in the negotiations.

Equilibrium outcomes

As de Fontenay and Gans prove (de Fontenay and Gans, 2014, Teorem 1), any perfect Bayesian equilibrium involves bilaterally efficient wholesale prices, i.e. wholesale prices that maximize the joint surplus of the negotiating pair conditional on the industry structure and the agreed wholesale prices of the other negotiating pairs. Because of passive beliefs, the bargaining pair expects all other pairs to agree on their (bilaterally efficient) equilibrium wholesale prices given the industry structure. If an offer involves an other than profit maximizing wholesale price, it won't be accepted because a bilaterally efficient wholesale price ensures higher payoff for the parties. In equilibrium, all agreed wholesale prices are, therefore, bilaterally efficient, i.e. satisfy

be signed that are negotiated later in the sequence, and this is also the case in equilibrium, it does not matter in which order they negotiate.

[13]Two-part prices can be a realistic assumption even if hospital care prices are unit prices. Insurers may give volume discounts or insurers and hospitals may bargain not only on unit prices but also on turnover ceilings; both strategies give the price the form of a two-part price.

$$\frac{\partial(\Pi_i + \Pi_j)}{\partial w_{ij}} = 0 \text{ for } i \in \{i, -i\} \text{ and } j \in \{j, -j\} \tag{4.8}$$

Furthermore, bargaining parties split the surplus generated by their contract as they would do in a two agent Nash-bargaining game (de Fontenay and Gans, 2014, Theorem 2), which is also the result in Binmore et al. (1986).

$$\Pi_i - \widetilde{\Pi}_i = \Pi_j - \widetilde{\Pi}_j \text{ for } i \in \{i, -i\} \text{ and } j \in \{j, -j\} \tag{4.9}$$

de Fontenay and Gans (2014, Appendix, viii.) also prove that the solution satisfying bilateral efficiency (equation 4.8) and Nash-bargaining-like split of the surplus (equation 4.9) is the unique perfect Bayesian equilibrium of the above described sequential non-cooperative game of pairwise negotiations. The equilibrium outcomes (w_{ij} and t_{ij}) can also be found by maximizing the Nash product $\Pi_{ij} = (\Pi_i - \widetilde{\Pi}_i)(\Pi_j - \widetilde{\Pi}_j)$, with respect to w_{ij} and t_{ij}.

When deriving our results, we will use this equilibrium property of the game instead of solving the sequential game explicitly. Furthermore, we check feasibility (individual participation constraints of all parties) recursively in all subgraphs of the industry structure to find the equilibrium network.

The generalized Myerson-Shapley value

It should be noted that in the presence of downstream competitive externalities, the bargaining equations (4.8, 4.9) equations do not decouple for different hospital-insurer pairs (i, j) since the profits of each downstream firm will depend on the actions of its competitors. Hence, for any given graph G of contracts, the bargaining equations (4.8) and (4.9) have to be solved simultaneously for all $(i, j) \in G$. Moreover, to account for all the outside options, these equations have to be solved recursively for all edges in all subgraphs of G. As explained earlier, individual participation constraints have to be checked recursively too.

Because of the recursive combinatorial structure of the bargaining equations, some intermediate subgraphs drop from the equilibrium expressions of payoffs of the connected duopoly. In particular, configurations where one supply relationship has been severed but otherwise all firms remain connected to each other (asymmetric duopoly) do not appear in the equilibrium payoffs.[14] The explicit expression for the equilibrium payoffs, when all individual participation constraints are satisfied, is given by the weighted sum of values of particular coalitions of agents called the generalized Myerson-Shapley value (de Fontenay and Gans, 2014).

[14] Here, connectedness is defined in the graph-theoretic sense: firm i is connected to firm j if there exist a chain of edges in the graph from the vertices of firms i and j.

$$\Pi_i(G) = \sum_{P \in P^N} \sum_{S \in P} (-1)^{|P|-1}(|P|-1)! \left[\frac{1}{N} - \sum_{\substack{i \notin S' \in P \\ S' \neq S}} \frac{1}{(|P|-1)(N-|S'|)} \right] \Pi_S(G^P). \tag{4.10}$$

In this expression, we have introduced some more notation from graph theory. Here, P denotes a partition of N players into non-overlapping non-empty coalitions. The set of all such partitions is denoted as P^N. The coalitions S and S' are members of P. The partitioned graph G^P contains only edges of the graph G that connect the members of the same coalition within P, but excludes links that connect members of different coalitions.[15] Finally, $\Pi_S(G^P)$ denotes the joint surplus of coalition S in the partitioned graph G^P.

Since expression (4.10) only involves coalitions which firm i is a member of, there are no inter-coalitional transfers, and that the intra-coalitional transfers add to zero. Hence, the total coalitional payoffs $\Pi_S(G^P)$ can be computed from the bilateral profit-maximizing equations (4.8) alone. The individual payoffs $\Pi_i(G)$ are then determined from the generalized Myerson-Shapley value (4.10).

Alternative payoff vector

It is important to stress that the generalized Myerson-Shapley value (4.10) is an equilibrium solution of the bargaining equations (4.8) and (4.9) if and only if the individual participation constraints (4.7) are satisfied for all edges of all subgraphs. Stole and Zwiebel (1996) term this absence of unilateral breakdowns of negotiations *feasibility*. For applications with competitive externalities, feasibility is something that will have to be explicitly verified in order to directly apply the equilibrium characterization above.

Whenever feasibility does not hold, either there is an other payoff vector in equilibrium than the generalized Myerson-Shapley value or there is no equilibrium in the given market structure. To find a potential alternative payoff vector, a modified set of recursive bargaining equations has to be solved explicitly. Let G' be a graph with a single edge $(i', j') \subset G'$ for which the individual participation constraint (4.7) is not satisfied. This means that the subgraph $G'' \equiv G' \setminus (i', j') \subset G'$ will be formed in equilibrium. In particular, G' itself cannot be a credible threat in the negotiations of an edge (i, j) in a graph G for which $G \setminus (i, j) = G'$. The replacement of an unfeasible graph G' by its largest feasible subgraph G'' has to be performed in every bargaining equation (4.9) in which G' occurs as an outside option. Moreover, the procedure of checking

[15] In effect, G^P is the set of edges in G consistent with the partition P.

feasibility and replacing unfeasible subgraphs has to be iterated until complete feasibility has been demonstrated.[16] In particular, it is possible no equilibrium candidate satisfying all the individual participation constraints will survive this iterative procedure.

4.3 Results

We present here the results of our main model, connected duopoly, and two alternative market structures, disconnected duopoly and upstream monopoly. Disconnected duopoly assumes that the two insurers contract exclusively with their hospital pair. The exclusivity of the contract becomes public in stage 0 of the game, i.e. before the bargaining game begins so that the bargaining pairs expect disconnected duopoly to arise in equilibrium. Upstream monopoly is the market structure when there is only one hospital in the hospital market and both insurers contract that single hospital. Also this is determined in stage 0.

Calculations are made in a Mathematica notebook and are available upon request.

4.3.1 Equilibrium prices

Because the outcome in the third stage of the game (realization of consumer demand for insurance) is fully determined by the results in the first two stages, we begin the backward induction at the second stage. In the second stage of our game, insurers engage in simultaneous price setting by optimizing their profit functions (4.6) with respect to p_i. Solving the corresponding first order conditions yields

$$\begin{aligned} p_i(w_i, w_{-i})|_{G_{CD}} &= \frac{1}{2}v_{broad} + \frac{1}{2}w_i - \frac{\mu^2}{2(4+\mu)(4+3\mu)}(v_{broad} - w_i) - \\ &- \frac{\mu\,(2+\mu)}{(4+\mu)(4+3\mu)}(v_{broad} - w_{-i}) \qquad (4.11) \\ p_i(w_i, w_{-i})|_{G_{DD}} &= p_i(w_i, w_{-i})|_{G_{UM}} \\ &= \frac{(4+3\mu)v_{narrow} + (2+\mu)((2+\mu)w_i + \frac{\mu}{2}w_{-i})}{(2+\frac{\mu}{2})(4+3\mu)} \qquad (4.12) \end{aligned}$$

where w_i is the average wholesale unit price that the insurer i pays to the hospital(s), i.e. $w_i = \frac{1}{2}w_{ij} + \frac{1}{2}w_{i,-j}$ in case of connected duopoly, and $w_i = w_{ij}$ in

[16] We discuss here only the case of a single edge that does not meet the individual participation constraint because this is what occurs in our specific model. If there were more unprofitable edges, there could be multiple feasible outside options in a given bargaining. That could implicate that the game has multiple equilibria.

case of disconnected duopoly and upstream monopoly. Note that $p_i(w_i, w_{-i})$ is the same in case of upstream monopoly and disconnected duopoly.

In all the three market structures, i.e. connected duopoly, disconnected duopoly and upstream monopoly, price is increasing in the marginal cost of the insurer (w_i) and in the marginal cost of the competitor (w_{-i}), as expected. Also, price (conditional on w_i) is decreasing in the degree of substitutability. Price is, ceteris paribus, the highest when the two health plans are independent from each other ($\mu \to 0$). In this case, we get the monopoly result, $p_i = \frac{v(G_i)+w_i}{2}$ in all three market structures. Price, given the wholesale prices w_{ij}, is the lowest when the health plans are perfect substitutes ($\mu \to \infty$). Then it approaches the weighted average of the marginal cost of the insurer and that of the competitor, namely $p_i = \frac{1}{3}(2w_i + w_{-i})$. These features of the premium (p_i) as function of the wholesale price (w_{ij}) are as expected from standard models of horizontal product differentiation.

Substituting the optimal prices of the second stage of the bargaining game (4.12) into the insurers' profit functions (4.6) together with the hospitals profit functions (4.5), yields a set of reduced profit functions of the first stage that depend on the average wholesale prices.

Connected duopoly has four different hospital-insurer pairs. The bargaining pairs maximize their joint surplus with respect to the variable fee w_{ij} as described by equation (4.8). In case of disconnected duopoly, each insurer bargains with a single, but not the same hospital. For example, insurer i bargains with hospital j over unit price w_{ij} by maximizing their joint surplus, and insurer $-i$ bargains with hospital $-j$ over unit price $w_{-i,-j}$. In contrast, when the market structure is upstream monopoly, the two insurers negotiate with the same hospital. For example, the two bargaining pairs are insurer i and hospital j and insurer i and hospital $-j$. The system of equations consists then of two equations.

The equilibrium wholesale unit prices, also illustrated in Figure 4.2, are:

$$w^*_{ij}|_{G_{CD}} = \frac{\mu^2}{\frac{1}{2}(16+\mu(16+5\mu))} v_{broad} \tag{4.13}$$

$$w^*_{ij}|_{G_{DD}} = \frac{2\mu^2}{(2+\mu)(16+\mu(12+\mu))} v_{narrow} \tag{4.14}$$

$$w^*_{ij}|_{G_{UM}} = \frac{\mu^2}{4(2+\mu)^2} v_{narrow} \tag{4.15}$$

When the two insurance policies are independent from each other ($\mu = 0$), the wholesale price is 0 in all three market structures, i.e. connected duopoly, disconnected duopoly and upstream monopoly, to avoid double marginalization and maximize the joint profit of the bargaining hospital-insurer pair. For low levels of μ, the wholesale unit price (w_{ij}) is increasing in substitutability (μ) in equilibrium. The wholesale unit prices of the two insurers with the same hospital, i.e. $w_{i,j}$ and $w_{-i,j}$, are strategic complements. Consequently, an increase in $w_{i,j}$ triggers an increase in $w_{-i,j}$, and similarly, decrease in $w_{i,j}$ triggers a decrease in $w_{-i,j}$. The interpretation is that when health plans are significantly differentiated, it is profitable to increase wholesale prices, knowing that the competitor will follow the strategy. However, wholesale prices will drop when substitutability on the insurance market reaches high levels and no hospitals contract both insurers. This is the case in disconnected duopoly; wholesale prices converge to the marginal costs of the hospitals when health plans are perfect substitutes, as expected. This is also clearly seen in Figure (4.2).

However, if there are hospitals contracting both insurers, cutting prices under the competitors' prices is not profitable any more. The bargaining pairs can increase their joint profit by increasing the wholesale unit prices even in case of high level of substitutability between health plans. Consequently, w_{ij} keeps growing with μ for any values of μ. When the two health plans are perfect substitutes ($\mu \to \infty$), the wholesale price approaches $w_{ij} = \frac{2v_{broad}}{5}$, which is also the average wholesale price because of the assumed symmetry in hospital shares. In the market structures connected duopoly and upstream monopoly, hospitals can serve as common agents and insurers can alleviate competition on the insurance market by contracting the same hospitals. A similar effect is shown in Gal-Or (1997) too. Insurers, when contracting a single hospital, are better off by selecting the same hospital than by contracting different hospitals in her model.

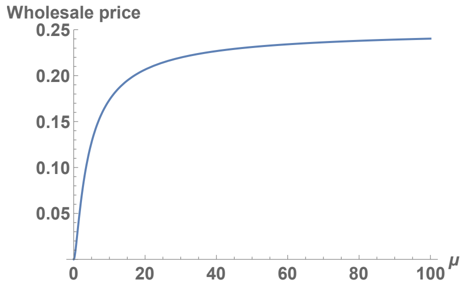
Substituting the equilibrium wholesale prices into (4.12) yields the following insurance premiums (assuming that the equilibrium network is connected duopoly, disconnected duopoly, upstream monopoly respectively for any value of μ)

$$p_i^*|_{G_{CD}} = \frac{2(4+\mu(3+\mu))}{16+\mu(16+5\mu)} v_{broad} \tag{4.16}$$

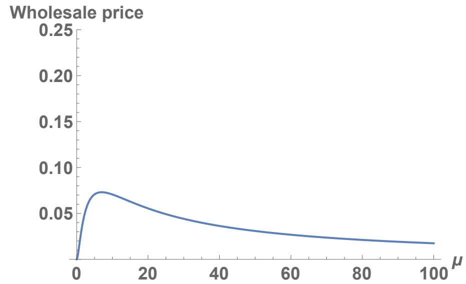
$$p_i^*|_{G_{DD}} = \frac{4(2+\mu)}{16+\mu(12+\mu)} v_{narrow} \tag{4.17}$$

$$p_i^*|_{G_{UM}} = \frac{(4+\mu)}{4(2+\mu)} v_{narrow} \tag{4.18}$$

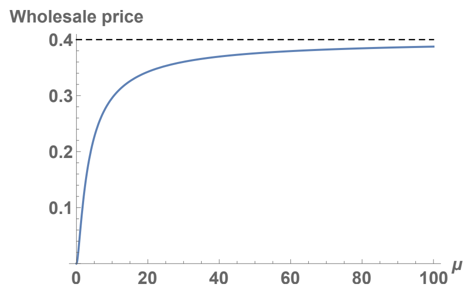
The equilibrium price is the function of substitutability (μ) and network value $v(G_i)$. However, it is also affected indirectly by the equilibrium wholesale price (w_i). As shown, equilibrium price is increasing in w_i and, in case of connected duopoly and upstream monopoly, w_i is increasing in μ in the whole parameter



(a) Upstream Monopoly



(b) Disconnected Duopoly

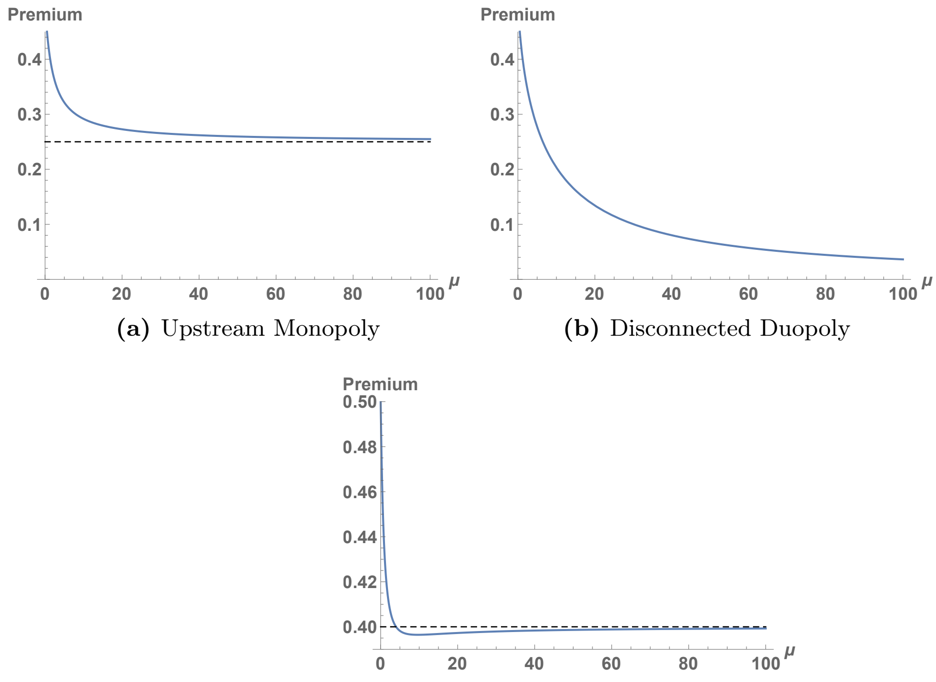


(c) Connected Duopoly

Figure 4.2: Equilibrium wholesale prices (w_{ij}) as a function of substitutability (μ) in case of connected duopoly (bottom), disconnected duopoly (top,right) and upstream monopoly (top, left). Parameters v_{broad} and v_{narrow} are set to 1. The scale of the y axis in the top and bottom row differs.

range. Consequently, there is an indirect upward pressure on p_i^* in these two market structures when substitutability (μ) increases. At the same time, the direct effect of μ on equilibrium price due to more intensive competition is a downward pressure. For low values of μ, the direct effect dominates and equilibrium price is decreasing in μ. In case of upstream monopoly, the direct effect keeps dominating for any value of μ and price keeps decreasing even though w_{ij} is monotonously increasing, as also seen in Figure (4.3). When the two health plans are perfect substitutes, the equilibrium premium converges to $\frac{1}{4}v_{narrow}$, which is above marginal cost of the hospital.

In case of connected duopoly, the equilibrium price has a minimum value of $p_i^* = \frac{3-\sqrt{2}}{4}v_{broad}$ at $\mu = 4 + 4\sqrt{2}$. After this point, growing substitutability (μ) increases price because the indirect effect of w_{ij} is strong enough to offset the downward pressure of more intensive competition. When substitutability approaches infinity, prices approach $\frac{2}{5}v_{broad}$ which is the average wholesale price.



(a) Upstream Monopoly

(b) Disconnected Duopoly

(c) Connected Duopoly

Figure 4.3: Equilibrium premium (p_i) as a function of substitutability (μ) in case of connected duopoly (bottom), disconnected duopoly (top, right) and upstream monopoly (top, left). Parameters v_{broad} and v_{narrow} are set to 1. The scale of the y axis in the top en bottom row differs.

In the market structure of disconnected duopoly, the common agency effect is not present. Equilibrium premiums are monotonously decreasing and converge to zero, when μ approaches infinity.

In order to get an intuition over the practical relevance of this study, we estimated the feasible value of μ in the Dutch health insurance market through the known actual value of price elasticity in the market. According to the estimates of Douven et al. (2017), the short term price elasticity (ϵ) was -0.9 to -2.2 in the period 2007-2015. In the year of the introduction of the new health insurance system (2006), the public attention to reforms and competition in the health insurance market yielded an all-time high price elasticity of -5.7 (Douven et al., 2017). We can calculate the price elasticity in this model from the demand function (4.1) by taking $\frac{\partial q_i}{\partial p_i}\frac{p_i}{q_i}$. In the equilibrium of connected duoply, we get the following expression

$$\epsilon = -\frac{4 + \mu(3 + \mu)}{4 + 3\mu} \tag{4.19}$$

If we substitute the estimated values of price elasticity into the expression and solve it for μ, we get a value between 0 and 4.6 for the period 2007-2015, and 15.3 for 2006. As Figure (4.3) shows, the equilibrium premium decreases sharply in the range $0 \leq \mu < 5$ and does not change significantly afterwards. The Dutch health insurance market seems to be in the range where stimulating higher degree of competition, i.e. substitutability (μ), is still rewarding. However, the fierce competition of 2006, seems to be excessive considering that μ exceeded its optimal value.[17]

The equilibrium values for the wholesale prices w_{ij} and premiums p_i determine the firms' profits modulo the lump sum transfers t_{ij}. Since these fixed fees sum to zero for every industry structure, we will not have to compute them explicitly to analyze consumer surplus and total welfare.

4.3.2 Industry profit and consumer surplus

The total industry profit is given by the sum of all firms' profits, in which all fixed fees and variable transfer payments cancel each other out.

$$\Pi(G) = \sum_i \Pi_i(G) + \sum_j \Pi_j(G) = \sum_i q_i p_i \tag{4.20}$$

The consumer surplus for a differentiated product market is most conveniently expressed through the aggregated indirect utility function

$$\mathrm{CS}(G) = U(q_i, q_{-i}) - \sum_i q_i p_i \tag{4.21}$$

Finally, total welfare is the sum of industry profit and consumer welfare.

$$\mathrm{TW}(G) = \Pi(G) + \mathrm{CS}(G) = U(G) \tag{4.22}$$

[17] In our model, μ is exogeneous, but, of course, in a real situation it can change by time. Insurers may, for example, try to differentiate their health plans to soften competition.

Substituting the equilibrium premiums into the hospitals' and insurers' profit functions yields for the industry profit

$$\Pi(G_{\mathrm{CD}}) = \frac{(4+2\mu)(4+3\mu)(4+\mu(3+\mu))}{(16+\mu(16+5\mu))^2} v_{broad}^2 \tag{4.23}$$

$$\Pi(G_{\mathrm{UM}}) = \frac{(4+\mu)(4+3\mu)}{16(2+\mu)^2} v_{narrow}^2 \tag{4.24}$$

$$\Pi(G_{\mathrm{DD}}) = \frac{4(2+\mu)(8+\mu(8+\mu))}{(16+\mu(12+\mu))^2} v_{narrow}^2 \tag{4.25}$$

Note that the industry profits divided by the total added value squared (v_{broad}^2, or v_{narrow}^2), just like consumer and total welfare in the following, depend on a single parameter: the downstream insurer differentiation μ. Similarly, we find for the consumer surplus and total welfare

$$\mathrm{CS}(G_{\mathrm{CD}}) = \frac{(2+\mu)^2(4+3\mu)^2}{2(16+\mu(16+5\mu))^2} v_{broad}^2 \tag{4.26}$$

$$\mathrm{CS}(G_{\mathrm{DD}}) = \frac{(8+\mu(8+\mu))^2}{2(16+\mu(12+\mu))^2} v_{narrow}^2 \tag{4.27}$$

$$\mathrm{CS}(G_{\mathrm{UM}}) = \frac{(4+3\mu)^2}{32(2+\mu)^2} v_{narrow}^2 \tag{4.28}$$

$$\mathrm{TW}(G_{\mathrm{CD}}) = \frac{(2+\mu)(4+3\mu)(24+\mu(22+7\mu))}{(2(16+\mu(16+5\mu))^2)} v_{broad}^2 \tag{4.29}$$

$$\mathrm{TW}(G_{\mathrm{DD}}) = \frac{(8+\mu(8+\mu))(24+\mu(16+\mu))}{2(16+\mu(12+\mu))^2} v_{narrow}^2 \tag{4.30}$$

$$\mathrm{TW}(G_{\mathrm{UM}}) = \frac{(4+3\mu)(12+5\mu)}{32(2+\mu)^2} v_{narrow}^2 \tag{4.31}$$

Industry profit, consumer welfare and total welfare are monotonous functions of μ in case of disconnected duopoly and upstream monopoly, as expected. However, the welfare functions, just like the equilibrium prices, have an optimum point in case of connected duopoly. This is interesting from the perspective of welfare.

Increasing substitutability, does not yield increasing consumer and total welfare in equilibrium in the whole range of μ. In case of complete differentiation ($\mu = 0$), consumer welfare and total welfare have their minimum at $\frac{v_{broad}}{8}$ and $\frac{3v_{broad}}{8}$ respectively. By increasing substitutability between the health plans, consumer welfare and total welfare increase in equilibrium, and at their maximum, they reach the value of $\frac{1}{32}(3+2\sqrt{2})v_{broad}^2$ and $\frac{1}{32}(5+6\sqrt{2})v_{broad}^2$ respectively. After their maximum, they decrease slightly and converge to $\frac{9v_{broad}}{50}$ and $\frac{21v_{broad}}{50}$ respectively when μ approaches infinity (perfect substitutability of health plans). Consequently, the welfare reduction between the optimum and perfect competition is relatively low compared to the welfare increase between complete differentiation and the optimum in our model.

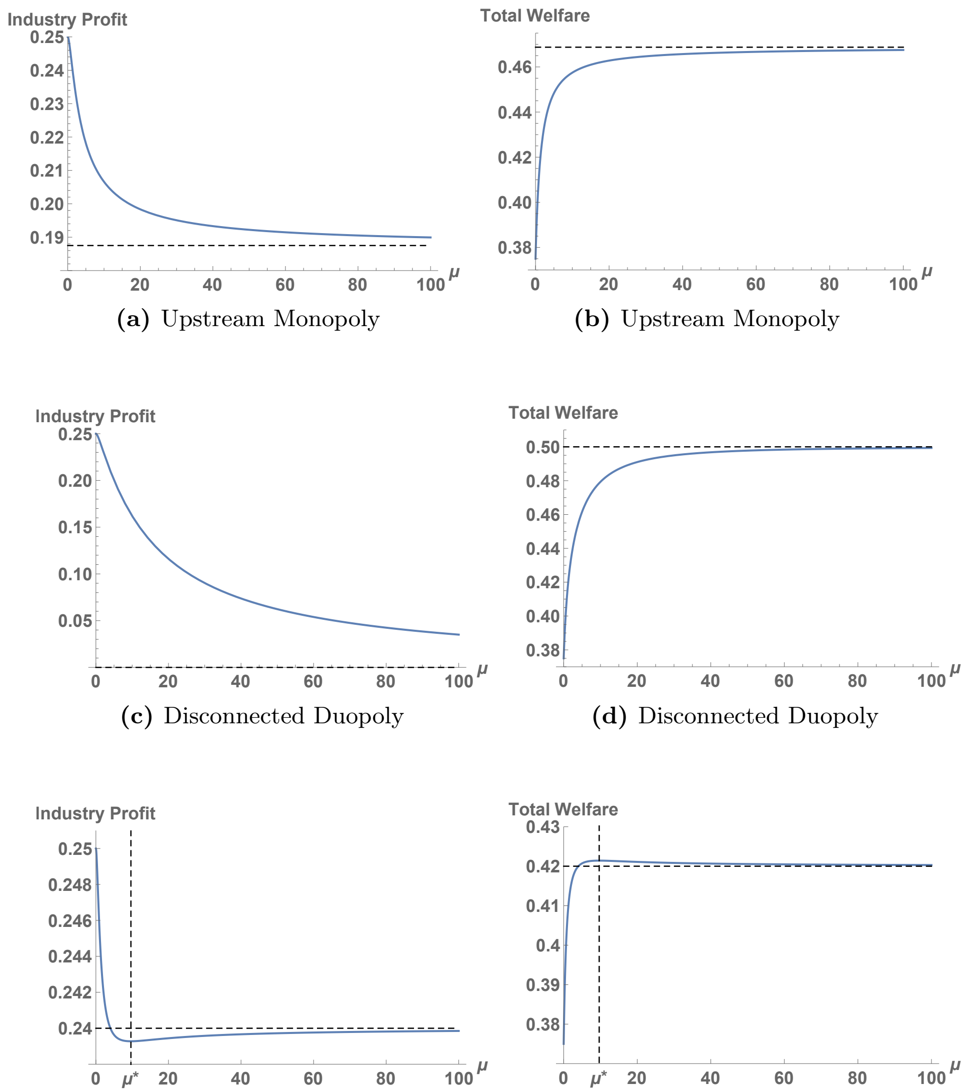
Industry profit also has its optimum at $\mu = 4 + 4\sqrt{2}$ when the industry structure is connected duopoly, which is a minimum value of $\frac{1}{16}(1 + 2\sqrt{2})v^2_{broad}$. The industry earns the highest profits ($\frac{v_{broad}}{4}$) when there is complete product differentiation ($\mu = 0$) on the health insurance market, as expected. When the health plans are perfect substitutes ($\mu \to \infty$), industry profit approaches $\frac{6v_{broad}}{25}$, which is slightly higher than its minimum value. These results are consequences of the upward pressure of w_{ij} on the equilibrium premiums, which is a revelation of common agency between insurers through contracting the same hospitals.

Figure 4.4 shows how industry profit and total welfare change with the substitutability of health plans (μ) in the equilibrium of connected duopoly, disconnected duopoly and upstream monopoly.[18] We set parameters v_{broad} and v_{narrow} to 1 so that the expressions only depend on μ. Although connected duopoly gives the lowest total welfare according Figure 4.4, this result depends strongly on the relative value of a broad network compared to a narrow network. If consumers value broad network highly, then welfare can be higher in connected duopoly than in the other market structures. However, we can conclude that substitutability of health plans affects the level of welfare less in case of connected duopoly than in case of disconnected duopoly or upstream monopoly.

To get an intuition how the relative value of broad network, compared to narrow network, affects total welfare, we show Figure (4.5). In Figure (4.5), we set $\mu = 4$ and present total welfare as a function of the relative value of broad and narrow network in case of the three different market structures. For this aim, we define the relative value of broad network as $\tau = \frac{v_{broad} - v_{narrow}}{v_{narrow}}$, and express the value of broad network as $v_{broad} = v_{narrow}(1 + \tau)$. When τ is 0, consumers do not not attach a higher value to a broad network of hospitals than to a narrow network. When $\tau = 0.1$ (the other extreem in Figure 4.5), consumers value a broad network 10% more than a narrow network.

As can be seen (Figure 4.5), total welfare in case of disconnected duopoly and upstream monopoly does not change with relative value of broad network (τ), as both insurers have a narrow contracted hospital network in these cases. Total welfare increases, in the case of connected duopoly, if we assume the relative value of a broad network to increase. This is not surprising, because insurers have a

[18]Consumer welfare function has a similar shape than the total welfare function in all three market structures.



(a) Upstream Monopoly **(b)** Upstream Monopoly

(c) Disconnected Duopoly **(d)** Disconnected Duopoly

(e) Connected Duopoly **(f)** Connected Duopoly

Figure 4.4: Industry profit (left column) and total welfare (right column) in market equilibrium as a function of substitutability (μ) in case of connected duopoly (bottom row), disconnected duopoly (middle row) and upstream monopoly (top row). Parameters v_{broad} and v_{narrow} are set to 1. The scale of the y axis in the left and right column differs.

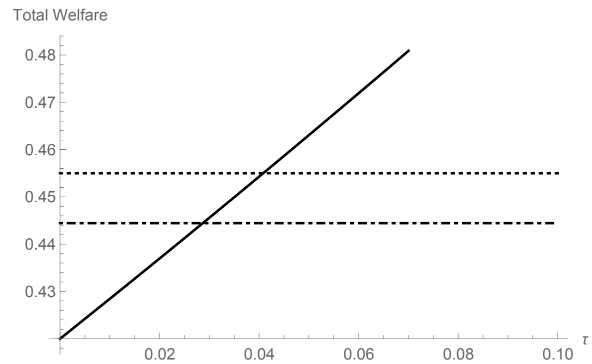


Figure 4.5: Total welfare in market equilibrium as a function of the relative value of broad network (τ) in case of connected duopoly (solid line), disconnected duopoly (dotted line) and upstream monopoly (dash-dotted line). In this figure, μ is fixed at 4 and v_{narrow} at 1.

broad network in connected duopoly; and if consumers have higher valuation for it, then, it increases welfare.

Figure 4.5 also shows, that total welfare in market structures with a narrow hospital network (disconneted duopoly and upstream monopoly) only exceeds the total welfare of connected duopoly, when consumers have a low valuation for a broad network. In a broad parameter range, connected duopoly ensures the highest total welfare among the studied market structures at a given level of substitutability.

In the following, we solve the bargaining equations to calculate profit distribution between insurers and hospitals and derive the equilibrium market structure. We show that connected duopoly is the equilibrium market structure in the whole parameter range of μ when two insurers and two hospitals are present in the market. The bargaining parties have incentives to sign all the contracts in the negotiations, and no exclusion will occur in equilibrium.

4.3.3 Profit distribution

When analysing the distribution of the industry profit between the market players, we use the generalized Myerson-Shapley value (4.10) to calculate the payoff vector of the players. We check feasibility in the graph and in all subgraphs to verify that the generalized Myerson-Shapley value is the equilibrium payoff vector. If one or more (sub)graphs are unfeasible, then we solve the bargaining equations (4.9 and 4.8) explicitly. In the equations, we replace the unfeasible outside options with their feasible subgraph to find an alternative payoff vector.

Also in this case, we check feasibility recursively to assure that the solution is the equilibrium outcome.

The generalized Myerson-Shapley value (4.10) expresses the firms' profits in the connected duopoly as a linear combination of industry profits of only three graphs

$$\Pi_j(G_{\mathrm{CD}}) = \Pi_{-j}(G_{\mathrm{CD}}) = \frac{1}{12}\left(3\Pi(G_{\mathrm{CD}}) + 2\Pi(G_{\mathrm{DM}}) - 2\Pi(G_{\mathrm{UM}})\right) \quad (4.32)$$

$$\Pi_i(G_{\mathrm{CD}}) = \Pi_{-i}(G_{\mathrm{CD}}) = \frac{1}{12}\left(3\Pi(G_{\mathrm{CD}}) - 2\Pi(G_{\mathrm{DM}}) + 2\Pi(G_{\mathrm{UM}})\right) \quad (4.33)$$

Note that the sum of all profits is equal to the industry profit. Furthermore, hospitals achieve higher profit than insurers since the downstream monopoly's industry profit is higher than (or equal to when $\mu = 0$) the upstream monopoly's profit.

To give an intuitive explanation to this last result, take the bilateral monopoly as starting point. If a second hospital enters the market and the insurer contracts with it, we get downstream monopoly. In the upstream (hospital) market, there is no competition in our model, because patients visit the hospitals in fixed proportions. Contracting a second hospital increases the industry profit because the value of the contracted network goes up from v_{narrow} to v_{broad}.

In contrast, when a new insurer enters the market, and the hospital of a bilateral monopoly contracts it, we get an upstream monopoly. On the downstream (insurance) market, there is competition in our model (except when $\mu = 0$). The entry of the second insurer, therefore, lowers the industry profit (or does not change it when $\mu = 0$). In summary, the industry profit of a bilateral monopoly is higher than that of an upstream monopoly, but lower than the industry profit of a downstream monopoly ($\Pi_{UM} \leq \Pi_{BM} < \Pi_{DM}$).

Feasibility

The generalized Myerson-Shapley value is the generating function of the equilibrium payoffs in the parameter range where all underlying graphs (i.e. market structures) are feasible. Recursively checking the feasibility of all graphs and their outside options shows that, except asymmetric duopoly, all market structures are feasible in the whole parameter range.

In case of asymmetric duopoly, besides horizontal differentiation (modelled via parameter μ), the two insurers are vertically differentiated through their network as well. Both v_{narrow} and v_{broad} occur in the industry profit. When consumers attach a high value to a broad network and the two insurance policies are substitutes, the demand of the insurer with a narrow network would become negative. Therefore, the insurer either contracts the other hospital as well or exits the market. The industry structure becomes connected duopoly in the first case, or downstream monopoly in the latter case.

The value of the substitutability parameter μ determines the maximal difference in the value of networks v_{narrow} and v_{broad} which insures that both insurers remain active:

$$\frac{v_{broad} - v_{narrow}}{v_{narrow}} \leq \frac{256 + 704\mu + 752\mu^2 + 392\mu^3 + 99\mu^4 + 9\mu^5}{64\mu + 160\mu^2 + 152\mu^3 + 68\mu^4 + 14\mu^5 + \mu^6} \tag{4.34}$$

The expression on the right hand side of inequality (4.34) approaches infinity when $\mu = 0$, which means that asymmetric duopoly exists for all relative values of hospital networks if the two health plans are independent from each other. When the two health plans are perfect substitutes ($\mu \to \infty$), the expression approaches 0.

There is an other reason why asymmetric duopoly is not always a credible threat in the bargaining: the incentive constraint of the hospital contracting both insurers is not always satisfied when negotiating about the second contract. By contracting the second insurer, the hospital can improve its bargaining position against the other insurer. At the same time, the total industry profit decreases because the entry of the second insurer to the insurance market creates competition and lowers profits. These two forces determine the parameter range where the hospital is better off to contract both insurers and so asymmetric duopoly is feasible.

Asymmetric duopoly is only feasible, if inequality (4.34) and the incentive constraint (4.7) of the hospital contracting both insurers do hold. The parameter range where asymmetric duopoly is feasible is depicted in Figure 4.6 as a coloured shape. Feasibility is important when we look at the outside options in the bargaining equations. In the negotiations of connected duopoly, asymmetric duopoly can be a credible threat only in the parameter range where it is feasible. Outside this parameter range, downstream monopoly will be the feasible outside option. Therefore, when deriving the profit distribution in connected duopoly, the outside options in the bargaining equation must be changed to downstream monopoly in the parameter range where asymmetric duopoly is not feasible. Therefore, the equilibrium payoff vector is not the Myerson-Shapley value in this range but an alternative payoff vector, which we derive in the next subsection.

As shown in Figure 4.6, asymmetric duopoly is not a credible threat if consumers have a high valuation for a broad hospital network compared to a narrow network. The intuition behind this result is that hospitals already have a strong bargaining power since consumers strongly prefer a broad network compared to a narrow network. The marginal increase in the bargaining position of the hospital cannot then outweigh the decrease in industry profit due to competition induced by the entry of the second insurer.

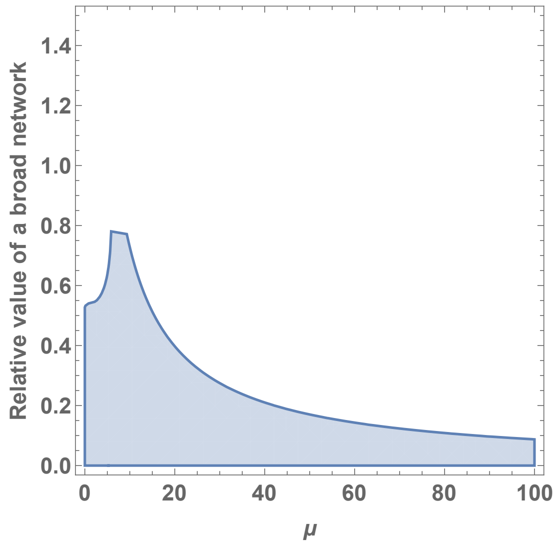


Figure 4.6: Parameter ranges of substitutability (μ) and relative value of a broad network compared to a narrow network where asymmetric duopoly is feasible (colored shape). In this parameter range, the generalized Myerson-Shapley value is the generating function of the equilibrium payoffs.

It is also seen in Figure 4.6 that the range of relative value of broad and narrow network where both insurers can have a positive demand, decreases with substitutability where the original level of substitutability is already high. In case of homogeneous products ($\mu \to \infty$), asymmetric duopoly is not a credible threat and the generalized Myerson-Shapley value is not a generating function of the equilibrium payoffs unless consumers do not value broad hospital networks at all $v_{broad} = v_{narrow}$.

Alternative profit distribution

By iteratively checking feasibility and replacing unfeasible outside options, we get an alternative set of bargaining equations. The solution of these equations gives the alternative equilibrium payoff vector:

$$\Pi_j(G_{\mathrm{CD}}) = \Pi_{-j}(G_{\mathrm{CD}}) = \frac{1}{12}\left(3\Pi(G_{\mathrm{CD}}) + 2\Pi(G_{\mathrm{DM}}) - \Pi(G_{\mathrm{BM}})\right) \quad (4.35)$$

$$\Pi_i(G_{\mathrm{CD}}) = \Pi_{-i}(G_{\mathrm{CD}}) = \frac{1}{12}\left(3\Pi(G_{\mathrm{CD}}) - 2\Pi(G_{\mathrm{DM}}) + \Pi(G_{\mathrm{BM}})\right) \quad (4.36)$$

The payoff of the hospitals is higher than in the case of the Myerson-Shapley value because $2\Pi(G_{\mathrm{UM}}) > \Pi(G_{\mathrm{BM}})$, which is $2\frac{(4+\mu)(4+3\mu)}{16(2+\mu)^2}v_{narrow}^2 > \frac{1}{4}v_{narrow}^2$. This reflects the higher bargaining power of the hospitals when asymmetric duopoly is not a credible threat in the bargaining.

In summary, the connected duopoly is feasible for all values of substitutability μ and relative consumer preference for broad network compared to narrow network. Changes in these parameters do not alter the equilibrium industry structure. The two parameters do affect the level of industry profit that hospitals and insurers together gain and the distribution of this profit among the market players.

4.4 Discussion

In our analyses, we focused on the competition aspect of decreasing product differentiation. However, there are other important aspects too, which we do not try to model. First, differentiation may provide better match between health plans and consumer preferences, which ceteris paribus increases welfare. Second, differentiation may induce adverse selection according to the healthcare risk of consumers, which decreases welfare. Third, differentiation may increase the complexity of health plan choice, which results in suboptimal choices of consumers. The net effect of the three above mentioned consequences of differentiation is ambiguous (e.g. Ericson and Sydnor (2017); Handel et al. (2015) discusses these issues).

As a side result, we have shown that insurers contract both hospitals at all levels of substitutability in our model, and so connected duopoly is the equilibrium market structure. This finding is in line with the results in the relatively few theoretical papers in health economics that discuss exclusive dealing in a bilateral oligopoly context. Gaynor and Ma (1996) find that neither insurers nor hospitals have individual incentives for exclusive dealing. The model of Gaynor and Ma consists of two homogeneous insurers, which would corresponds to $\mu \to \infty$ in our model, and two differentiated hospitals. In their bargaining game, insurers offer contracts to the hospitals that accept or reject these contract proposals.

Also, Capps et al. (2003) show that in their framework, there is no exclusive dealing by insurers as long as there are no large cost or quality differences between hospitals and as long as consumer willingness to pay for ex-post hospital choice is homogeneous. However, they do not model the insurance market, and so insurers and hospitals do not consider market conditions in the insurance market during the bargaining. Gal-Or (1997) finds that exclusion of a hospital occurs in equilibrium when hospitals are less differentiated then insurers. Nevertheless, both hospitals will be contracted by both insurers when the level of hospital differentiation exceeds the level of differentiation between insurers. The exclusionary equilibrium follows from the assumption on unit price. When an insurer contracts a single hospital, it can bargain volume discounts in return for exclusivity, which favours the insurer. In contrast, an exclusive contact to a hospital weakens the bargaining position of the insurer in our model.

Finally, our results on the equilibrium market structure are also in line with the result of Douven et al. (2014). The bargaining game coincides in both papers, but we deviate from their fixed demand (defining a downward sloping demand function) and their Hotelling model of insurer and hospital differentiation (assuming that consumers have a preference for free hospital choice before falling ill, but visit hospitals in fixed proportions after falling ill, and defining differentiation parameter μ in the demand function of insurers). Also, applying a downward sloping demand curve allow us to use a fully general two-part price rather than restricting the variable reimbursement to an exogenously regulated tariff. In Douven et al. (2014) connected duopoly is the equilibrium market structure as long as the industry profit of downstream monopoly is less than twice as high as the industry profit of connected duopoly and hospital differentiation does not exceed a certain level compared to insurer differentiation. These two conditions hold in our model and we see indeed the same result regarding the equilibrium market structure.

In our theoretical model, we applied the bargaining framework of de Fontenay and Gans (2014), while recent empirical work often applies the Nash-in-Nash model (Ho and Lee, 2017; Gowrisankaran et al., 2015). In a Nash-in-Nash model, insurer and hospital pairs engage in a Nash-bargaining independently from each other and assuming agreement by the other pairs (like in the Nash-equilibrium). Contracts are binding and do not depend on the realized market structure, i.e. the contracts do not change even if an other pair fails to reach agreement. This contrasts our model where the break down of a contract triggers the renegotiation of all other contracts. Models based on Nash-in-Nash bargaining approach are empirically tractable, but as Collard-Wexler et al. (2017) shows, the solution is not necessarily a sub-game perfect equilibrium. In situations when the Nash-in-Nash solution is not sub-game perfect, the outside-options in the bargaining are not credible threats.

One can also think of the Nash-in-Nash approach as a representation of bounded rationality, while our model assumes full rationality. In Nash-in-Nash bargaining, parties can induce what would be the equilibrium if they broke the contract (outside option) but they are not capable to induce the equilibria further (e.g. what would happen if also some one else broke a contract). Dranove et al. (2011) argue that it is not reasonable to assume that bargaining parties will pursue the induction to the lowest level. The number of bargaining equations that has to be solved simultaneously to calculate a perfect Bayesian equilibrium increases indeed exponentially with the number of parties in the market. At the same time, we see that insurer and hospital markets are getting more and more concentrated. When there are only a few (dominant) hospitals and insurers in the market (as it is often the case in the Netherlands), it is acceptable to assume that the bargaining parties take all possible network outcomes into account.

We finally stress some limitations of our model and indicate directions for further research. Our analysis discusses the case of a bilateral duopoly. We assume a specific downward sloping demand curve in the insurance market, and do not explicitly model the competition in the hospital market. It would be interesting to analyze a general demand function in future research so that general conditions can be identified under which welfare is decreasing in the degree of substitutability. An other extension of our model could be the explicit modelling of hospital competition, which would allow to incorporate co-payments for consumers to steer their hospital choice or to loosen the assumption of symmetric hospitals. Finally, further research would be needed to analyze the case when hospitals treat the uninsured people as well, which would mean a second market for the hospitals outside the insurance market.

4.5 Conclusions

With our theoretical bilateral duopoly model, we have shown that total welfare and consumer welfare may decrease when substitutability between health insurance plans increases. The unconventional effect of decreasing differentiation follows from the bargaining between insurers and hospitals. Contracting the same hospital(s) allow insurers to coordinate their prices (just like in common agency and strategic delegation models) and alleviate competition on the insurance market raising prices above the competitive level. This upward pressure on prices may outweigh the downward pressure of increased competition due to less differentiation. Consequently, it is not always welfare enhancing to increase substitutability because higher levels of competition not necessarily yield better market outcomes.

We have also shown that insurers contract both hospitals at all levels of substitutability in our model, and so connected duopoly is the equilibrium market structure. At a given level of substitutability, it is also the connected duopoly that ensures the highest level of total welfare in a broad parameter range in our model. Translating this to policy, it means that free contracting between insurers and hospitals probably yields broad contracted hospital networks, and that also ensures higher welfare than an HMO-like insurance market. Only in a small parameter range, where consumers value a broad contracted hospital network only a slightly higher than a narrow contracted network, is an alternative market structure, like HMOs, desirable from social perspective.

Acknowledgements

We would like to thank Jan Boone, Cédric Argenton, Martin Gaynor, Johan Polder, Marco Varkevisser and Michiel Bijlsma for their valuable comments and suggestions.

Bibliography

Bakx, P., O'Donnell, O., and Doorslaer, E. Spending on health care in the netherlands: not going so dutch. *Fiscal Studies*, 37(3-4):593–625, 2016.

Bernheim, B. and Whinston, M. Common agency. *Econometrica*, 54(4):923–942, 1986.

Bijlsma, M., Boone, J., and Zwart, G. Competition leverage: how the demand side affects optimal risk adjustment. *The RAND Journal of Economics*, 45(4): 792–815, 2014.

Binmore, K., Rubinstein, A., and Wolinsky, A. The Nash bargaining solution in economic modelling. *The RAND Journal of Economics*, 17(2):176–188, 1986.

Capps, C., Dranove, D., and Satterthwaite, M. Competition and market power in option demand markets. *The RAND Journal of Economics*, 34(4):737–763, 2003.

Chernev, A., Böckenholt, U., and Goodman, J. Choice overload: A conceptual review and meta-analysis. *Journal of Consumer Psychology*, 2(25):333–358, 2015.

Collard-Wexler, A., Gowrisankaran, G., Arizona, H., and Lee, N. R. S. "nash-in-nash" bargaining: A microfoundation for applied work. 2017.

Corts, K. S. and Neher, D. V. Credible delegation. *European Economic Review*, 47(3):395–407, 2003.

de Fontenay, C. C. and Gans, J. S. Bilateral bargaining with externalities. *The Journal of Industrial Economics*, 62(4):756–788, 2014.

Douven, R., Halbersma, R., Katona, K., and Shestalova, V. Vertical integration and exclusive behavior of insurers and hospitals. *Journal of Economics & Management Strategy*, 23(2):344–368, 2014.

Douven, R., Katona, K., Schut, F. T., and Shestalova, V. Switching gains and health plan price elasticities: 20 years of managed competition reforms in the netherlands. *The European Journal of Health Economics*, pages 1–18, 2017.

Dranove, D., Satterthwaite, M., and Sfekas, A. Bargaining and leverage in option demand markets. 2011.

Ericson, K. M. M. and Starc, A. How product standardization affects choice: Evidence from the massachusetts health insurance exchange. *Journal of health economics*, 50:71–85, 2016.

Ericson, K. M. and Sydnor, J. The questionable value of having a choice of levels of health insurance coverage. *Journal of Economic Perspectives*, 31(4):51–72, 2017.

Fanti, L. Cross-ownership and unions in a cournot duopoly: When profits reduce with horizontal product differentiation. *Japan and the World Economy*, 27:34–40, 2013.

Gal-Or, E. Exclusionary equilibria in health-care markets. *Journal of Economics & Management Strategy*, 6(1):5–43, 1997.

Gal-Or, E. A common agency with incomplete information. *The RAND Journal of Economics*, pages 274–286, 1991.

Gaynor, M. and Ma, C. Insurance, Vertical Restraints, and Competition. *GSIA Working Papers*, 1996.

Gaynor, M., Ho, K., and Town, R. J. The industrial organization of health-care markets. *Journal of Economic Literature*, 53(2):235–84, 2015.

Gowrisankaran, G., Nevo, A., and Town, R. Mergers when prices are negotiated: Evidence from the hospital industry. *The American Economic Review*, 105(1): 172–203, 2015.

Halbersma, R. and Katona, K. Vertical restraints in health care markets. *TILEC Discussion Paper No. 2011-005*, 2011.

Handel, B. R., Kolstad, J. T., and Spinnewijn, J. Information frictions and adverse selection: Policy interventions in health insurance markets. Technical report, National Bureau of Economic Research, 2015.

Ho, K. and Lee, R. S. Insurer competition in health care markets. *Econometrica*, 85(2):379–417, 2017.

Kopel, M. and Pezzino, M. Strategic delegation in oligopoly. In Corchón, L. C. and Marini, M. A., editors, *Handbook of Game Theory and Industial Organization: Volume II, Applications*, chapter 10, pages 248–285. Edward Elgar Publishing, Cheltenham, UK, 2018.

Loewenstein, G., Friedman, J. Y., McGill, B., Ahmad, S., Linck, S., Sinkula, S., Beshears, J., Choi, J. J., Kolstad, J., Laibson, D., et al. Consumers' misunderstanding of health insurance. *Journal of Health Economics*, 32(5): 850–862, 2013.

NCHS, N. C. f. H. S. *Health, United States, 2016: With Chartbook on Long-term Trends in Health.* National Center for Health Statistics, Hyattsville, MD. 2017., 2017.

Schut, F. T. and Varkevisser, M. Competition policy for health care provision in the netherlands. *Health Policy*, 121(2):126–133, 2017.

Shubik, M. and Levitan, R. Market structure and behaviour. *Harvard Uni*, 1980.

Stole, L. and Zwiebel, J. Intra-firm bargaining under non-binding contracts. *The Review of Economic Studies*, 63(3):375–410, 1996.

Zanchettin, P. Differentiated duopoly with asymmetric costs. *Journal of Economics & Management Strategy*, 15(4):999–1015, 2006.

Chapter 5

Vertical Integration and Exclusive Behavior of Insurers and Hospitals

Rudy Douven, Rein Halbersma, Katalin Katona, Victoria Shestalova

This chapter was published as:
Douven, R., Halbersma, R., Katona, K., and Shestalova, V. Vertical integration and exclusive behavior of insurers and hospitals. *Journal of Economics & Management Strategy*, 23(2): 344–368, 2014

Abstract

We examine vertical integration and exclusive behavior in health care markets in which insurers and hospitals bilaterally bargain over contracts. We employ a bargaining model of two hospitals and two health insurers competing on premiums. We show that asymmetric equilibria exist in which one insurer contracts one hospital while the other insurer contracts both hospitals, even if all players are equally efficient in their production. Asymmetric equilibria arise if hospitals are sufficiently differentiated. In these cases, total industry profits increase and consumer welfare decreases in comparison to the case in which both insurers have contracts with both hospitals. Vertical integration makes these equilibria possible for a wider range of parameters.

JEL Classification Numbers: G22, G34, I11, L14, L42
Keywords: insurer-provider networks, vertical integration, exclusive behavior

5.1 Introduction

One of the main challenges in health care is to reduce costs by providing health care more efficiently. A market-oriented approach has been followed in the United States and also recently in Europe, an approach in which competing health insurers and health care providers should in theory achieve an efficient allocation of the production and consumption of health care. This approach may stimulate the appearance of new forms of institutional and contractual arrangements in the health-care sector. For example, the growth of managed care in the United States has led to tighter vertical relationships between health insurers and providers in the form of health maintenance organizations (HMO). In these organizations insurers' enrollees typically receive full reimbursement for services from providers within the network, while sometimes facing co-payments when visiting providers outside the network. Vertical restraints or integration can be a tool for insurers and providers to gain efficiency, but they may also have anticompetitive effects. The potential consequences of vertical integration between an insurer and a hospital are currently cause for concern in Europe. For instance, the question of these consequences has been discussed by the Dutch Parliament, and the current government in the Netherlands now plans a per se prohibition of vertical integration between an insurer and a hospital.[1]

In this paper, we examine under which market conditions exclusive behavior and vertical integration can arise and whether this may harm competition. We study this question with a bilateral-duopoly model of a competitive health care market, a model in which insurers and hospitals bilaterally bargain over contracts.[2] We show that two types of exclusive behavior can occur in this setting. First, one of the insurers may be excluded. This happens if decreasing competition (i.e., monopolization) in the insurer market leads to a substantial increase in total industry profits. Second, one hospital may not, in equilibrium, engage in a contract with one of the insurers. In such a case, a managed-care insurer (contracting one hospital) and a conventional insurer (contracting both hospitals) can coexist in the market, even if both insurers and both hospitals are equally efficient in their production. We show that the range of parameters under which the latter outcome can occur grows if one insurer-hospital pair integrates vertically. The modeled result that both managed-care and conventional health plans can coexist in the market is new, since earlier literature on this topic (Gal-Or (1997) and Ma (1997)) focused on exclusionary outcomes. Our model draws from Gal-Or (1997) but incorporates the bargaining concept developed by

[1] Source: The coalition agreement of the Dutch government of September 30, 2010 (http://www.rijksoverheid.nl/regering/het-kabinet/regeerakkoord).

[2] In the remainder of the paper we will use the term hospitals, since we primarily focus on this subject, but a more general term, such as health care providers, would be applicable as well.

de Fontenay and Gans (2005). The key difference with Gal-Or (1997) is that both hospitals and insurers are modeled as strategic players; they bargain over profit allocation, rather than over linear contracts.

The paper proceeds as follows. In Section 2, we describe the literature and our contribution to it. We develop our model in Section 3, which we use in Section 4 to characterize equilibrium outcomes in the cases with and without vertical integration of a hospital and an insurer. Section 5 provides the analysis of consumer welfare. In the last section, we draw conclusions and outline directions for further research. We relegate to the appendix more technical details concerning the bargaining approach used.

5.2 Literature and contribution

Managed-care organizations use various forms of vertical arrangement to reduce the cost of providing health care and to improve the quality of care. There is evidence that a vertically integrated network may enhance consumer welfare by providing health care more efficiently. Many studies on the United States have shown that insurance provided by managed-care organizations cost 10 to 20 percent less than indemnity insurance. Whether these cost reductions comprise all efficiency savings is, however, still unclear; some of these cost reductions may also be related to a lower quality of care or the selection of low-risk enrollees (Getzen Thomas, 2007). There is also a welfare loss associated with vertically integrated networks. On the demand side, two recent empirical papers of Capps et al. (2003) and Ho (2006) report welfare losses from restriction in provider choice. On the supply side, welfare losses are associated with strategic behavior of insurers and providers. Gaynor and Vogt (2000) provide an overview.

The anticompetitive effects of horizontal mergers are well-known, but the literature regarding vertical relationships is less developed. In their overview, Bijlsma et al. (2008) argue that vertical relationships may result in anticompetitive foreclosure of competitors, but only in the presence of market power in the insurance and/or hospital market. Recently, Ho (2009) provides empirical evidence that, in the United States, market power of hospitals is sometimes responsible for vertical restraints. Some hospitals may demand high prices that not all insurers are willing to pay.

In Europe, these issues play a role in countries with market-oriented approaches to health care, such as the Netherlands and Switzerland. In the Netherlands, there was a political debate regarding a per se prohibition of vertical mergers between hospitals and health insurers. The final report by an independent commission concluded that a ban on vertical integration was not necessary and that antitrust policy should assess intended vertical mergers case by case (Baarsma et al., 2009).

A related subject in health care concerns the effects on welfare of vertical arrangements within the health care provision chain, for example, between physicians and hospitals. Here, too, the empirical literature is mixed. For example, Cuellar and Gertler (2006) find that in many U.S. markets in which managed care grew rapidly, hospitals and physicians integrated, raising hospital prices. Ciliberto and Dranove (2006) find no evidence, however, in California that vertical activity of hospitals and physicians led to significant changes in hospital prices.

There are a few theoretical papers in the health economics literature that study exclusive contracting and vertical integration between insurers and hospitals in a duopoly setting. Important contributions are papers by Gaynor and Ma (1996); Ma (1997); Gal-Or (1997, 1999). Gaynor and Ma (1996) study exclusive dealing in a model of two homogeneous insurers and two differentiated hospitals. They assume a situation in which insurers can grant an exclusive contract to a single hospital to treat all insured enrollees. Gaynor and Ma find that neither insurers nor hospitals have individual incentives for this type of exclusive dealing. If customer foreclosure of the non-contracted hospital did occur, however, the reduced choice would be detrimental to consumer surplus.

Gal-Or (1997) studies a bargaining model of two insurers and two hospitals. Both insurers and hospitals are differentiated along Hotelling lines. In the downstream market, two insurers simultaneously choose the hospital networks that they contract, as well as the premiums for the associated insurance policies. For each pair of insurer strategies in the downstream market, the insurers' profits are determined through simultaneous, bilateral Nash bargaining between the various hospital-insurer pairs. Even though hospitals and insurers are treated symmetrically in each bargaining subgame, only insurers are strategic players that can optimize between the various bargaining subgames. Hospitals, in contrast, must take the insurers' choices of network and premium as given. Gal-Or finds that an exclusionary outcome (foreclosure of a hospital) can arise in equilibrium, because insurers are able to receive larger price discounts if they contract a single hospital. This behavior is profitable to insurers over a small range of parameters in which hospital differentiation is much smaller than insurer differentiation. In this exclusionary outcome, consumers are better off, because insurers obtain a favorable price by offering exclusivity to one hospital and partly transfer these gains to consumers. In a subsequent paper, Gal-Or (1999) extends her model to arbitrary numbers of hospitals and insurers located on Salop circles, largely confirming the analysis of the bilateral duopoly case.

Ma (1997) analyzes vertical integration in a model of two homogeneous insurers and two differentiated hospitals, similar to the model of Gaynor and Ma

(1996). He demonstrates that a vertical merger can result in foreclosure of the competing insurer. Such foreclosure can subsequently lead to downstream monopolization, in which case Ma (1997) shows that the effect on consumer welfare is ambiguous.

As discussed above, both Gal-Or (1997) and Ma (1997) focus on exclusionary outcomes, when one of the firms exits the market. In contrast, in our paper, we show that both a managed-care health plan (contracting with only one hospital) and a conventional insurer plan (contracting with both hospitals) can coexist in the market. Our model draws from Gal-Or (1997), but incorporates the bargaining concept developed by de Fontenay and Gans (2005).

The model designed by Gal-Or provides us with a starting point, since our interest lies in a model that will be applicable to situations with fixed total demand for health insurance, such as under the condition of mandatory insurance. This set-up features a fixed population distributed on a Hotelling line. Mandatory insurance, in combination with a large basic benefit package, is applied in many European countries.[3] The United States still has many uninsured individuals, but the proposed health reform by the Obama administration is intended to provide insurance to those who currently do not have it.[4] The key difference from Gal-Or (1997) is that both hospitals and insurers are seen as strategic players; they bargain over profit allocation, rather than over linear contracts. We implement these features by using a more advanced bargaining concept, recently developed by de Fontenay and Gans (2005, 2007). It treats both insurers and hospitals as strategic players and is suitable for modelling non-cooperative bargaining among multiple parties, resulting in a unique Bayesian-Nash equilibrium. The bargaining concept deviates from the assumption of strictly linear contracts between insurers and hospitals, allowing for the use of a two-part tariff.

A two-part tariff consists of a per-unit price and a fixed fee. Such combinations are observed in regulated hospital markets,[5] but they are not directly observed in liberalized markets, for which hospital-insurer reimbursement claims typically feature linear prices. Therefore, earlier theoretical studies assume lin-

[3]A model with these features is applied in many European countries, for example, in Germany, Switzerland, and the Netherlands. Also, in the United States, the Medicare HMO plans for citizens over 65 years of age represent an example of such a system. In health economics terms, this type of arrangement represents a model of managed competition with community-rated premiums, open enrollment, and a risk-adjustment system.

[4]Recently, two papers appeared that concern vertical restraints in a voluntary insurance setting. In Bijlsma et al. (2008), exclusive contracting of hospitals by insurers raises the costs of self-insurance by consumers. In Halbersma and Katona (2011), the total demand for treatments in hospitals is not fixed. Higher prices in their model raise the number of non-insured individuals and lower hospital demand.

[5]For example, this reimbursement scheme used to be applied in the Netherlands under the old budgetary system for hospitals.

ear contracts in insurer-hospital bargaining (Gaynor and Ma, 1996; Gal-Or, 1997, 1999). However, empirical evidence on the presence of quantity discounts (e.g., Wu, 2009; Melnick et al., 1992) corroborates the assumption that insurer-hospital contracts are not strictly linear, but rather that the negotiated average prices depend on purchased quantities. Furthermore, the recent empirical contribution by Wu (2009) shows that patient channeling adds to the volume discount that health plans negotiate with hospitals; hence, plans that are more successful at channeling patients extract greater discounts. This supports the idea that an insurer and a hospital bargain over the division of their profits, rather than over the per-unit price itself.[6]

One way to reconcile our assumed two-part tariffs with the observably linear claim-reimbursement data is a scheme in which insurer-hospital contracts specify a two-part tariff, but in which the fixed fee is amortized over the contracted (expected) number of patients. Effectively, a two-part tariff with an amortized fixed fee is equivalent to a volume discount over an estimated number of patients. Such volume discounts are not explicitly observed in claims data, but the implicit relationship could be estimated from the observed prices and volumes.

Apart from the bargaining game, our model design (specification of consumer preferences and hospital costs) otherwise closely follows Gal-Or (1997).[7] The incentive structure in our model is similar to that of Gal-Or (1997) and Ma (1997). Both Gal-Or and Ma find symmetric equilibria where one downstream (hospital) or upstream (insurer) firm is excluded. In contrast, we find asymmetric equilibria in which one insurer contracts both hospitals while the other insurer contracts one hospital. Both insurers and hospitals gain from the asymmetry because it creates product differentiation in the insurance market, which yields additional profits in the total market. In one equilibrium, a (vertically integrated) hospital-insurer pair increases their profits by not allowing enrollees of the competitive insurer to visit their hospital. The competing insurer cannot engage an exclusive arrangement with the other hospital, because this hospital would thereby lose market share and obtain less profit. We find that the total industry profit is higher in this equilibrium than when both insurers contract both hospitals. The asymmetric equilibrium leads to lower consumer welfare; it raises total premium payments and restricts hospital choice for consumers.

[6] While linear contracts can be seen as one extreme, another extreme is a pure *capitation* contract that specifies only the amount of the transfer payment from the insurer to the hospital and no quantity. Such a contract has been recently concluded between Dutch insurer Menzis and the Haga-hospital located in the Hague. Source: Skipr Magazine, 6 December 2011: "Menzis contracteert Haga met vaste 'aanneemsom'" by Philip van de Poel. See http://www.skipr.nl/actueel/id9416-menzis-contracteert-haga-met-vaste-aanneemsom.html.

[7] Also note that neither model incorporates moral hazard or selection effects. Moral hazard and selection effects are prominent in any health care market. Relaxing one of these above assumptions would make our model richer, but also more complex.

5.3 Model and the bargaining game

The set-up of our model of the health-care market is similar to that considered by Gal-Or (1997). For the bargaining game, we follow de Fontenay and Gans (2005, 2007), who designed a bilateral bargaining framework that is specifically suitable for situations with externalities. Our setup includes externalities resulting from horizontal competition between insurers and hospitals respectively, as well as externalities from hospital network size. In this section, we describe the model and define the bargaining outcome for the case without vertical integration.

5.3.1 General setup

In our model, two health insurers and two hospitals serve a certain population. We assume that each health insurer and each hospital negotiate bilaterally about the amount that the insurer will pay to the hospital for providing health care to its enrollees. If the negotiation succeeds, the hospital joins the insurer's network. After the networks have been established, the insurers compete for individuals by setting a uniform insurance premium for health insurance that fully covers the care from the respective network.[8]

We designate the insurers by I_1 and I_2 and assume them to be located at the end points of a Hotelling line of unit length. The population is distributed uniformly on the line between the two insurers, with transportation cost parameter M reflecting the degree of differentiation between insurers. Consumers know their location $y \in [0, 1]$ at the downstream Hotelling line before buying insurance. The reason why we assume the insurers to be horizontally differentiated ex ante is as follows. In practice, consumers often obtain care through collective contracts of their employers, and employers may prefer one insurer over another as a result of additional services offered by that insurer, or as a result of switching costs. In addition, individuals' perceived switching costs, such as status quo bias, may cause horizontal differentiation in the insurer market. The presence of horizontal differentiation is consistent with the low cross-price elasticities of demand for health insurance reported by many studies (e.g. Douven et al., 2007; Strombom et al., 2002).

After buying insurance, a consumer falls ill with fixed probability θ ($0 < \theta < 1$) and learns his illness. In the upstream market, two hospitals, H_A and H_B, provide medical services. To ease exposition, we consider a symmetric case in which both hospitals have zero fixed costs and constant marginal costs, and hence the same average treatment cost $c_A = c_B = c$.[9] Although both hospitals

[8]Hence, we assume that the individuals do not pay any copayments for receiving care. However, including copayments would not change the results.

[9]The model can be also generalized for asymmetric costs.

are able to treat all types of diseases, they are differentiated in their effectiveness of treating different diseases. We model the differentiation between hospitals by using the Hotelling model. The hospitals are located at the end points of the upstream Hotelling line of unit length. Patients are uniformly distributed between them with the transportation cost parameter t, which reflects the degree of ex-post differentiation between hospitals. The transportation parameter t is larger, when there are large differences between hospitals in their effectiveness for treating different diseases. The location of a patient on the upstream Hotelling line is denoted by $x \in [0, 1]$. While the downstream location y is known ex-ante (i.e. at the moment of buying health insurance), the upstream location x is revealed only later, after the consumer has fallen ill. Therefore, the product that consumers buy is essentially a bundle of options for access to the hospitals contracted by an insurer (Capps et al., 2003).

Both transportation parameters, M and t, are considered fixed in the short run, during which contracts are concluded and service delivery takes place. However, they can change in the long run. For example, in the long run, hospitals can extend or narrow their specialization ,which changes the level of their differentiation, insurers can adjust their policies, which influences the switching cost of consumers, etc. In our paper, we do not model such multi-period extensions of our single-shot game.

The two hospitals and the two insurers bargain bilaterally over contracts. In the case of a successful negotiation, the hospital enters insurer I_i's network, $G_i \subseteq \{H_A, H_B\}$. The network of the two insurers can overlap. Individuals who buy insurance from insurer I_i have access only to the hospitals that are in network G_i. Insurers engage in price competition in the downstream market by setting insurance premiums F_i for access to the respective insurance network G_i. If $G_i = \varnothing$, insurer I_i has no contract with any hospital, i.e., insurer I_i exits the market. Then consumers can buy health insurance only from the other insurer. In the latter case, we assume that the remaining insurer cannot set monopoly prices, but a regulator will cap the premium to a maximum of $\overline{F}$ that guarantees a certain minimum expected utility level to consumers. We assume the regulation to be light-handed in the sense that the regulated monopoly premium cannot be less than the equilibrium premium of insurance duopoly (which will be derived in Section 3.4). The assumption that monopoly premiums will be regulated at an affordable level is realistic and ensures that our model is still applicable for the analysis of markets with mandatory insurance, in which the complete population needs to be covered by health insurance.

5.3.2 Consumer preferences

We specify consumer preferences by the same indirect utility function as in Gal-Or (1997). The individual ex-ante expected indirect utility depends on the insurance premium, the hospital network to which the individual will have access, and consumer location y at the Hotelling line between the two insurers. In particular, a consumer who buys health insurance from insurer I_i (offering access to network G_i at price F_i) derives an ex-ante expected indirect utility of:

$$U_i = \theta(v - T(G_i)) - (F_i + My_i). \tag{5.1}$$

The first term represents the ex-ante expected indirect utility from the treatment in a hospital from network G_i. Here v is a fixed parameter, and $T(G_i)$ is the Hotelling's transportation cost to the network G_i, which we will derive below. The last term reflects the insurance premium F_i and the transportation cost My_i between the consumer and insurer I_i, where $y_1 = y$ and $y_2 = 1 - y$.

As explained, the individual falls ill with fixed probability θ, after which he learns his location x on the Hotelling line between the two hospitals. Therefore, the individual's ex-post transportation costs to hospitals H_A and H_B are expressed as tx and $t(1-x)$ respectively. If network G_i includes both hospitals, then the individual will select the one which is closer to x (i.e. hospital H_A if $x \in [0, \frac{1}{2}]$ and H_B if $x \in [\frac{1}{2}, 1]$), whereas if network G_i consists of only one hospital, there is no choice and the individual goes to that hospital. Therefore the expected transportation costs $T(G_i)$ in expression (5.1) are the following:

$$\begin{aligned} T(\{H_A, H_B\}) &= t\int_0^{1/2} x\mathrm{d}x + t\int_{1/2}^1 (1-x)\mathrm{d}x = \frac{t}{4}, \\ T(\{H_A\}) = T(\{H_B\}) &= t\int_0^1 x\mathrm{d}x = \frac{t}{2}. \end{aligned}$$

We assume that v is a fixed value and that it is sufficiently large, so that the ex-ante expected utility in equilibrium will always be positive. Our model is now fully determined by the exogenous parameters M, t, θ, c, and $\overline{F}$.

5.3.3 Timing

When modeling the bargaining process and the payoff allocation, we use the bargaining framework of de Fontenay and Gans (2005) and de Fontenay and Gans (2007). The timeline of the game consists of a single exogenous stage followed by two endogenous stages:

- **Stage 0:** The ownership of assets and the set of allowed negotiations among the four players are fixed exogenously.[10]
- **Stage 1:** The bargaining takes placeunder the given ownership structure. The equilibrium network of contracts is established as (a) a subset of the exogenously allowed negotiations and (b) a corresponding payoff allocation.
- **Stage 2:** Insurers offer insurance policies to consumers with a premium and unrestricted access to the contracted hospitals. Consumers choose an insurer. After this they may fall ill with probability θ and then choose the closest hospital from their insurer's hospital network, and receive treatment. Insurers transfer the respective payments to hospitals.

Throughout this article, we consider only vertical contractual relationships, in the sense that insurers buy services produced by hospitals (i.e., they do not buy insurer services produced by the other insurers), while hospitals do not buy services of other hospitals. At stage 0, it is decided once for all which bilateral contracts are allowed. In this paper, we always assume that all the four bilateral contracts are initially allowed.[11] The asset allocation process is modelled as exogenous. In Section 4.1, we assume that each firm is owned by a different owner. In Section 4.2, we consider the case in which one insurer-hospital pair integrates, therefore, one of them obtains the control over both the hospital and the insurance firm. Treating the ownership structure as exogenous, we compare the joint profits of a hospital-insurer pair both with and without integration to analyze the incentive to integrate. This partial analysis allows us to conclude whether vertical integration is profitable, and what type of integration is likely to occur. We then check the robustness of this partial analysis by considering all possible equilibrium outcomes and showing that a countermerger of the other two parties will be infeasible in equilibrium.

5.3.4 Total industry profits

In this section, we consider all potential alternative configurations of insurer-hospital contractual relationships that may arise in this model and derive total industry profits for each configuration. Since each insurer's network may include none, one, or both hospitals, there are fifteen different (not empty) configurations

[10] The set of allowed negotiations must be consistent with the ownership structure. In particular, in case of independent firms owned by their managers, an insurance firm's manager negotiates with a hospital's manager. However, if one hospital-insurer pair integrates, all negotiations of this firm will be done by the owner of the integrated firm. We will elaborate on this further in section 4.

[11] However, some of these allowed contracts may turn out to be infeasible and will not be concluded in equilibrium. This case is different from the case of naked exclusion (not considered in this paper), in which certain bilateral contracts are not allowed from the beginning.

possible, shown in Appendix 1. Six configurations correspond to an insurer monopoly in which one insurer is out of business and the other insurer contracts one or two hospitals. Nine networks represent duopoly cases, in which both insurers contract either one or two hospitals. Due to symmetry, some of these networks result in the same total profits. Hence, there are only six different configurations to consider; they correspond to different rows in Table 5.A.1 from Appendix 1.

We start with the insurer duopoly case, in which insurer I_1 contracts network G_1 and insurer I_2 contracts network G_2. In a Hotelling setup insurer I_i's demand is determined by the marginal consumer who is indifferent between the two insurers. It is easy to show that insurer I_i's demand, q_i, is expressed as:

$$q_i(F_i, F_{-i}|G_i, G_{-i}) = \frac{1}{2} + \frac{F_{-i} - F_i}{2M} + \frac{\theta\left[T(G_{-i}) - T(G_i)\right]}{2M}.$$

Here the labeling $-i$ denotes the other insurer. The expression is symmetric for both insurers. Insurers can increase demand by lowering their premium F or by decreasing Hotelling's transportation cost to their network $T(G)$. We consider efficient bargaining (this assumption will be discussed in Section 3.5), which means that insurers internalize hospital production costs in their pricing decisions. Therefore, the profit-maximizing insurers maximize their revenue minus production cost. Since the marginal costs of both hospitals are assumed to be the same[12], this yields the following condition on optimal duopoly premiums F_i^*:

$$F_i^* = \arg\max_{F_i} q_i(F_i, F_{-i}|G_i, G_{-i})(F_i - \theta c)$$

A straightforward calculation yields that in any fully symmetric duopoly configuration, the premiums are $M + \theta c$ and the total industry profits are M. The intuition behind this result is that in a symmetric case the indifferent consumer in the insurance market is located in the middle of the Hotelling line and total industry profits are equal to the degree of insurer differentiation M, as in the standard Hotelling model (Tirole, 1988). For an asymmetric duopoly networks, in which one insurer enters into a contract with one hospital and the other one with two hospitals, our model generates the respective insurers' premiums $\theta c + M - \frac{\theta t}{12}$ and $\theta c + M + \frac{\theta t}{12}$. The total profits equal $M + \frac{(\theta t/12)^2}{M}$. Since the insurer that contracts both hospitals becomes more attractive to consumers, the indifferent consumer in the insurance market is now located closer to the insurer with a smaller network. Therefore, the insurer that contracts both hospitals is

[12] The model can be generalized to the case of asymmetric costs. For a given network state, hospital H_j's demand q_{ij} from the insured by insurer I_i equals $q_{ij} = \theta q_i(F_i, F_{-i}|G_i, G_{-i})s_j(G_i)$, where $s_j(G_i)$ is the expected share of consumers at hospital H_j from insurer I_i. and the condition on optimal insurance premiums F_i^* becomes $F_i^* = \arg\max_{F_i} q_i(F_i, F_{-i}|G_i, G_{-i})(F_i - \theta \sum_{j \in G_i} c_j s_j(G_i))$

able to charge a higher premium, which results in higher total industry profits. The additional profits relate positively to the degree of hospital differentiation, t (as compared to the degree of insurer differentiation M), and to the probability of contracting illness, θ.

Next, we consider insurer monopoly configurations. With only one insurer being present in the market, say insurer I_i, the demand function and the optimal premium are the following:

$$q_i(F_i|G_i) = \frac{\theta(v - T(G_i)) - F_i}{M},$$
$$F_i^* = \arg\max_{F_i} q_i(F_i|G_i)(F_i - \theta c).$$

As explained in Section 3.1, to avoid the possibility that the health insurer monopoly rations demand, we assume in this case that a regulator caps the premium level. Therefore, the regulated monopoly premiums for one- and two-hospital networks are $\overline{F}$, and the corresponding profits are $\Pi = \overline{F} - \theta c$. The regulation is light-handed in the sense that the regulated premium cannot be less than the minimum premium earned in a duopoly setting, i.e., $\Pi \geq M$. The latter assumption agrees with the general insight from industrial organization models that prices are lower in less concentrated markets. Note that for the sake of simplicity we assume that total industry profits in the insurer monopoly case do not depend on the number of hospitals in the insurer's network. This assumption matches with the outcome obtained for symmetric insurer duopoly cases, in which less hospital choice in both insurers's policies does not reduce total industry profits either. However, our results will still hold if we assume that the regulatory cap depends on the number of hospitals contracted by the insurer.

5.3.5 Bargaining and payoff allocation

Several recent studies have applied cooperative game theory concepts to determine the payoff allocation among players in the context of bilateral bargaining. Most of these studies adopt an axiomatic approach to derive the equilibrium payoffs in the form of the Shapley value (Stole and Zwiebel, 1996; Inderst and Wey, 2003). De de Fontenay and Gans (2005, 2007) have shown that the cooperative Shapley-value outcomes can also be derived from a non-cooperative Bayesian-Nash game with fully specified (passive) beliefs for all players; and the model can be extended to the case with externalities in the downstream market. The payoff allocation in the extended model is expressed by the generalized Myerson-Shapley value.[13] Similarly to the Shapley value, the generalized Myerson-Shapley

[13] The generalized Myerson-Shapley value extends the Shapley-value concept to games in which a coalition value depends on other players' actions. In particular, the generalized Myerson-Shapley-value

value can be justified on axiomatic grounds, as well as can be obtained from an extensive form game. For the sake of simplicity, we adopt the axiomatic approach here. See de Fontenay and Gans (2005) and de Fontenay and Gans (2007) for a description of the respective extensive form game and all formal assumptions supporting it.

The axiomatic approach is based on the general result of de Fontenay and Gans (2007) that the individual payoffs resulting from the bargaining satisfy the axioms of bilateral efficiency and fairness. Here bilateral efficiency means that - for each hospital-insurer pair - the intermediate tariffs maximize the joint surplus of the two players; fairness means that the net surplus derived from each bilateral relationship is split equally between the two players. Next to bilateral efficiency and fairness, the equilibrium allocation must be feasible, i.e. satisfy the participation constraints for both players. Since the relationship is only feasible when it is profitable, a contract will only be concluded if it generates a non-negative net surplus. The feasibility conditions are analyzed in more detail in the next section.

The main difference with de Fontenay and Gans (2005, 2007) is that we consider downstream price competition while in their paper, downstream firms compete on quantities. In particular, we assume that hospitals and insurers negotiate two-part tariff contracts: insurers pay a lump sum amount and a price per treatment. (The relevance of this assumption has been discussed in Section 2). Furthermore, we impose that the per-unit price is set equal to the hospital's marginal cost of providing the treatment, c.[14] This restriction on the per-unit price is comparable to a regulatory regime, in which the regulator requires that the per-unit price must cover hospital production costs, while allowing the firms to negotiate freely over the transfer payment from the insurer to the hospital. In such a case the insurers will internalize the production cost when choosing their optimal policies. Therefore, under this regulatory restriction, bilateral efficiency implies overall efficiency.

The application of the efficiency and fairness axioms leads to the system of bargaining equations (Myerson, 1977a):

$$\Phi_i(K) - \Phi_i(K \backslash ij) = \Phi_j(K) - \Phi_j(K \backslash ij)$$

$$\sum_i \Phi_i(K) = v(K) \text{ for all } K \subseteq G;\ (ij) \in K;\ i, j = A, B, 1, 2$$

reflects the average marginal contribution of the player to various coalitions when the coalition value may depend on the players' partition into coalitions.

[14]This restriction is mainly chosen to a) simplify the analysis for the symmetric graph where both insurers contract both hospitals, and b) to focus on the incentives for exclusive behavior.

where K runs over all the possible network configurations of the complete graph G (see Appendix 1); (ij) runs over all possible contracts within each configuration; $v(K)$ denotes the total profit of all players in configuration K; $\Phi_j(K)$ and $\Phi_j(K \backslash ij)$ are the payoffs of firm j both with and without the contract (ij). The first equation expresses the fairness of each bilateral bargaining relationship: both parties have an equal amount to gain from contract (ij). The second equation says that all the payoffs sum up to the total industry profits, taking into account the competitive externalities that the firms impose on each other.

The resulting payoffs are obtained by solving the equations recursively for each configuration. We start with the simplest configuration (in which there are no contracts, which yields every player a zero payoff), after which we consider configurations with only one contract using the previous configuration as the outside option. We continue in this way, until we reach the most complete configuration of four contracts. The solution is unique and expressed by the generalized Myerson-Shapley value (Myerson, 1977a,b). The resulting generalized Myerson-Shapley value expressions are provided in Appendix 2.

Table 5.1 summarizes the results of the recursive computations for the benchmark case, in which there is no vertical integration between firms. Since both the production costs and the consumer preferences are symmetric, we need to consider only six different configurations. The first column of the table shows all the different configurations possible. The second column shows the respective total industry profits. They depend on the industry configuration, and can take three values: Π, M, and $M + \Delta M$, where $\Delta M = \frac{(\theta t/12)^2}{M}$. The third column presents the resulting individual payoffs, Φ_i, computed by solving the bargaining equations recursively. The fourth column gives the feasibility conditions per configuration that are assumed to hold. These feasibility conditions will be derived in the next section.

Several general results follow from the outcomes in Table 5.1:

- A hospital or an insurer with more links in a given network obtains higher profit than its counterpart with less links since establishing more contracts improves the bargaining position of the firm by improving its outside option.
- Only in an insurer duopoly with an asymmetric network, the hospital differentiation parameter t enters the profit expression. The reason is that, in symmetric networks, hospital choice by a consumer is independent of its insurer choice.
- In an insurer duopoly with two hospitals in both insurers' networks, an increase in Π increases hospital profits but decreases insurer profits. The reason for this is the (attractive) outside option of hospitals to eliminate one of the insurers.

- In an insurer duopoly with an asymmetric hospital networks, additional profit ΔM arises because of the increased differentiation of insurer policies. The bargaining game allocates the additional profits $\Delta M = \frac{(\theta t/12)^2}{M}$ equally among the four players. Therefore, both total and individual profits increase in this case if consumers perceive relatively more differences between hospitals (higher t) than between insurers (lower M).

5.4 Feasibility conditions: incentive for exclusive behavior

In addition to bilateral efficiency and fairness, the bargaining outcome also needs to be feasible. An industry configuration is feasible, if and only if each firm benefits from signing each contract under each possible subconfiguration, i.e., the participation constraints are satisfied for both contracting parties. Before the negotiations start, the complete set of allowed links is exogenously determined by the ownership structure. The bargaining equilibrium can however result in a subnetwork of the full network of allowed links if not every individual contract is mutually profitable. The equilibrium outcome of the game in that case is an exclusive network. In this section, we derive conditions for these exclusivity outcomes. We formulate conditions under which all links will be retained in equilibrium ('the complete industry equilibrium'), and under which conditions some players will break links ('exclusive industry equilibrium'). We consider two ownership structures: no vertical integration (Section 4.1) and vertical integration (Section 4.2).

5.4.1 Benchmark: no vertical integration

We derive the feasibility conditions for each industry configuration shown in Table 5.1 under the assumption that each firm is owned by its asset manager, i.e., there is no vertical integration. We denote the complete configuration (shown in the last row of Table 5.1) by G, and any part of this configuration by $K \subseteq G$. Then the feasibility condition implies that for all K, any two firms i and j that have a relationship in K should benefit from this relationship:

$$\Phi_i(K) \geq \Phi_i(K \backslash (ij)) \text{ and } \Phi_j(K) \geq \Phi_j(K \backslash (ij)), \text{ where } K \subseteq G,\ (ij) \in K. \quad (5.2)$$

When some of these conditions are not satisfied, the complete configuration is infeasible, and one or more insurer-hospital pair(s) have not concluded a contract. We will call this situation 'exclusive behavior'.

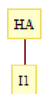
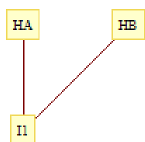
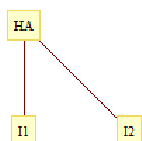
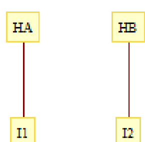
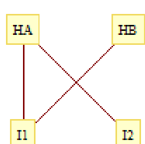
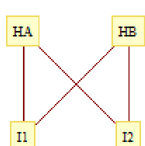
Network configuration	Total profit $\widehat{v}$	Individual payoffs	Feasibility conditions
	$\Pi = \overline{F} - \theta c$	$\Phi_A = \frac{1}{2}\Pi$ $\Phi_1 = \frac{1}{2}\Pi$	$\Pi \geq 0$
	$\Pi = \overline{F} - \theta c$	$\Phi_A = \Phi_B = \frac{1}{6}\Pi$ $\Phi_1 = \frac{2}{3}\Pi$	$\Pi \geq 0$
	M	$\Phi_A = \frac{1}{3}(M + \Pi)$ $\Phi_1 = \Phi_2 = \frac{1}{6}(2M - \Pi)$	$\Pi \geq 0$ $2M \geq \Pi$
	M	$\Phi_A = \Phi_B = \frac{1}{4}M$ $\Phi_1 = \Phi_2 = \frac{1}{4}M$	$M \geq 0$
	$M + \Delta M =$ $M + \frac{(\theta t/12)^2}{M}$	$\Phi_A = \frac{1}{12}(4M + 3\Delta M)$ $\Phi_B = \frac{1}{12}(3\Delta M + 2\Pi)$ $\Phi_1 = \frac{1}{12}(4M + 3\Delta M)$ $\Phi_2 = \frac{1}{12}(4M + 3\Delta M - 2\Pi)$	$\Pi \geq 0$ $2M \geq \Pi$
	M	$\Phi_A = \Phi_B = \frac{1}{12}(M + 2\Pi)$ $\Phi_1 = \Phi_2 = \frac{1}{12}(5M - 2\Pi)$	$\Pi \geq 0$ $2M \geq \Pi$ $\frac{\theta t}{M} \leq 4\sqrt{3}$

Table 5.1: Total industry profits and individual payoffs under the respective feasibility conditions

The feasibility conditions shown in the fourth column of Table 5.1 have been derived using the payoff allocations shown in the third column of the same table. These conditions need to be derived for all subgraphs $K \subseteq G$ and all links. For example, suppose that all the subgraphs that are included in G, $K \subset G$, are feasible and show how to derive the last feasibility condition for $K = G$ (this condition is shown in the last row of Table 5.1). From symmetry, it is sufficient to check the feasibility condition for one link of this graph, say link between H_B and I_2. Moreover, since the two firms share the net surplus equally, it is sufficient to check the condition for one firm only, say for hospital H_B. Link $(H_B I_2)$ is only profitable for hospital H_B if $\Phi_B(G) \geq \Phi_B(G \backslash (H_B I_2))$. Substituting the payoff values from the third column of Table 5.1, we obtain $\frac{1}{12}(M + 2\Pi) \geq \frac{1}{12}(3\Delta M + 2\Pi)$, which reduces to $M \geq 3\Delta M$ (equivalent to $\frac{\theta t}{M} \leq 4\sqrt{3}$). Applying this procedure to all subgraphs $K \subset G$ and all links, we derive all the other feasibility conditions shown in Table 5.1.

Proposition 1 *As long as conditions (5.3)-(5.4) hold, the bargaining game results is a unique equilibrium in which both insurers contract both hospitals.*

$$\begin{aligned} 2M &\geq \Pi, && (5.3) \\ \frac{\theta t}{M} &\leq 4\sqrt{3}, && (5.4) \end{aligned}$$

A violation of the first condition leads to elimination of one insurer in equilibrium. If the first condition holds, while the second condition is violated, the number of contractual relationships decreases by one.

The proof of Proposition 1 follows directly after eliminating the overlapping feasibility conditions and filling in the value $\Delta M = \frac{(\theta t/12)^2}{M}$. If one of these conditions is not satisfied then at least one player has no incentive to enter into a contract and foreclosure occurs. The equilibrium configuration then contains less than four contracts or one of the insurers exits the market. Foreclosure may occur for two reasons in equilibrium.

First, if feasibility condition (5.3) is violated, it is individually rational for hospitals to contract with a single insurer and to eliminate the other insurer from the market. This is because in such a case the hospitals' loss from weakening their outside options is compensated through the gains from the increased total industry profit because of the monopolized insurance market. The exclusion occurs when monopoly profits $\Pi = \overline{F} - \theta c$ are at least twice as high as duopoly profits M.

Second, if only condition (5.4) is violated, then it is not profitable anymore to establish a contract for the last bargaining pair. The equilibrium is a three-link configuration, in which one insurer contracts both hospitals and the other insurer contracts one hospital. In the insurer market, the indifferent consumer is no longer located in the middle of the Hotelling line but closer to the insurer which contracts only one hospital. The insurer with two hospitals exploits the fact that the consumers value choice, and raises the insurance premium. This generates extra industry profit that are allocated among all the players in such a way that the profit for the hospital and insurer with one contract are higher than they would be with an additional contract. Thus, the fourth relationship does not arise. Note that each of the four possible three-link networks is equally likely to occur due to our symmetry assumption. All the four firms are still active in the market in this case.

The outcome that one hospital is fully excluded never occurs in equilibrium in our model. Each insurer has as incentive to deviate from such an outcome, because contracting more hospitals increases its profit by improving its bargaining position vis-a-vis each hospital. This is an important insight that contrasts with the result of Gal-Or (1997). She considers a similar model, but a different bargaining procedure featuring linear contracts between hospitals and insurers. In her model, if $\frac{\theta t}{M}$ is sufficiently small, both insurers have an incentive to contract one hospital and exclude the other one from the market, because "if a payer chooses to exclude one of the hospitals from its approved list, its bargaining position vis-a-vis the remaining hospital is improved, since this hospital may be willing to accept lower reimbursement rates in return for a larger volume of patients that such an exclusion guarantees" (Gal-Or, 1997, p.6). However, since in our case the contract is non-linear, the bargaining position of an insurer always deteriorates when a hospital is excluded from its network.

5.4.2 Vertical integration

We will now focus on the complete industry configuration (satisfying the feasibility conditions (5.3)-(5.4)), and analyze whether vertical integration may create incentives for exclusive behavior.[15] In our analysis of vertical integration, we adopt the approach of de Fontenay and Gans (2005). It presents a specific view on vertical integration, assuming that the four managers of the firms' assets have essential skills, and therefore, integration does not eliminate the need to negotiate with those managers. In such a case, there may be two types of integration between a hospital and an insurer, depending on who becomes the owner of the integrated firm: forward (FI) and backward integration (BI). Under FI, the hospital takes over the insurer and becomes the owner of the integrated firm, while the insurance firm's manager becomes an employee. Under BI, the reverse happens and the insurer takes over the hospital.

[15] The cases in which some of the conditions (5.3)-(5.4) do not hold are less interesting, since foreclosure would occur anyway.

The decision on integration is assumed to be exogenous. This decision immediately affects the allocation of the asset ownership rights in stage 0 of the game. From now on, the owner of the integrated firm will conduct all the external negotiations of the integrated firm with other firms, while the manager can only negotiate with the owner of his firm. Therefore, the owner considers three contractual links: with the manager of the taken-over firm and with both independent firms; while the manager of the taken-over firm considers only the link to the owner. The other two firms consider the links to the owner and to each other. The analysis of the feasibility conditions for the equilibrium outcome shows under which circumstances the integrated firm has incentives for exclusion.

For example, let us consider FI in more detail. Since both hospitals and both insurers are symmetric, it does not matter which pair integrates. For the sake of clarity, we suppose that integration occurs between hospital H_A and insurer I_1. Table 5.2 (column FI, full graph) shows the new configuration of contractual relationships after hospital H_A takes over insurer I_1. When H_A acquires the ownership rights, it takes over the premium-setting decision of its insurer I_1 and represents this insurer in negotiations with the other hospital. Insurer I_1 receives a transfer payment from hospital H_A for managing the insurance firm, while all the profits accrue to H_A. This changes the graph because insurer I_1 can no longer negotiate with hospital H_B directly, but it does it via hospital H_A. Therefore, in the event of a breakdown in negotiation between hospital H_A and insurer I_1, not only the link between insurer I_1 and hospital H_A would be broken, but no arrangements would occur between insurer I_1 and hospital H_B either. The essential difference with the benchmark case (Section 4.1) is that under forward integration a breakdown between hospital H_A and insurer I_1 has a deeper impact because after such a breakdown insurer I_1 would exit the market, while without integration insurer I_1 would still be able to send its enrollees to hospital H_B.

When backward integration takes place between insurer I_1 and hospital H_A, the logic is similar, but the insurer gets all the ownership rights. The graph changes as shown in Table 5.2 (column BI, full graph) and all the profit of the integrated firm accrues to the insurer.

If the full graphs of BI and FI are feasible, both insurers are able and willing to buy from both hospitals. Solving bargaining equations for this cases results in payoff allocation shown in Table 5.2 (full graph). The next row in Table 5.2 shows the outcome corresponding to the case in which the feasibility conditions do not hold. As we will show, the owner in this case forecloses its competitor by not negotiating with him. In the remainder of this section we will analyze the feasibility of the full graph, i.e., whether the integrated firm has incentive to foreclose its competitor.

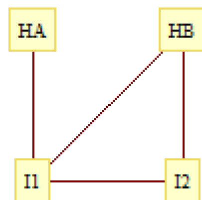
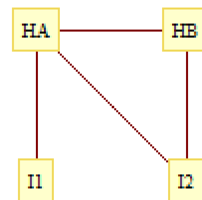
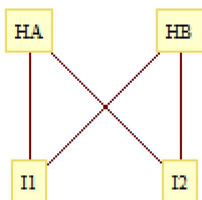
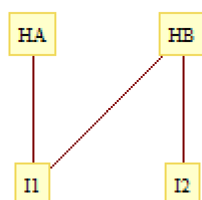
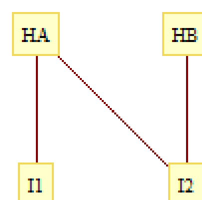
NI: No integration	BI: Backward integration	FI: Forward integration
	Full graph	Full graph
		
	$\Phi_A=\frac{1}{12}(M+\Pi)$ $\Phi_B=\frac{1}{12}(M+3\Pi)$ $\Phi_1=\frac{1}{12}(5M-\Pi)$ $\Phi_2=\frac{1}{12}(5M-3\Pi)$	$\Phi_A=\frac{1}{12}(4M)$ $\Phi_B=\frac{1}{12}(2\Pi)$ $\Phi_1=\frac{1}{12}(4M-2\Pi)$ $\Phi_2=\frac{1}{12}(4M)$
$\Phi_A=\frac{1}{12}(M+2\Pi)$ $\Phi_B=\frac{1}{12}(M+2\Pi)$ $\Phi_1=\frac{1}{12}(5M-2\Pi)$ $\Phi_2=\frac{1}{12}(5M-2\Pi)$	Foreclosure	Foreclosure
		
	$\Phi_A=\frac{1}{12}(3\Delta M+2\Pi)$ $\Phi_B=\frac{1}{12}(4M+3\Delta M)$ $\Phi_1=\frac{1}{12}(4M+3\Delta M)$ $\Phi_2=\frac{1}{12}(4M+3\Delta M-2\Pi)$	$\Phi_A=\frac{1}{12}(4M+3\Delta M)$ $\Phi_B=\frac{1}{12}(3\Delta M+2\Pi)$ $\Phi_1=\frac{1}{12}(4M+3\Delta M-2\Pi)$ $\Phi_2=\frac{1}{12}(4M+3\Delta M)$

Table 5.2: Payoffs with vertical integration (valid only if the respective graphs are feasible)

We start with the analysis of BI. The feasibility conditions derived for this case show that the full graph for BI is never feasible for $t > 0$[16]. Various links may be unprofitable. In particular, link I_1I_2 is unprofitable to insurer I_1 (the owner of the integrated firm), because his payoff without this link is larger than

[16] In particular, suppose that $2M \geq \Pi$, so that it is feasible to have two insurers. Then link I_1I_2 becomes unfeasible if either $M \geq 3\Delta M$ or $M + 3\Delta M \geq \Pi$. However, there are also other links that are infeasible on the BI graph. Detailed computations of all payoffs and all the feasibility conditions on all subgraphs of BI and FI (not included here) are available upon request.

with the link: $\frac{1}{12}(5M - \Pi) < \frac{1}{12}(4M + 3\Delta M)$. This occurs for two reasons. Firstly, without link I_1I_2, insurer I_2 will not be able to provide its enrollees access to hospital H_A. This increases differentiation in the insurer market, thereby raising the total industry profit. Secondly, the insurer I_1's profit share increases, because it attracts more consumers in the insurance market. Since link I_1I_2 is unprofitable, vertically integrated insurer I_1 will prefer not to negotiate with I_2. In other words, backward vertical integration will result in foreclosure of the other insurer. This equilibrium is stable in the sense that insurer I_2 cannot undertake a countermerger with hospital H_B and foreclose its competitor, because this strategy would be unattractive to hospital H_B, which would then lose patients from the vertically integrated insurer-hospital pair. Note also that the joint profit of the vertically integrated pair in this case exceeds their profit in the non-integrated case: $\frac{1}{6}(3\Delta M + \Pi + 2M)) > \frac{M}{2}$ (if either $\Pi > M$ or $\Delta M > 0$). This implies that BI in combination with foreclosure of the competing insurer is profitable for any $t > 0$.

The feasibility conditions for the complete graph for FI can be expressed as $2M \geq \Pi$ and $0 \geq \Delta M$. Therefore, here again, the complete graph is only feasible if there is no differentiation between hospitals ($t = 0$). Our interest, however, lies in the case of differentiated hospitals ($t > 0$). In that case, the full graph collapses. The most restrictive feasibility condition arises for the link H_AH_B. It can be seen from Table 5.2 (Column FI), that the link H_AH_B is unprofitable since the payoff is higher without this link: $\Phi_A + \Phi_B = \frac{1}{12}(4M + 6\Delta M + 2\Pi) > \frac{1}{12}(4M + 2\Pi)$. This shows that the owner of the vertically integrated firm, hospital H_A, will not negotiate with hospital H_B,.to ensure that the enrollees of insurer I_1 will only visit H_A.This strategy creates product differentiation at the insurer market and raises total industry profits. Also the profit allocation changes, because the integrated insurer-hospital pair will gain more customers in the hospital market, but it will loose customers in the insurance market. Under FI (and exclusive strategy), the profits are $\Phi_A + \Phi_1 = \frac{1}{12}(8M + 6\Delta M - 2\Pi)$; while without FI, the profits are $\Phi_A + \Phi_1 = \frac{1}{2}M$. The integrated-firm profits are larger only if $\Delta M > \frac{\Pi - M}{3}$. This implies that the profits from product differentiation, ΔM, need to be sufficiently high. Therefore, this type of integration is profitable only under a smaller range of parameters.

There are no additional gains if both insurers vertically integrate with a hospital. Vertical integration is only profitable if it also adds asymmetry in the hospital or insurer markets. If both insurer-hospital pairs integrate, the symmetry is restored; therefore, total industry profits decrease and nobody yields additional profits.

We conclude that vertical integration creates circumstances under which at least one pair of firms does not reach a contractual agreement, leading to an incomplete graph. In the case of FI, the hospital, as the owner of the vertically integrated firm, will ensure that the enrollees of its insurer will not visit the competing hospital. While in the case of BI, the insurer, as the owner of the vertically integrated firm, will prevent the enrollees of the competing insurer to obtain access to its own hospital. Once we take into account that the owner will break the link, BI is always profitable, but FI is profitable only under the condition: $M + 3\Delta M > \Pi$. This leads to the following proposition.

Proposition 2 *As long as hospital products are not perfect substitutes, regulated monopoly profits are at least as large as duopoly profits, and feasibility conditions (5.3)-(5.4) hold, vertical integration will always result in an incomplete industry configuration. FI will be accompanied by preventing the own enrollees from accessing the competing hospital, while BI will be accompanied by preventing the enrollees of the competing insurer from accessing the hospital of the vertically integrated pair. Once we take into account that the owner will break the link, BI is always profitable for the pair; and FI is profitable only if the profits from product differentiation are sufficiently high ($M + 3\Delta M > \Pi$).*

5.4.3 Discussion of results on foreclosure

Our results show that vertically integrated hospital-insurer firms can coexist with independent insurers and hospitals. This outcome differs from Gal-Or (1997) and Ma (1997), who also consider the case of two insurers and two hospitals, but obtain monopolization in one of the markets.

In the model of Ma (1997), insurers buy services from hospitals, bundle hospital products with their own service and offer consumers an insurance policy. The insurance policy can be either a simple contract (offering access to one hospital only, at a fixed price), or an option contract (giving the possibility to choose between hospitals and specifying the respective prices that the consumer will pay). Similarly to our model, if one insurer-hospital pair, say I_1H_A, vertically integrates, it always stops selling services of hospital H_A to the competing insurer. Therefore, the second insurer will only be able to offer consumers a simple contract. However, in contrast to our model, there nobody buys such a contract in equilibrium, which leads to monopolization of the insurance market. The main reason for this difference with our model is that, in Ma (1997), the insurers themselves are not differentiated (except for their networks). Since consumers do not derive any additional value from having a particular insurer, but they value the option to choose their hospital, a simple contract is less attractive. Since both insurers buy services of hospital H_B at the same price, and insurer I_1 is able to differentiate hospital prices in its insurance policy, this insurer always wins all the customers. Therefore, only one insurer will remain in the market.

In the model of Gal-Or (1997), the outcome is that one hospital exits the market. Her model is similar to ours, except for the bargaining part. There insurers are 'first movers': they offer hospitals linear contracts and bargain with them about the price. Gal-Or (1997) derives the condition under which it is attractive to both insurers to offer a contract to the same hospital and exclude the other hospital. The reason is that the remaining hospital is charging a lower price for obtaining more volume. This outcome does not arise in our model, because in our case, by excluding one hospital, insurers would increase the bargaining power of the remaining hospital, which would reduce their own profits. Besides, they would also eliminate the possibility to extract extra profits by differentiating their networks.

Our result on foreclosure under vertical integration differs from the result of de Fontenay and Gans (2005), which they derive under the assumption of perfect symmetry and substitutability upstream and downstream. They show that under this assumption, foreclosure after forward integration of H_A with I_1 is immaterial in the sense that the "foreclosed" downstream firm I_2 can still receive a non-zero payment for no services rendered. The reason for this is that even though I_2 plays no actual productive role, it does provide the integrated firm owner (in its internal negotiations under FI) an outside option in case of a bargaining breakdown with I_1. Thus, while there is technical foreclosure in terms of the elimination of downstream competition, H_A still cedes rents to I_2 so as to improve its bargaining position with respect to I_1's manager. In our model with differentiated downstream firms, such outcomes cannot occur, however. This excludes the possibility of "not talking" to firms while still keeping the option of contracting them within the future.[17]

5.5 Consumer welfare

Consumer welfare depends on the number of hospitals in each insurer's network and the premiums paid to the insurers. Note, that total industry profits and consumer allocation between firms depend only on the final industry configuration. For example, it is unimportant whether a particular network state arises spontaneously or as a result of vertical integration. As explained, three configurations are possible in equilibrium, depending on parameters' values: (i) an insurer monopoly with two hospitals; (ii) an insurer duopoly with asymmetric networks; and (iii) an insurer duopoly with both hospitals in each network. Table 5.3 provides information about premiums and consumer surplus for these three configurations.

[17] To encompass such possibilities, our model would have to be extended to a multi-period dynamic framework. We consider this beyond the scope of this current paper.

Consumer welfare is the highest in the duopoly case with both hospitals in both networks. The reason is that consumers value both free hospital choice and lower premiums. If both insurers' networks include both hospitals, all the consumers can visit their preferred hospital and buy a contract at their preferable insurer. Therefore, Hotelling transportation costs are minimized in both markets. Furthermore, competition in the insurance market is the strongest in a duopoly in which both insurers have symmetric networks. Therefore, total industry profits, and thus the average insurance premiums, are also the lowest.[18] Our result that consumer welfare is the highest when both insurers contract both hospitals contrasts with the result of Gal-Or (1997), where consumer welfare is higher in an exclusive equilibrium.[19] Note, however, that the exclusive outcome that arises in Gal-Or (1997) is characterized by the exclusion of a hospital. Such exclusive equilibria do not arise in our model.

In order to compare the two remaining equilibrium outcomes (ii) and (iii) to each other, we need to compare the average premium and transportation costs. The average premium is higher in the case of monopoly, but the total transportation cost may be lower under an insurer monopoly with two hospitals, depending on model parameters. Since we have assumed that the regulator caps the monopoly premium, the result of the comparison will depend on the level of this cap. Asymmetric insurer duopoly is preferable to monopoly as long as $\overline{F} > \theta c + M + \frac{7\theta t}{8} - \frac{1}{2M}\left(\frac{\theta t}{12}\right)^2$.

Proposition 3 *As long as regulated monopoly profits are at least as large as duopoly profits, hospital products are not perfect substitutes, hospitals do not differ in cost efficiency, and the complete graph is feasible, the highest consumer welfare is achieved on the complete graph.*

Since vertical integration prevents the possibility of a 4-link configuration (provided that feasibility conditions (5.3)-(5.4) hold), it also reduces consumer welfare in this case. Therefore, we conclude with the following proposition.

Proposition 4 *Suppose regulated monopoly profits are at least as large as duopoly profits, hospital products are not perfect substitutes, hospitals do not differ in cost efficiency, and feasibility conditions (5.3)-(5.4) hold. Then vertical integration (if profitable) decreases consumer welfare.*

[18] More formally, it follows from Table 5.3 that the consumer surplus for the complete configuration (four contractual relations) exceeds the consumer surplus for an asymmetric network (three links) if and only if $\theta t/M < 36$. However, this condition must hold in the situation when the complete graph is feasible.

[19] In Gal-Or (1997), consumers benefit from the exclusion of a hospital, because this allows the insurers to negotiate a lower transaction fee with the other hospital, and the competing insurers (partly) pass these benefits to consumers.

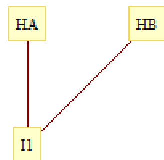
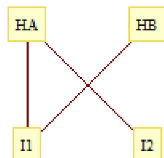
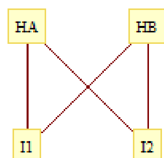
	Insurer premiums	Consumer surplus
	$\overline{F}$	$\frac{1}{2}(\theta v - \overline{F} - \frac{\theta t}{4})$
	$\theta c + M + \frac{\theta t}{12}, \theta c + M - \frac{\theta t}{12}$	$\frac{1}{2}(\theta v - \theta c - M - \frac{3\theta t}{8} + \frac{1}{2M}\left(\frac{\theta t}{12}\right)^2)$
	$\theta c + M, \theta c + M$	$\frac{1}{2}(\theta v - \theta c - M - \frac{\theta t}{4})$

Table 5.3: Welfare analysis

5.6 Conclusions

We have shown that insurers and hospitals may find it profitable to adopt exclusive behavior or to integrate vertically in order to increase their profit. However, these strategies lower consumer welfare. Even if consumer preferences regarding hospitals and insurers are symmetric and even if both hospitals are equally efficient, we find that, under some circumstances, insurers and hospitals may choose to contract selectively in order to secure more favorable contractual terms. The first exclusive strategy that follows from our theoretical duopoly model is that hospitals may choose to exclude one insurer from the market. This strategy will eliminate competition in the insurer market, raising total industry profits. Removing one insurer from the market, however, implies a loss of hospitals' bargaining power. Therefore, total industry profits must be substantially higher in the new situation (in our duopoly model, at least twice as high) to compensate the hospitals for this loss. The second strategy that can be adopted by market players involves the increased differentiation in the insurance market resulting from the differentiating hospital networks of the insurers. We find that if hospitals' differentiation is much larger than insurers' differentiation (see the conditions in Section 4.1), one hospital and one insurer will choose not to enter into a contract with each other. The principle behind this mechanism is that the insurer that contracts two hospitals benefits the most from its premium increase; the insurer is willing to share these gains with the other players via bargaining, which compensates the hospital-insurer pair that does not enter into a contract.

We show that vertical integration improves opportunities for increasing total industry profits by differentiating provider networks. We find that vertical integration makes it profitable for the owner of the integrated firm to foreclose its competitor. Therefore, a hospital takeover by an insurance firm will be accompanied by foreclosure of the competing insurer (i.e., by preventing the competing insurer's enrollees from access to this hospital). This raises profits for the hospital-insurer pair and decreases the profits of the competing insurer, because that insurer becomes less attractive to consumers. Similarly, an insurer takeover by a hospital will be accompanied by foreclosure of the competing hospital. This strategy also introduces asymmetry in the insurance market, which raises total industry profits. Furthermore, this strategy reduces the outside options of the competing hospital, thereby weakening its bargaining power. However, it also weakens the insurer's own position, this type of vertical integration can occur only if the profit increase achieved by network differentiation is very large. Our results show that both vertically integrated and independent firms can be present in the market in equilibrium, which differs from earlier theoretical papers, such as Ma (1997) and Gal-Or (1997), that have demonstrated the monopolization of one of the markets in equilibrium.

Our results provide guidance for the policy debate in countries that are moving towards a more market-oriented approach to health care. As the role of competition increases, these markets tend to reveal more information about the performance of individual hospitals and especially about the quality differences between hospitals. If consumers start to perceive more differentiation among hospitals, this could trigger insurers and providers to adopt exclusive strategies with possible anticompetitive effects. However, provider and insurer competition also has important efficiency effects, which could outweigh the possible anticompetitive effects of exclusive vertical restraints. The introduction of more competition in health care also stimulates providers and insurers to search for new organizational forms. For example, in the United States, this process has led to the appearance of a variety of managed-care organizations, featuring selective networks, that are based on vertical arrangements such as integration or exclusive restraints between insurers and providers. Our model suggests that these exclusionary networks could reduce consumer welfare unless these disadvantages are compensated by the efficiencies of integrated health care delivery.

We finally stress some limitations of our analysis and outline directions for further research. The analysis in this paper covers a symmetric, bilateral-duopoly model with fixed consumer demand for insurance and health care. In other words, there is market power both upstream and downstream. As our analysis shows, this means that any exclusionary equilibrium is (likely to be) anticompetitive.

Our model can be extended to incorporate variable consumer demand (see e.g. Halbersma and Katona, 2011) and asymmetries across both insurers and hospitals. Cost differences between hospitals and potential efficiency gains may be the reason for exclusive relationships (see e.g. Glied, 2000). Furthermore, we do not consider capacity constraints or vertical quality differences in the hospital sector that might also play a role in the bargaining outcome. The model could be further extended to incorporate these features. Another possible direction for further research could be the extension of our framework towards more flexible, empirically viable frameworks, as has recently emerged in the empirical literature (e.g. Capps et al., 2003; Ho, 2009).

Acknowledgement

We would like to thank to Paul de Bijl, Jan Boone, Marcel Canoy, Martin Gaynor, Katherine Ho, Viktoria Kocsis, Erik Schut, Misja Mikkers for their comments on preliminary versions of our paper. Furthermore, we thank participants of the IHEA conference in Beijing (2009) and EARIE conference in Ljubljana (2009), and our colleagues at the CPB and NZA for their comments during seminars at our institutes. Martin Gaynor and Katherine Ho were commissioned by the Dutch Healthcare Authority as external reviewers of an earlier draft of this paper. We are grateful to two anonymous referees and the co-editor for their careful and detailed comments and suggestions.

Appendix

5.A Netwerk configurations

Table 5.A.1 shows fifteen possible network configurations. We distinguish six monopoly and nine duopoly configurations. In the symmetric case each row in Table 5.A.1 represents the same type of network. These six different types of network configurations are also listed in Table 5.1.

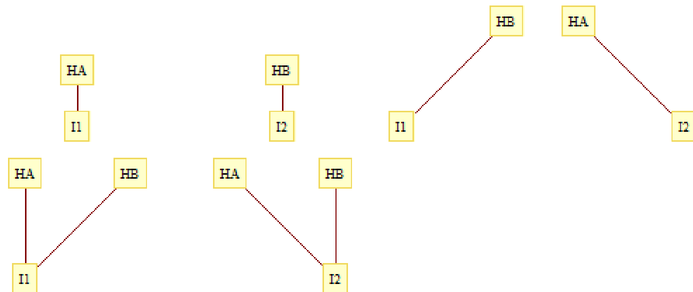
5.B Difference between our model and the model of de Fontenay and Gans (2005)

de Fontenay and Gans (2005) consider a model with Cournot competition in the downstream market. The firms negotiate over quantities and transfer payments. The contracted quantities fully determine the industry production pattern and, thus, the outcome of the quantity competition in the downstream market. Differently from that model, our model considers Hotelling competition in the downstream market. The contract specifies the total payment that will be paid to the hospital on top of the production cost related to medical services to the enrollees of this insurer choosing for this hospital. Therefore, quantities are not fixed by contracts, but determined as a result of the Hotelling competition between insurers, which takes place in the next stage of the game.

Generalized Myerson-Shapley values

As shown by Myerson (1977a), the solution of the bargaining equations presented in Section 3.5 is expressed by the generalized Myerson-Shapley value:

Insurer monopoly



Insurer duopoly:

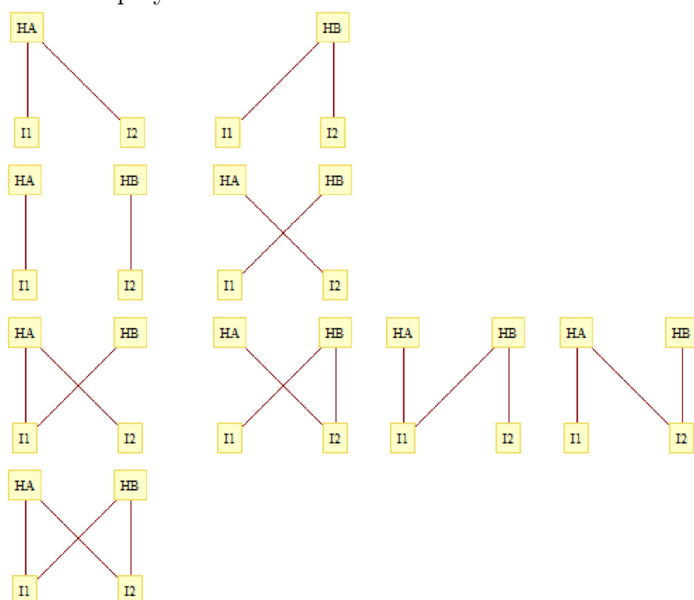


Table 5.A.1: All network configurations

$$\Phi_i = \sum_{P \in P^N} \sum_{S \in P} (-1)^{|P|-1} (|P|-1)! \left[\frac{1}{N} - \sum_{\substack{i \notin S' \in P \\ S' \neq S}} \frac{1}{(|P|-1)(N-|S'|)} \right] v^S_{L_P}. \tag{5.B.1}$$

Here P denotes a partition of N players into coalitions, P^N is a set of all partitions, S and S' are coalitions in P, $|.|$ is the notation for the number of elements in the respective set of elements, L_P is the partitioned graph (the graph what contains all links of L connecting the members of the same coalition within P, but excludes those connecting the members of different coalitions), and $v^S_{L_P}$ are the coalition profits on a partitioned graph L_P.

The expressions shown in Tables 1 and 2 can be also directly obtained by applying this formula and substituting the respective coalition values. In particular, we provide below the generalized Myerson-Shapley value expressions in the configuration in which both insurers contract both hospitals and there is no vertical integration:

$$\begin{aligned}
\Phi_1 &= \frac{1}{12}(-2v_{(1A)} - 2v_{(1B)} + 2v_{(2A)} + 2v_{(2B)} + v_{(1A)(1B)} + v_{(1A)(2A)} + v_{(1B)(2B)} - \\
&- 3v_{(2A)(2B)} + 3v^{1A}_{(1A)(2B)} - 3v^{2B}_{(1A)(2B)} + 3v^{1B}_{(1B)(2A)} - 3v^{2A}_{(1B)(2A)} + \\
&+ 3v_{(1A)(1B)(2A)(2B)})
\end{aligned}$$

$$\begin{aligned}
\Phi_2 &= \frac{1}{12}(+2v_{(1A)} + 2v_{(1B)} - 2v_{(2A)} - 2v_{(2B)} - 3v_{(1A)(1B)} + v_{(1A)(2A)} + v_{(1B)(2B)} + \\
&+ v_{(2A)(2B)} - 3v^{1A}_{(1A)(2B)} + 3v^{2B}_{(1A)(2B)} - 3v^{1B}_{(1B)(2A)} + 3v^{2A}_{(1B)(2A)} + \\
&+ 3v_{(1A)(1B)(2A)(2B)})
\end{aligned}$$

$$\begin{aligned}
\Phi_A &= \frac{1}{12}(-2v_{(1A)} + 2v_{(1B)} - 2v_{(2A)} + 2v_{(2B)} + v_{(1A)(1B)} - 3v_{(1A)(2A)} + v_{(1B)(2B)} + \\
&+ v_{(2A)(2B)} + 3v^{1A}_{(1A)(2B)} - 3v^{2B}_{(1A)(2B)} - 3v^{1B}_{(1B)(2A)} + 3v^{2A}_{(1B)(2A)} + \\
&+ 3v_{(1A)(1B)(2A)(2B)})
\end{aligned}$$

$$\begin{aligned}
\Phi_B &= \frac{1}{12}(+2v_{(1A)} - 2v_{(1B)} + 2v_{(2A)} - 2v_{(2B)} + v_{(1A)(1B)} + v_{(1A)(2A)} + v_{(1B)(2B)} - \\
&- 3v_{(2A)(2B)} - 3v^{1A}_{(1A)(2B)} + 3v^{2B}_{(1A)(2B)} + 3v^{1B}_{(1B)(2A)} - 3v^{2A}_{(1B)(2A)} + \\
&+ 3v_{(1A)(1B)(2A)(2B)})
\end{aligned}$$

The coalition profits in our model are as follows:

$$v_{(1A)} = v_{(1B)} = v_{(2A)} = v_{(2B)} = v_{(1A)(1B)} = v_{(2A)(2B)} = \Pi$$

$$v_{(1A)(2A)} = v_{(1B)(2B)} = v_{(1A)(1B)(2A)(2B)} = M$$

$$v^{1A}_{(1A)(2B)} = v^{2B}_{(1A)(2B)} = v^{1B}_{(1B)(2A)} = v^{2A}_{(1B)(2A)} = \frac{M}{2}$$

It is easy to see that substituting these coalition profits in the generalized Myerson-Shapley value expressions will yield the payoff values $\Phi_A = \Phi_B = \frac{1}{12}(M + 2\Pi)$ and $\Phi_1 = \Phi_2 = \frac{1}{12}(5M - 2\Pi)$, which coincide with the values shown in the last row of Table 5.1.

Bibliography

Baarsma, B., van Damme, E., van Dijk, T., Ottervanger, T., and VerLoren van Themaat, V. Eindrapport commissie verticale integratie tussen zorgverzekeraars en zorgaanbieders, 2009. In Dutch.

Bijlsma, M., Meijer, A., and Shestalova, V. Vertical relationships between health insurers and healthcare providers. Technical report, CPB Netherlands Bureau for Economic Policy Analysis, 2008.

Capps, C., Dranove, D., and Satterthwaite, M. Competition and market power in option demand markets. *The RAND Journal of Economics*, 34(4):737–763, 2003.

Ciliberto, F. and Dranove, D. The effect of physician–hospital affiliations on hospital prices in california. *Journal of Health Economics*, 25(1):29–38, 2006.

Cuellar, A. and Gertler, P. Strategic integration of hospitals and physicians. *Journal of health economics*, 25(1):1–28, 2006.

de Fontenay, C. and Gans, J. Vertical integration in the presence of upstream competition. *The RAND Journal of Economics*, 36(3):544–572, 2005.

de Fontenay, C. and Gans, J. Bilateral bargaining with externalities. http://works.bepress.com/joshuagans/14/, 2007. Working paper.

Douven, R., Lieverdink, H., Ligthart, M., and Vermeulen, I. Measuring annual price elasticities in dutch health insurance: a new method. CPB Discussion Paper 90, CPB Netherlands Bureau for Economic Policy Analysis, 2007.

Douven, R., Halbersma, R., Katona, K., and Shestalova, V. Vertical integration and exclusive behavior of insurers and hospitals. *Journal of Economics & Management Strategy*, 23(2):344–368, 2014.

Gal-Or, E. Exclusionary equilibria in health-care markets. *Journal of Economics & Management Strategy*, 6(1):5–43, 1997.

Gal-Or, E. Mergers and exclusionary practices in health care markets. *Journal of Economics & Management Strategy*, 8(3):315–350, 1999.

Gaynor, M. and Ma, C. Insurance, Vertical Restraints, and Competition. *GSIA Working Papers*, 1996.

Gaynor, M. and Vogt, W. B. Antitrust and competition in health care markets. *Handbook of health economics*, 1:1405–1487, 2000.

Getzen Thomas, E. Health Economics and Financing, 2007.

Glied, S. Managed care. In *Handbook of health economics*, volume 1, pages 707–753. Elsevier, 2000.

Halbersma, R. and Katona, K. Vertical restraints in health care markets. *TILEC Discussion Paper No. 2011-005*, 2011.

Ho, K. Insurer-provider networks in the medical care market. *American Economic Review*, 99(1):393–430, 2009.

Ho, K. The welfare effects of restricted hospital choice in the us medical care market. *Journal of Applied Econometrics*, 21(7):1039–1079, 2006.

Inderst, R. and Wey, C. Bargaining, mergers, and technology choice in bilaterally oligopolistic industries. *RAND Journal of Economics*, pages 1–19, 2003.

Ma, C. Option contracts and vertical foreclosure. *Journal of Economics & Management Strategy*, 6(4):725–753, 1997.

Melnick, G. A., Zwanziger, J., Bamezai, A., and Pattison, R. The effects of market structure and bargaining position on hospital prices. *journal of Health Economics*, 11(3):217–233, 1992.

Myerson, R. Graphs and cooperation in games. *Mathematics of Operations Research*, pages 225–229, 1977a.

Myerson, R. Values of games in partition function form. *International Journal of Game Theory*, 6(1):23–31, 1977b.

Stole, L. A. and Zwiebel, J. Intra-firm bargaining under non-binding contracts. *The Review of Economic Studies*, 63(3):375–410, 1996.

Strombom, B., Buchmueller, T., and Feldstein, P. Switching costs, price sensitivity and health plan choice. *Journal of Health economics*, 21(1):89–116, 2002.

Tirole, J. *The theory of industrial organization*. MIT press, 1988.

Wu, V. Y. Managed care's price bargaining with hospitals. *Journal of Health Economics*, 28(2):350–360, 2009.

Chapter 6

Welfare Standards in Hospital Mergers

Katalin Katona, Marcel Canoy

This chapter was published as:
Katona, K. and Canoy, M. Welfare standards in hospital mergers. *The European Journal of Health Economics*, 14(4):573–586, 2013

Abstract

There is a broad literature on the consequences of applying different welfare standards in merger control. Total welfare is usually defined as the sum of consumer and provider surplus, i.e. potential external effects are not considered. The general result is then that consumer welfare is a more restrictive standard than total welfare, which is advantageous in certain situations. This relationship between the two standards is not necessarily true when the merger has significant external effects.

We model mergers on hospital markets and allow for not-profit-maximizing behavior of providers and mandatory health insurance. Mandatory health insurance detaches the financial and consumption side of health care markets, and the concept 'consumer' in merger control becomes non-evident. Patients not visiting the merging hospitals still are affected by price changes through their insurance premiums. External financial effects emerge on not directly affected consumers.

We show that applying a restricted interpretation of 'consumer' (neglecting externality) in health care merger control can reverse the relation between the two standards; consumer welfare standard can be weaker than total welfare. Consequently, applying the wrong standard can lead to both clearing socially undesirable and to blocking socially desirable mergers. The possible negative consequences of applying a simple consumer welfare standard in merger control can be even stronger when hospitals maximize quality and put less weight on financial considerations. We also investigate the implications of these results for the practice of merger control.

6.1 Introduction

Competition authorities that need to decide on mergers have some leeway as to the criteria for their assessment of welfare. Most authorities made a choice for consumer welfare. However, there are some countries such as Canada and Australia where merger control seems to follow the principles of total welfare Renckens (2007). There is a long standing discussion in the economic literature on whether one or the other standard is preferable. The question is addressed both from practical and theoretical perspectives. We contribute to this discussion by pointing out that specific features of the health care sector, making externalities to emerge, reverse some key results.

In the simplest case, total welfare in an economy is calculated as the sum of consumer surplus and producer surplus Motta (2004). In some industries however, other components of welfare such as external effects on not directly affected consumers may arise as significant. We could take examples from environmental economics (assuming that the merger influences the magnitude of the externality, e.g. pollution), from the financial sector (thinking on the effect of a merger on the stability of the whole system) or as we do in this article, from the health care (and insurance) sector. Facing a merger case with significant external effects, competition authorities should not lean on general results but take into account the peculiarities of the market. As shown in this article, neglecting the external effects can alter the conclusion on the preference of consumer or total welfare standard in merger control.

The question whether to explicitly include externalities in merger analysis of health care markets raises another ambiguity. The relationship between health services and health insurance markets makes non-evident who the consumers are in a merger of providers (e.g. hospitals). The exact definition of the consumer determines whether externalities are implicitly taken into account or their explicit consideration in the analysis is needed. Besides presenting the consequences of neglecting to include important externalities in the analysis, we also embrace the topic of how potential non-profit-maximizing behavior of hospitals affects the results.

Merger control of hospitals is a relatively new branch for competition authorities, and is less analyzed in the literature on optimal welfare standards. Hospital markets are liberalized in a number of countries (e.g., Switzerland, Netherlands, US) making it a sector that falls under competition law scrutiny.[1]

Hospital markets have certain specific features that require attention. First, hospitals are not necessarily striving for maximum profits. When providers at-

[1]The standpoint that competition law applies to health care just like to any other sector in the US stems from the case Goldfarb v. Virginia State Bar in 1975 Sage et al. (2003).

tach much weight on quality arguments in their merger decisions[2], it is more probable that mergers of socially undesirable (i.e. excessive) quality improvements are initiated. Second, the market (in many countries) is characterized by the fact that patients are insured. The concept 'consumer' becomes non-evident when insurance is mandatory, premiums are uniform regardless of heterogeneity among consumers, and in absence of significant co-payments. Some of these elements, i.e. mandatory insurance, uniform premiums and absence of significant co-payments, apply to other (non-health) insurance markets as well. However, the combination of all the three characteristics particularly features health care insurance (e.g., in Germany, France and The Netherlands).

Patients receive treatment from the providers but do not directly pay for the services. Insurers reimburse providers and collect premiums from their clients. The level of the premium is, however, independent from the individual consumption[3] making the financial and consumption side of the market for hospital services detached.

Taking a hospital's perspective, a hospital has on the one hand clients to whom it offers services (patients) and on the other hand has clients from whom it receives reimbursement for the same services (insurers). The group of consumers receiving the services and the group paying for it through health insurance do typically not coincide. Such level of detachment of monetary transactions and service flow makes the correct definition of the consumer non-evident. Theoretically, there could be many different definitions of a 'consumer'. Moving from a narrowly to a broadly defined group, a consumer can be: one that actually visits the given hospital, one that might visit it, or one that through the insurance premium pays for the services offered by the given hospital. The subsequent definitions broaden the group of consumers and consider the financial effect in an increasing level. Since insurance spreads the health care expenditure across all its clients, only the most extended definition allows one to consider the whole financial effect of a merger. If we use a narrower definition, we disregard the external effects of the uniform insurance premiums.

Our contribution to the literature is that we explicitly model these specific hospital market features by which we consider the consequences of potential externality and not-profit-maximizing behavior on the analysis. We define two groups of consumers. First, the broadly defined group 'indirectly affected con-

[2]There is no consensus in the literature about what (not-for-profit) hospitals maximize. See section 6.2.4 for a overview of proposed alternatives.

[3]The level of premium depends on the total consumption of health care services (at macro level). Increased consumption due to, for example, moral hazard problems makes the premiums rise. In this article, we do not consider the problem of moral hazard and abstract from possible links between consumption level and the premium. This is an acceptable assumption in health care systems where patients need the GP's referral to visit a hospital, like in the United Kingdom, the Netherlands, Denmark and Norway.

sumers' contains all clients of insurance companies. Some of these consumers are likely to never visit the merged hospital (e.g., because they live far away). They are not affected by possible changes in the service level of the hospital. These patients, for example, do not benefit from a quality improvement or from the introduction of a new technology in the merged hospital. They are, however, indirectly affected through the financial effects. Since they purchase health insurance, all these consumers share the burden of a possible rise in insurance premium; they are exposed to an external effect of the merger. (The previously mentioned group of consumers who at the moment are not sick but in the future might visit the hospital are also included in this category. They draw some benefits from the quality improvement - we will define this as option value in section 6.4 - because in case of illness they can expect better treatment. They are also affected by potential price increases through their health insurance.) Second, the narrowly defined group 'directly affected consumers', which is also part of the previous broader group, contains only the clients of insurance companies that actually visit the merged hospital. These consumers are not only exposed to financial effects, but they are also directly affected by possible changes in quality. However, when using this narrow definition, we do not consider the whole financial effect of the merger. Specifically, the premium paid by consumers who do not visit the merged hospital is not considered. A part of the financial effect appears as an externality in the analysis.

Depending on the definition of 'consumer', the exact meaning of consumer welfare standard also changes. Applying a narrow definition, that is neglecting the externality, the consumer welfare (CW) standard can result in a more lenient criterion than the total welfare (TW) standard, which is in contrast to the general view in the literature. In markets where consumers obtain the benefits and pay the costs of their consumption, the CW standard has been equal to or tougher than the TW standard. Many theoretical models (e.g., Besanko and Spulber (1993); Lyons (2003)) have built on this feature by showing the superiority of CW standard in a number of circumstances. We show that the externality effects on health care markets stemming from insurance can reverse the relationship between CW and TW standard, which questions the generality of the claims in the literature.

From a policy perspective, a narrow definition of the 'consumer' in the CW standard can lead to both approving socially undesirable mergers as well as blocking socially desirable mergers. A CW standard, which applies the most extended definition of 'consumer' and so implicitly includes the external effects, repairs this problem. However, it requires one to consider effects that are potentially external to the relevant hospital market of the merger. In the process of merger control, this dilemma appears in the evaluation of the potential positive (quality) effects of a merger against the negative (price) effects. The externality caused by insurance is reflected in the diverging valuation of quality improvements by different groups of consumers. The method to value potential quality improvements and aggregate it across all consumers influences the effectiveness of the merger standard.

The next subsection reviews the economic literature that investigates the reason for different welfare standards in merger control. Section 2 describes the model. Section 3 summarizes the results of our model and discusses the consequences of applying different merger standards from a theoretical perspective. Policy implications and relevance to the current practices are described in Section 4. In Section 5, we conclude.

6.1.1 Related literature

There are two branches of economic literature analyzing which welfare standard is optimal in competition policy analysis. One compares practical and direct effects of applying TW or CW while the other branch applies the agency framework to analyze the decision problem of competition authorities. We summarize the arguments of both approaches. At the end of this section, we relate the results of the literature on non-for-profit hospital mergers to our findings.

Articles of the practical approach (e.g., Heyer (2006); Salop (2005); Pittman (2007)) focus broadly on two aspects of a merger: changes in efficiency of production and (re)distribution among different groups of society such as providers and consumers. Regarding efficiency, it is analyzed whether a merger offers opportunities to produce more quantity or better quality given the scarce resources in the economy. The TW and CW standard differs in what exactly is considered as efficiency gains. While the TW standard values every efficiency improvement, the CW standard merely acknowledges gains that are passed on to consumers. Fixed cost savings, for example, may outweigh the anticompetitive effects of a merger according to the TW standard but do not contribute to the CW standard. The CW standard ignores some efficiency gains and gives priority to distributional aspects. Using the CW standard can lead to a situation where consumers are the final beneficiaries of the merger, but it forgoes some efficiencies that would solely benefit providers. The TW standard does not consider (re)distributional effects but evaluates mergers solely on efficiency considerations. There is no clear conclusion in the literature which approach is in line with the goals of competition policy.

Articles in the other branch of the literature look at the merger control process as a whole and apply principal-agent theory (e.g., Neven and Röller (2005); Fridolfsson (2007); Besanko and Spulber (1993); Lyons (2003)). They assume that the ultimate goal of society is to maximize TW. In these models the competition authority is an agent that controls mergers according to a given welfare standard. The focus of these analyses lies in the consequence of choosing TW or CW standard as the objective function of the competition authority. In spite of the fact that the final goal is to maximize total welfare, it can sometimes be achieved by defining CW as the objective function for the agent authority. This can be explained by a general result according to which CW is a more restrictive standard than TW. Besanko and Spulber Besanko and Spulber (1993) and Lyons Lyons (2003) both build on this characteristic of CW and show under which conditions the CW standard achieves higher TW than the TW standard.

Besanko and Spulber Besanko and Spulber (1993) apply a model of asymmetric information to show that a tougher merger standard than the TW standard increases the societal gain from a merger. Since authorities cannot perfectly estimate the welfare consequences of a merger, their decision is a random variable from the firms' perspective. Furthermore, rejection of a merger has a higher probability when the CW standard (a tougher criterion) is applied. Because preparing and submitting a merger proposal is costly, firms initiate mergers that they expect to be accepted. In the case of a tougher merger standard, it results in a self selection toward socially preferable merger alternatives. The key elements in this model leading to the preference of CW standard are the costly procedure and asymmetric information.

Lyons Lyons (2003) derives the relative advantage of one or the other welfare standard from the diverging treatment of changes in fixed costs. CW standard incorporates cost reductions only if they are passed on to consumers. Therefore, fixed cost reductions are excluded from the CW analysis. Welfare gains from mergers in Lyons' article is described as the ratio of anticompetitive effects and fixed cost reductions. The CW standard is a tougher standard because potential anticompetitive effects cannot be compensated by cost reductions. Since firms prefer mergers with anticompetitive effects (higher prices), the CW standard is more likely to reject the first proposal of firms than the TW standard. The desirability of rejection in the long term depends on alternative mergers. If the subsequent proposal of firms yields a higher TW, then rejection was a desirable decision. If the alternative is a socially less beneficial merger, then approval by TW standard is a better strategy. Lyons Lyons (2003) models a given industry structure and analyses sequential mergers to find the equilibrium structure conditional on the merger standard.

These articles have considered general sectors without significant externalities, a set-up that does not fit health care markets. Calem et al. Calem et al. (1999) focus on distinguishing different welfare measures specific to health care markets. They emphasize two distinguishing characteristics of hospital (and in general health care) markets. First, health care insurance causes moral hazard in the consumption of hospital services to the extent of the co-payment rate. Second, hospitals may be non-profit; specifically, they may maximize output instead of profit. Considering these characteristics of hospital markets, they compare the effects of a merger on consumer surplus (gain from hospital services minus co-payments paid by consumers), net social surplus (gain from hospital services minus price paid by the insurer) and gross social surplus (gains from hospital services minus costs born by hospitals). They model quality competition among hospitals, which may yield over-production of quality because of moral hazard or the non-profit nature of hospitals. Consequently, a merger may be gross social welfare enhancing since it reduces quality competition and restricts excess quality. Considering only consumer surplus, which reduces with decreasing quality, can be misleading when evaluating hospital mergers. These results are health care specific but are not explicitly linked to the literature on merger control. Our article makes this last step too. Similar to Calem et al. Calem et al. (1999), we compare various welfare concepts applied to health care markets. In addition, we look explicitly at the consequences of using these measures in merger control.

Finally, we discuss the strand of literature that investigates whether not-for-profit (NFP) organizations should be treated differently in merger control than for-profits (FP). Several articles show that the behavior of NFP firms can be interpreted as a profit-maximizing behavior with lower perceived costs. Beside monetary profit, NFP firms gain additional utility from production, which makes them accept higher costs for the same level of production. This attribute of NFP organizations appears in our model as well. Both theoretical (e.g., Philipson and Posner (2006); Richman (2007)) and empirical (e.g., Vita and Sacher (2001); Gaynor and Vogt (2003)) articles conclude that not-for-profit hospitals exploit their market power in a similar way than their for-profit counterparts.

Prüfer Prüfer (2010) however shows that this result depends on the assumptions made on the maximand of the NFP firms. In his article, NFP firms with an owner preferring high quality produce excessive quality (from the societal perspective). The merger (to monopoly) eliminates competition, which indirectly makes the firm produce lower quality, in this way increasing the total welfare. Prüfer Prüfer (2010) draws the attention to the importance of examining the objective of owners of the merging NFP firms when assessing effects on society.

It is important to notice that the not-for-profit status of hospitals is different from the possible not-profit-maximizing behavior. Non-for-profit status is a definition used in the context of taxation and refers roughly to two rules. NFP hospitals enjoy exemption from taxation, and they are not allowed to pay rents to their owners (non-distribution constraint). In contrast, the not-profit-maximizing behavior refers to the objective function of the hospital that is revealed in its decisions. In this article, we consider this second possibility but do not discuss the case of not-for-profit status. Similar to the literature above, our model shows that merged quality-maximizing hospitals exploit their market power just as their profit-maximizing counterparts but their decisions on merger can differ significantly.

6.2 Model

Our static model includes three players; hospitals deciding whether to initiate a merger, consumers paying the insurance premium and choosing a hospital when they fall ill, and a competition authority blocking or approving the merger. The standard that the competition authority applies in merger control is either the TW standard or a version of the CW standard. In both cases, the authority approves the merger if the standard indicates net gains and blocks the merger if the standard indicates net losses. We do not model the insurance market explicitly. This assumption is not restrictive for the purpose of this study.

6.2.1 Hospital market and consumer preferences

We characterize the market for hospital services and consumer preferences by applying the circular city model. Distance to a hospital is an important choice factor of consumers (see, e.g., McGuirk and Porell (1984)) which makes substitutability of providers asymmetric, i.e., dependent on distance from the consumer's location. Location models fit this characteristic of the market. Furthermore, we focus on effects of a merger and consider hospitals in any other aspect symmetric. The circular model, in contrast to the linear model, allows for this.[4]

Let the n hospitals offering treatment to patients be located on a circle of unit circumference at equal distance from each other. Besides horizontal differentiation, hospitals may also vary in quality of services offered. Patients[5] are uniformly distributed on the circumference of the circle. We assume that every patient prefers hospital services of higher quality to that of lower quality and that they dislike traveling. They trade-off quality and distance from the hospital uniformly in ratio t. Patients do not pay directly for their treatment; therefore, the price does not play a role in their hospital choice. Thus, the utility derived from receiving hospital treatment includes two terms; the quality of the services in the visited hospital (q_i) and the distance to the hospital (x).

[4]The leading condition on the model of the hospital market that implies our main result is that clients of an insurer do not equally benefit from quality improvements in a specific hospital. This condition assumes a model in which not all consumers choose the same hospital even if that hospital is of higher quality. If there is diversity in hospital choice among consumers (e.g., due to horizontal differentiation of hospitals as in the circular city model), then our main result holds. Therefore, the circular city model is not restrictive.

[5]We use the term patient to refer to a consumer who needs hospital treatment.

$$U = q_i - tx \tag{6.1}$$

The demand for hospital i's services (equation 6.2) consists of the sum of two 'half demands': the demand in the market segments where hospital i competes with hospital $i+1$ and $i-1$ respectively. In each segment, the demand is derived by determining the position of the indifferent patient based on the utility function (equation 6.1).

$$D_i(q_i, q_{i-1}, q_{i+1}, t) = \frac{1}{n} + \frac{q_i - q_{i-1}}{2t} + \frac{q_i - q_{i+1}}{2t} \tag{6.2}$$

Equation (6.3) describes the utility that patients of hospital i derive from their visit. We refer to this value as patient welfare produced by a given hospital (PW_i). Equation (6.4) defines the (total) patient welfare (PW) which is the sum of the welfare produced by each hospital.

$$PW_i = \int_0^{\frac{1}{2n}+\frac{q_i-q_{i-1}}{2t}} (q_i - tx)\,dx + \int_0^{\frac{1}{2n}+\frac{q_i-q_{i+1}}{2t}} (q_i - tx)\,dx \tag{6.3}$$

$$PW = \sum_{i=1}^{n} PW_i \tag{6.4}$$

6.2.2 Insurance market and hospital-insurer bargaining

We do not model the insurance market and the hospital-insurer relationship explicitly but make some simplifying assumptions. In the insurance market, we assume Bertrand competition among symmetric firms, which results in premiums at the level of the uniform marginal cost. The single role of the insurance market in this model is to pool patients' health care expenditure and set a uniform premium for all consumers. In this model prices between insurers and hospitals are assumed to be results of negotiations. Instead of explicitly modeling the negotiations, we make two assumptions on the outcomes and incorporate these simplified solutions in the further steps of the model. These assumptions are common in the literature and are not restrictive for this model.

First, negotiated prices between hospitals and insurers consists of two parts: reimbursement of the costs of the hospital and a share of the net gain from concluding the contract[6]. The net gain is here defined similarly to Capps et al. Capps et al. (2003), i.e., it equals the added value that the given hospital brings

[6] We do not assume a specific solution of this game but use the general efficiency condition of cooperative games: the negotiating hospital-insurer pair shares the total net gain from their cooperation (contract).

to the insurer's network minus the additional expenditure (or saving) caused by including the given hospital in the network. Including an extra less efficient hospital to the insurer's network yields added value to patients because they need to travel less. At the same time, including a less efficient hospital means extra costs for the insurer. Therefore, when concluding one more contract, the insurer's clients have to travel less but are cured more expensively. The difference of these two effects is the net gain from concluding the contract. Negotiating hospital-insurer pairs bargain about the division of this net gain.

Formally, the added value of a given hospital is defined as the total patient welfare assuming patients may attend all hospitals in the market minus the total patient welfare assuming that patients may attend any but the given hospital in the market. This value represents the added utility that patients derive from the existence of that given hospital. Note that this formula yields higher added value for a hospital of high quality or in an isolated location than for a hospital of average quality in a densely populated location. The underlying intuition is that dropping the high-quality low-density hospital from the market leads to consumers substituting it for a hospital of considerably lower quality or for one lying relatively far away. Such substitution means loss of utility for consumers. Additional expenditure is defined similarly; costs of the insurer when the hospital is included in its network minus the costs assuming that the hospital is not part of the network.

Second, in order to keep the model simple, we assume that the insurer and the hospital share the net gain from concluding the contract in a given proportion, namely 50-50%[7]. We apply the following formula to determine the price of hospital i's service (w_i)

$$w_i = c_i + \left(\frac{PW - PW_{-i}}{D_i} - \frac{TE - TE_{-i}}{D_i} \right) \frac{1}{2}$$

where c_i denotes the constant average cost of hospital i, TE denotes the total expenditures of the insurer, which is defined as $TE = \sum_{i=1}^{n} D_i w_i$. PW_{-i} denotes the total PW when hospital i is not in the market and similarly TE_{-i} is the total expenditure of the insurer when hospital i is not in the market.[8] Note that we assume hospitals to agree on the same price with all the insurers, i.e., a hospital has a single price. Since insurers are symmetric in the model, this is a logical assumption.

When modeling a merger between hospitals, we apply the notion of a merger as defined in the property rights literature (e.g., Hart and Moore (1990)). The

[7] Any other proportion would be possible and would not alter the qualitative results of our model.

[8] Note that $PW_{-i} \neq PW - PW_i$ (and $TE_{-i} \neq TE - TE_i$). The left hand side describes the patient welfare (total expenditure) calculated for $n-1$ hospitals in the market, while the right hand side represents the patient welfare (total expenditure) in a part of the market.

idea is that integration implies the shift of ownership rights between the merging entities. The hospital that is overtaken does not act as an individual entity anymore but the new owner of the merged hospital has disposal of both hospital locations. Therefore, a single decision maker negotiates with the insurer following the merger. If they fail to agree, both hospital locations become unavailable for patients. PW_{-i} is thus calculated by dropping both hospitals from the network. Therefore, patients have to travel further for a substitute than before merger, and the added value of the hospital increases. A merger leads to higher prices, ceteris paribus, reflecting the increased market power of the merged hospital.

Note that the price depends only on the value that hospitals add to patient welfare both in the integrated and the non-integrated situation. Hospitals of higher than average quality produce more added value; therefore, they have higher prices. This is, however, independent of the hospital's objective function[9]. Similarly, the price of the merged hospital increases regardless its maximand. This result is in line with the literature on NFP hospitals, which shows that NFP hospitals exploit their market power similarly to their FP counterparts. Our results coincide with this, although we model a bargaining outcome in contrast to the usually assumed price setting behavior.

6.2.3 Welfare measures

Based on the exact definition of the 'consumer' used in the merger analysis, we identify two different consumer welfare standards. We define 'simple CW' as the difference of PW and the share patients pay from health care expenditures. Simple CW measures the direct effect of a merger, and does not consider external effects introduced by health insurance. The concept of the 'consumer' is defined here as patients visiting the hospital, which is a narrow definition because it excludes a large group of healthy consumers or consumers in other hospital markets.

The extended definition of 'consumer' that we use involves everyone affected by the merger (including the previously defined smaller group), which implies all people covered by the same insurance because through the uniform premium they pay they are affected by changes in the hospital prices. Furthermore, we assume that consumers from other hospital markets can also be pooled by the same insurance and so can be affected by the merger. Expenditures can be shared among consumers in a larger region than the hospital market each patient considers. We will call the welfare measure calculated as PW derived from hospital services minus total health care expenditure as 'extended CW'.

[9]The hospital's objective function (profit-, quality-maximization or partially both) matters in the merger decision of hospitals. See also section 6.2.4 and 6.2.5.

The difference between simple and extended CW is in the cost component, i.e., the implicit in- or exclusion of external effects. Patients visiting the hospital enjoy all benefits of a potential quality increase (reflected in increasing PW), but pay only a proportion of potential extra costs[10]. When considering all consumers affected by the merger, PW is still considered, and costs are also fully taken into account. Calculating the extended CW, we internalize the external effects of insurance on consumers paying premium but not visiting the given hospital.

To formalize the concept of simple CW and extended CW, let us define the ratio

$$S = \frac{\text{Number of patients directly affected}}{\text{Number of all affected consumers}}$$

Equation (6.5) defines simple CW, while Equation (6.6) defines extended CW. Note, that S can also be interpreted as the ratio of consumers that are included in the merger analysis. In this way, S is a continuous variable that determines the level of externality and the distortion introduced by partial analysis that excludes a group of consumers paying premiums. Equation (6.6) shows that the extended CW can be written as the sum of simple CW and the externality effect caused by health insurance.

$$CW_{simple} = PW - S\sum_{i=1}^{n} D_i w_i \tag{6.5}$$

$$CW_{ext} = PW - \sum_{i=1}^{n} D_i w_i = PW - S\sum_{i=1}^{n} D_i w_i - (1-S)\sum_{i=1}^{n} D_i w_i \tag{6.6}$$

TW is defined as the difference of PW and the cost of its production, which equals the sum of the welfare of all groups in the society.[11] Similar to the theoretical strand of the literature, we use TW as benchmark.

[10] This is the consequence of our assumption on uniform insurance premiums. If premium differentiation among consumers is possible, then every patient (group) can bear its own financial burden; i.e., there would not be externality effect and hospital mergers would not differ from non-health care mergers. Co-payments are a tool of premium differentiation. However, co-payments generally do not shift the whole cost difference to patients but only a part of it, so only affecting the quantititave nature of our results.

[11] First, considering the extended definition of consumer, TW is the sum of CW and suppliers surplus, i.e. $TW = (PW - \sum_{i=1}^{n} D_i w_i) + (\sum_{i=1}^{n} D_i w_i - \sum_{i=1}^{n} D_i c_i)$. Second, considering the restricted definition of consumer and so calculating with the simple CW, we can write TW as $TW = (PW - S\sum_{i=1}^{n} D_i w_i) + (\sum_{i=1}^{n} D_i w_i - \sum_{i=1}^{n} D_i c_i) - ((1-S)\sum_{i=1}^{n} D_i w_i)$. The first term equals the simple CW, the second term is the suppliers surplus, while the third term is the externality effect, i.e. effect on other consumer groups in the society. In the calculations of TW, externalities have to be taken into account since they form costs (or benefits) for society although not for the group defined as 'consumers' or 'suppliers' in the analysis. It can be compared to the textbook example of production or consumption externalities. The individual and social costs of consumption diverge, which results in externalities that are not

$$TW = PW - \sum_{i=1}^{n} D_i c_i$$

6.2.4 Objective function of hospitals

Several articles in the literature test possible assumptions on the objective function of non-for-profit hospitals. The assumption of non-profit maximization is investigated by several empirical articles Chang and Jacobson (2010); Deneffe and Masson (2002); Horwitz and Nichols (2009). They confirm that the maximand of NFP hospitals is not the monetary profit. Deneffe and Masson Deneffe and Masson (2002) and Horwitz Horwitz and Nichols (2009) find that NFP hospitals are most likely to maximize output or a mix of monetary profit and output, while Chang and Jacobson Chang and Jacobson (2010) conclude that the data is the most consistent with the theory of perquisite maximization. Others (e.g., Horwitz (2005); Clement et al. (2002)) at the same time find that NFP hospitals provide more charity or unprofitable care or higher quality than for-profits, which can be the result of some kind of welfare maximization. Although there are several hypotheses on the real objective function of NFPs, there is no consensus in the literature about it. Malani et al. Malani et al. (2003) overview the empirical literature and conclude that there is not enough evidence to distinguish among different theories on the NFP objective function.

In our model, we assume that hospitals maximize a combination of monetary profits and quality.[12] In one extreme case of our model, hospitals maximize purely the quality level of their care regardless of monetary profits. The other extreme is the pure profit-maximizing behavior. We will refer to hospitals following exclusively the previous strategy as purely-quality-maximizing hospitals, while to hospitals following the latter strategy as purely-profit-maximizing hospitals.

Quality improvement can be a strategy for both the purely-quality- and purely-profit-maximizing type. In contrast to quality-maximizing hospitals where quality improvements directly increase the objective function of the hospital, quality improvement only has an indirect effect on profit-maximizing hospitals. Quality increases the monetary profit through higher market shares and higher

considered by the consumers but that are costs for society. See, e.g., in (Begg and Vernasca, 2008, p. 304).

[12]The article closest to ours, Calem et al. Calem et al. (1999) defines the goal of non-profit hospitals in outcome maximization. Quality maximization is an alternative hypothesis that is also supported by a number of empirical studies. Quality maximization is a form of perquisite maximization defined and tested in Chang and Jacobson Chang and Jacobson (2010). Further, Malani et al. Malani et al. (2003) could also not reject the hypothesis that hospitals maximize non-contractible quality.

prices. The motive of a profit-maximizing hospital for quality improvement essentially differs from the motive of a purely-quality-maximizing hospital.

Specifically, we assume that hospitals maximize the weighted average of monetary profit (Π) and quality (q). The objective function of hospital i is thus:

$$G_i = \alpha \Pi_i + (1-\alpha) q_i = \alpha D_i(q,t)(w_i - c_i) + (1-\alpha) q_i \tag{6.7}$$

where Π_i denotes the monetary profit of the hospital, which is demand multiplied by price minus cost. We denote the relative weight of profit maximization to quality maximization in the decisions of the hospital by α. In case of a purely-profit-maximizing hospital, $\alpha = 1$, while in case of a purely-quality-maximizing hospital, $\alpha = 0$. Hospitals may aim at both objectives, i.e., they have an eye on costs, but increase quality just for its intrinsic value as well (and not to seek higher market share). We assume that the value of α is the same for all hospitals in the market.

Our definition of the objective function is similar to the model of profit-deviating firms used in Lakdawalla and Philipson Lakdawalla and Philipson (1998) though they assume 'output-preferring' hospitals. It is important that, similar to the model mentioned above, we do not model the not-for-profit status of hospitals (i.e., the nondistribution constraint), but focus on the profit-deviation behavior. Budget constraint for hospitals such as a constraint for positive monetary profit is not assumed. Lakdawalla and Philipson Lakdawalla and Philipson (1998) argues that donors who gain utility from the profit-deviation can cover the negative profits. For the ease of presentation, we also disregard a budget constraint, the inclusion of which would not change our qualitative results.

6.2.5 Merger decision of hospitals

Two hospitals will only initiate a merger if it results in an increase in their objective function, i.e., their joint gain (G_{i+j}) is higher after the merger than before it.

$$\begin{aligned} G_{i+j}^{before} &< G_{i+j}^{after} \\ \alpha(D_i(w_i - c_i) + D_j(w_j - c_j)) + (1-\alpha)\frac{q_i + q_j}{2} & \\ &< \alpha D_{i+j}(w_{i+j} - c_{i+j}) + (1-\alpha) q_{i+j} \end{aligned} \tag{6.8}$$

Where q_{i+j} is the quality level after merger, c_{i+j} is the cost after the merger, D_{i+j} is the demand of the merged hospital, while w_{i+j} is the price of hospital services calculated for the merged hospital.

As described at Equation (6.7), α gives relative weights to monetary and non-monetary benefits in the objective function of hospitals. From Inequality (6.8) can be seen that a purely-quality-maximizing hospital will initiate any merger with $\frac{q_i+q_j}{2} < q_{i+j}$ irrespective to costs. Other hospitals ($0 < \alpha \leq 1$) consider both quality and cost consequences of a merger. If there are several possible mergers the hospital chooses the one that ensures the highest gain.

Specifically, we assume two effects of a merger: 1) a merger specific change in quality, Δq, and 2) a merger specific change in marginal cost, Δc.[13] The change in quality can be an increase, assuming that, e.g., larger hospitals have a better reputation and so attract better personnel. It can also be a decrease in quality, assuming that cultural differences in the two institutions lower quality in the short run. Costs may decrease or increase owing to a merger. One reason for a decrease can be the scale economy of certain activities. A cost increase, however, is possible, for example, due to higher organizational costs after the merger.

We assume that quality differences after the merger are not so large that it is worthwhile for patients to pass by the nearest hospital (that is to travel more than 1/n of the circle): $\frac{t}{n} > |\Delta q|$. Furthermore, we only consider mergers between two neighboring hospitals, and assume that the merged hospital does not close any of its locations. The only strategic action on the part of the hospital that we explicitly consider is the decision on the merger: whether and with which (neighboring) hospital to merge.

6.3 Results

To keep the presentation of results simple, we set the initial values of quality and cost level uniform for all hospitals in the market. In this way, we have two quality and two cost levels in the model: q and c for all hospitals before the merger, which changes to $q + \Delta q$ and $c + \Delta c$ for the merged hospital after the merger (but remains q and c for all other hospitals). Furthermore, we set the transformation rate between quality and traveling at 1 ($t = 1$). Finally, we fix the number of hospitals in the market before the merger at 4 ($n = 4$).

The following equation produces the gain (G_{i+j}) that merging hospitals can obtain:

$$\Delta G_{i+j} = \alpha \left[(\Delta q - \Delta c) \left(\frac{1}{4} + \frac{\Delta q}{2} \right) - \frac{(\Delta q)^2}{4} + \frac{1}{32} \right] + (1 - \alpha)\Delta q$$

As can be seen, even in the absence of cost and quality effects, hospitals gain $\frac{\alpha}{32}$ because of their increased bargaining power. Increasing costs lower the gains from

[13] The assumption that the change in quality is merger specific means that other hospitals cannot change their quality level. This is a realistic assumption when, for example, scale effects play a role.

the merger or do not influence it in case of purely-quality-maximizing hospitals. Effects of quality improvements depend on the level of Δc and α. Purely-quality-maximizing hospitals ($\alpha = 0$) always gain from quality improvements. Hospitals that (partially) maximize monetary profit can gain as well since better quality attracts more patients. More patients generate monetary profit if the hospital has a positive price cost margin. For $\alpha = 1$, for example, the condition for increase in gains is $\Delta c < \frac{1}{2} + \Delta q$.

6.3.1 Effects on different welfare measures

We will compare the effects of a merger for alternative welfare measures: simple and extended CW, and TW. Each of these measures can be used in merger control. However, they yield different conclusions. Simple CW is specific to health care markets. Extended CW corresponds to CW in non-health care markets, and the definition of TW does not differ either. Reasons for and effects of applying CW or TW standard in non-health care markets have been extensively analyzed in the economic literature, while Calem et al. Calem et al. (1999) have described some health care specific welfare measures. We relate our findings to these previous results.

Change in the (simple) consumer welfare is given by the following equation:

$$\Delta CW = \Delta q \left(\frac{1}{2} + \Delta q\right) - \frac{(\Delta q)^2}{2} - \frac{S}{2}\left[\Delta c + (\Delta q - \Delta c)\left(\frac{1}{4} + \Delta q\right) - \frac{(\Delta q)^2}{2} + \frac{1}{32}\right]$$

The first term represents the quality gain due to merger, while the last term shows the price effects. The term between ($\frac{(\Delta q)^2}{2}$) is the loss in travel time; consumers travel more because they sometimes opt for a hospital of better quality that is further away than the nearest hospital. This loss from traveling is always compensated by the gain from higher quality (in the first term) otherwise patients would not choose the far away hospital.

The last term indicates that S mitigates the costs effects, i.e., simple CW is a harder constraint than TW when cost effects are advantageous and it is a weaker constraint when price effects are disadvantageous. Simple CW reaches its extreme value in S when the number of consumers affected directly and indirectly coincides ($S = 1$). Consumers cover all the health care expenditures, which also yields that simple CW equals the extended CW. Generally, extended CW is the boundary of simple CW since it considers the complete cost effects instead of partial analyses.

The concept of 'consumer welfare' and 'net social welfare' in Calem et al. Calem et al. (1999) is similar to our simple and extended CW, respectively. They diverge, however, in modeling the insurance market since Calem et al.

Calem et al. (1999) do not include insurance premiums in consumer surplus, but consider out-of-pocket co-payments of consumers. Consequently, their consumer welfare concept captures the cost effects of a merger only to the extent of co-payments. They classify net social welfare as total welfare excluding hospital profits because those are hard to observe or verify. While the calculation of this and our extended CW concept coincides, the underlying idea is different. Our extended CW is not a kind of total welfare, since no profits of providers or insurers are included. If insurers made profit, we would not include it in the calculation. Extended CW is a consumer surplus taking into account all effects of a merger on consumers.

Change in total welfare is given by the following equation:

$$\Delta TW = (\Delta q - \Delta c)\left(\frac{1}{2} + \Delta q\right) - \frac{(\Delta q)^2}{2}$$

The first term represents the net gain from quality and cost effects of a merger while the second term is the loss from further traveling because of quality differences.

Considering non-health care markets, consumer and total welfare changes in the same direction and in the same instances (disregarding some special cases). Exceptions are, for example, the changes in fixed costs (not included in consumer welfare) and price discrimination among consumers (total welfare increases while consumer welfare decreases). The reason for the discrepancy between extended CW and TW in this model is the fixed demand ($\sum_i D_i = 1$) which is specific to health care markets. Because of the insurance market (and the absence of co-payments), patients do not react to price increases, i.e. there is no dead weight loss in the presence of prices above marginal cost. Changes in the price are purely redistributional in terms of gains between hospital sector and consumers; TW remains unchanged while extended CW changes.

The health care specific differences in welfare measures become larger if the group of consumers considered in the analysis is widened.[14] Patients directly affected by quality changes is the narrowest definition. We expect that they are less concerned with the potential costs of a merger (lowest S) because of the high externality effect. Widening the considered group, the potential clients of hospitals, consumers on the hospital market can be included in the analysis. The cost effects of the merger are considered to the extent this group will bear it. External effects due to insurance still can be present if insurers have patients on (and consequently spread health care expenditures across) more hospital markets.

[14] As explained in the introduction, we find the non-evident definition of consumer the specialty of health care markets. Different definitions of the consumer yield different result in the welfare analysis. Therefore considering hospital mergers requires careful analysis.

A complete consumer welfare analysis would embrace all consumers covered by the same insurance (extended CW). The only welfare effect excluded then is the profit of providers. TW considers this last aspect of welfare as well.

6.3.2 Welfare standards in merger control

Formally, the simple CW standard ignores providers' profit and the insurance externality. Ignoring providers' profit makes the standard stronger than the TW standard as can be seen from the results on the extended CW standard. The neglected externality can however be negative (when $\Delta c > 0$), which can outweigh the previous effect and make the simple CW standard weaker than the TW standard. In summary, the simple CW standard may both clear undesirable mergers and block desirable ones.

While previous literature found reasons why a stronger standard (CW) can be advantageous in merger control, we state that CW can also be too lenient in health care markets, which could lead to clearing undesirable mergers. We describe the intuition of the externality and provide a numerical example illustrated on Figures (6.1-6.2). In this way, we disprove that CW standard is always a stronger condition than TW standard. To disprove a claim on generality, a numerical example is sufficient. Notice that the example is by no means an anomaly. We do not provide the general parameter regions where simple CW standard is a weaker standard than TW standard, since that would be technically quite cumbersome and not needed to show the point. Our goal is to show that applying the simple CW standard in health care may have unexpected results.

Figure (6.1) shows in function of Δc and Δq where a change in TW, CW and merger related hospital gains turns out to be positive. The shaded area depicts merger alternatives that are profitable for purely-profit-maximizing hospitals. This is the set of merger proposals that a competition authority can expect. Below the TW line, the change in total welfare is positive. Applying TW standard, the authority would clear a merger in this parameter range. Applying CW standard, the authority approves every merger alternative below the CW line. On the left panel, CW is calculated as extended CW ($S = 1$), while on the right panel, CW is the simple CW with $S = \frac{1}{2}$. Figure (6.2) shows similar parameter ranges for partially-quality-maximizing hospitals ($\alpha = 0.7$).

As seen on the left panel of Figure (6.1), the extended CW standard is tougher than the TW standard. Since we have chosen TW as the benchmark, we can say that the extended CW standard commits type I errors: it rejects mergers that would increase TW. Besanko and Spulber Besanko and Spulber (1993) build on this characteristic of CW and state that it can contribute to the self-selection of merger proposals that increases total welfare. In the model of Lyons Lyons (2003), CW is again a tougher standard than TW. That model takes account

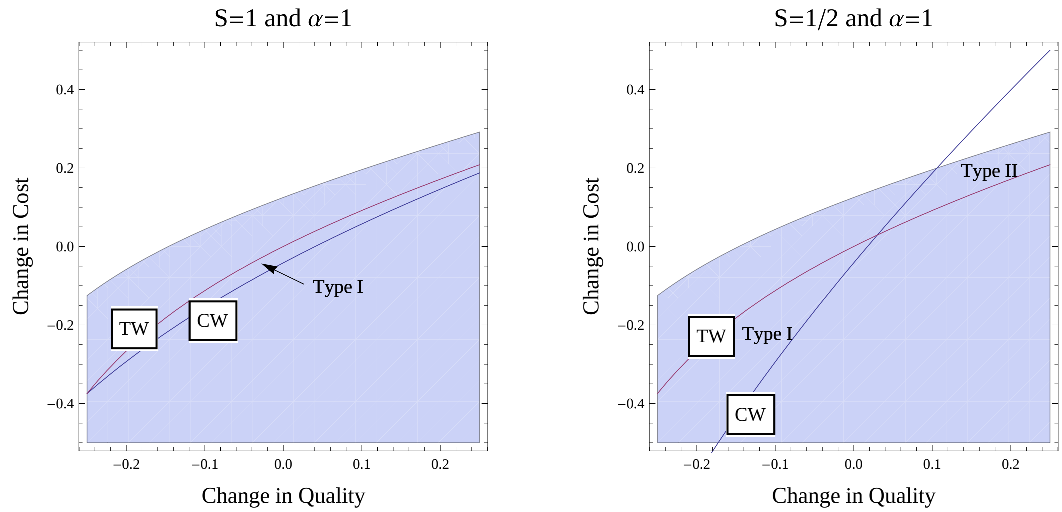


Figure 6.1: Parameter ranges in which the CW standard commits type I and type II errors, respectively. The shaded area depicts the parameter ranges where the merger is profitable for a profit-maximizing hospital ($\alpha = 1$). We assume no externalities ($S = 1$) in the left panel and externalities ($S = \frac{1}{2}$) in the right panel.

to alternative mergers and study the decision of a competition authority in a dynamic perspective. Under CW standard, the authority is more likely to reject a merger than under TW standard. This is an advantage when an alternative merger gives higher TW. Thus in both models, the CW standard can commit type I errors, but no type II errors; i.e., it may reject mergers that increase TW, but it does not approve mergers that decrease TW. This idea coincides with our results on extended CW standard. Simple CW standard, however, may commit type II errors as well as shown in the right panel of Figure (6.1). This merger standard accepts higher cost increase for given level of quality improvement than the TW standard and in this way can wave socially undesirable, costly mergers. The mechanisms described in Besanko and Spulber Besanko and Spulber (1993) and Lyons Lyons (2003) do not work any more; simple CW standard cannot be preferable to TW standard.

The probability of type II errors is, however, limited as long as hospitals show profit-maximizing behavior as in Figure (6.1). When hospitals intrinsically value quality as in Figure (6.2), they tend to accept more cost increase in turn of a given level of quality improvement. As can be seen, this is the region where the simple CW standard commits type II errors. In other words, there is more chance that the simple CW standard approves socially undesirable mergers when hospitals are (partially) quality maximizing than in case of profit-maximizing hospitals. The probability of committing type I error, in contrast, decreases. Quality-maximizing hospitals initiate less mergers that lowers quality. In this respect, quality-maximizing hospitals are more restrictive in proposing a merger than the standard of the competition authority and that is why there are less proposals that the standard rejects.

Regarding to the extended CW standard in case of partially-quality-maximizing hospitals (left panel of Figure 6.2), it does not commit Type II error. Because the probability of Type I error is also decreasing due to hospitals' behavior, the decisions based on TW and extended CW welfare standard converge.

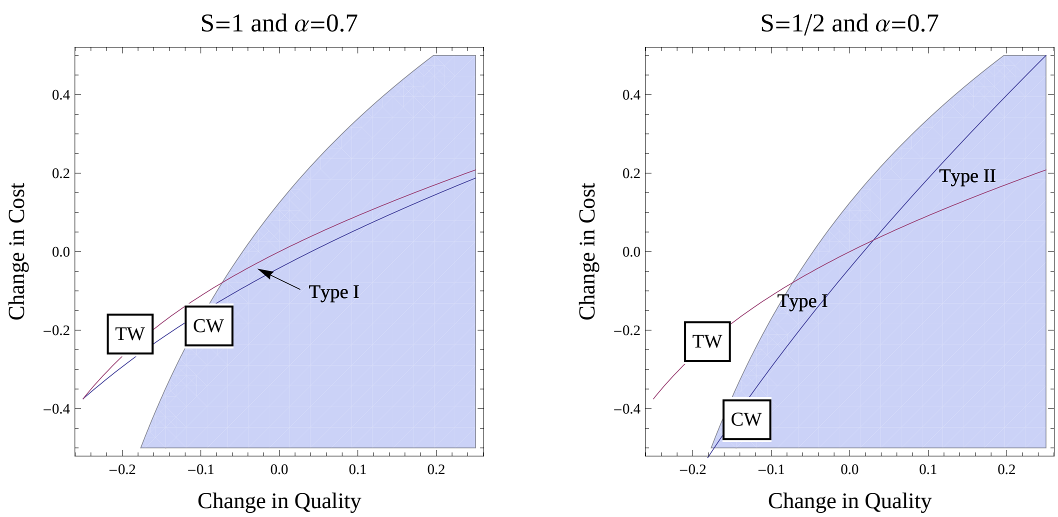


Figure 6.2: Parameter ranges in which the CW standard commits type I and type II errors, respectively. The shaded area depicts the parameter ranges where the merger is profitable for a partially-quality-maximizing hospital ($\alpha = 0.7$). We assume no externalities ($S = 1$) in the left panel and externalities ($S = \frac{1}{2}$) in the right panel.

In conclusion, quality-maximizing behavior of hospitals has similar effects to the application of the simple CW standard, i.e. the cost effects play a less role in decisions. Scrutinizing whether cost increase is proportional to quality improvement is thus essential when hospitals attach high intrinsic value to quality. The simple CW standard does not concern the complete cost effects of a merger; therefore, this standard may not be hard enough to block costly, quality improving mergers. The use of extended CW is essential in this case because the external effects ignored by the simple CW are significant.

As shown, we can draw a parallel between a merger analysis based on extended CW (including the external effects) in hospital mergers and a merger analysis based on CW standard in non-health care mergers. Consequently the advantages and disadvantages of extended CW standard in case of hospital mergers compared to TW standard coincide with that of the CW standard compared to TW standard in non-health care mergers since explicit consideration of externality is in none of the cases needed. The results of general literature on the application of CW or TW in merger analysis, summarized in 6.1.1, apply to the considerations on extended CW standard versus TW standard in hospital mergers. We argue here for applying extended CW or TW standard for health care markets, but avoiding the usage of the simple CW standard, which neglects significant external effects.

6.4 Application in merger control

Most experiences with merger control in the hospital market have been in the US, Germany, and the Netherlands. In order to assess the practical value of our analysis, we briefly review how our findings can be related to the current practice. We also help the interpretation of our suggestions by a short hypothetical example.

It is important to note that there is some discrepancy between the approach of theoretical models to merger analysis and the practice of competition authorities. First, competition authorities only make the trade-off between positive and negative effects required by the welfare standard analysis when the expectation is that the merger significantly harms competition on the relevant market. The logic behind this is that a more ‘intrusive’ or far-reaching merger control would require information that authorities typically do not have. Second, in many jurisdictions, again for understandable and pragmatic reasons, most merger decisions are made by and large on purely legalistic grounds.[15] For example, if market shares are deemed low, a merger is waived without looking at the substantive side of a case, implicitly assuming that mergers between firms with relatively low market shares are unlikely to cause problems. This implies that discussions on the choice of welfare standards are only relevant for those cases where a real welfare analysis is made.

[15] Here we refer to the sequential approach in welfare analysis used, for example, in the EU, the UK and Australia. In the first step, general presumptions (e.g., on market shares) are applied to filter out unambiguous cases, i.e., when a positive or negative outcome of the analysis is highly probable. A thorough analysis is made only to the remaining more ambiguous cases. The US applies a case-by-case approach, i.e., every merger case is subject to a welfare analysis Renckens (2007).

In practice, the primary concern of competition authorities in hospital merger cases is the expected price changes on the hospital market. Although other aspects of competition such as quality might be considered as well,[16] evaluations of the effects of mergers are in first instance based on the expected effects on the financial side of the market. Furthermore, effects are not geared at consumers (i.e., to the premiums), but the analysis stops at the intermediate stage of insurer (i.e., hospital prices). As shown in this article, price effects at this intermediate level do not include externalities since the insurer has to bare the entire health care costs of enrollees. Using the notation of our model, the calculations address in first instance the expenditure of insurers ($\sum_{i=1}^{n} D_i w_i$) and not the consumer welfare ($PW - S\sum_{i=1}^{n} D_i w_i$). Consequently, the calculations of competition authorities are not distorted as long as the merger does not effect the patient welfare.

Concerns may arise when positive effects of a merger have to be weighed against negative price effects. Merging parties may want to show through an efficiency defense that benefits from quality improvements outweigh the welfare losses of possible price increase. Notice that the burden of proof that quality improvements will emerge lies with the merging parties, but authorities have to trade off these gains against the competitive harm of the merger[17] The exact definition of the 'consumer' then becomes essential in order to calculate the exact gains for consumers. When calculating the PW derived from improved quality, there are two essential questions to answer: who benefits from the quality improvement and what is their willingness to pay for it. Competition authorities should be aware that through health insurance a large group of patients is affected by the merger but their willingness to pay for potential quality improvements probably varies. By calculating the financial effects of a merger, the 'consumer' is defined as all enrollees since the price effect is only calculated on the insurers' level and not derived to consumers. Considering the quality effects, many of these consumers may not benefit directly from the potential quality improvement and have, therefore, low valuation for it.

The fact that clients of an insurer do not equally benefit from a given quality improvement of a provider can be captured in the divergence of their reported willingness to pay for it. Diener Diener et al. (1998) distinguishes three sources of willingness to pay that can also be related to the different consumer groups

[16] Quality as an attribute of competition is mentioned only in the third of court cases in the US between 1985 and 1999 Hammer and Sage (2002).

[17] Efficiencies have to fulfill three criteria in order to be taken into account by the authorities: they have to benefit consumers, be merger specific and be verifiable Commission et al. (2004). Even if these requirements are fulfilled, there are some clear-cut cases when explicit trade-off of the price and quality effects are not necessary (e.g., a marginal quality improvement that increases the costs substantially). Weighing positive and negative effects against each other is the most relevant when both are significant and the net effect cannot be foreseen without a thorough analysis.

in an insurer's population.[18] First, people value a good or service because of its 'use value'. If they directly consume it, they are probably willing to pay a certain price for it. Second, people may expect that in the future they probably will need the good or service. In this case they may be willing to pay an 'option value' for the possibility of access in the future. Finally, consumption of some goods and services has external effects. For example, high grade of vaccination in a population gives protection also to those not receiving the vaccine. People may be willing to pay for such 'externality value' of a good or service. Smith Smith (2007) shows that these three values differ. Although use value dominates, option value and externality value are also a significant source of the total value. Clients of an insurer are willing to pay either use or the option value dependent whether they are current patients of the hospital. In a merger control process is therefore necessary to consider the heterogeneities in willingness to pay in order to come to an appropriate aggregation of patient benefits owing to the merger.

To illustrate this point, take the example of two hospitals that promise to invest in the latest technology in their dialysis center after they merge. We assume that this implies a real quality improvement in their dialysis treatment, which cannot be achieved without the merger due to scale efficiencies. The competition authority concludes from its research that the concentration is likely to considerably lessen competition on the hospital market. The merging parties however claim that this negative price effect would be compensated by the improved quality for consumers (the superior technology in the dialysis center). How should the authority trade off the price increase and the quality improvement?[19] Clearly, patients currently visiting the dialysis center are going to value the improvements (use value). Consumers on the given hospital market are also going to value it because might they need kidney dialysis, they can expect a better quality service in their hospital. Their valuation (option value) is however lower. Consumers on other hospital markets, but in the same insurance pool, do not benefit from the improvement; and therefore, have a very low (or zero) valuation. The competition authority has to take into account this diversity in willingness to pay when calculating the aggregated gains for consumers. It could be a pitfall to measure the willingness to pay of patients directly benefiting from the quality improvement and generalize it to all consumers on the hospital market. Increase in PW calculated in this way overestimates the real benefits

[18] Diener discusses in his article the application of contingent valuation methodology in health care. The contingent valuation method aims to elicit information on willingness to pay for, for example, a health care program or medical device. In this sense, the principals of eliciting the correct willingness to pay in the population raises the same question as applying the appropriate price-quality trade-off in merger control.

[19] This is a hypothetical example. To our best knowledge, quality and price effects have not yet been explicitly set against each other by competition authorities.

of consumers. It should not be automatically assumed that all patients on the relevant hospital market have the same willingness to pay. Therefore, it should be scrutinized which consumer groups are included in the analysis and how their welfare is aggregated.

The important conclusion of this article is twofold for the legal practice. First, potential gains for consumers have to be aggregated across the whole group covered by insurance. The reason is that the financial side, the price effects are also calculated for this group because the price of hospital services is considered in the analysis rather than the insurance premium. Second, consumers probably attach diverging values to given improvements in health care according to their current situation and expected future needs, which has to be taken into account in the analysis.

6.5 Conclusions

Our paper contributes in four ways to the literature. First, we have shown that consumer welfare in health care markets should be interpreted in a different way than in standard markets. This difference is prompted by the fact that mergers may have consequences for other consumers than just the ones directly affected. The indirect effect runs through health insurance premiums, i.e., if after a merger the bargaining position of the merged entity drives the premium up, all clients of the insurers are affected, not just patients of the merged hospital. This external effect on consumers paying premium but actually not receiving health care has to be taken into account in the welfare standard of competition authorities.

Second, we have more specifically shown that failing to incorporate the above-mentioned external effect can make CW in case of quality enhancing mergers a weaker standard than total welfare. This result contrasts with results in markets without significant externalities. Consumer welfare standard in such markets is always at least as tough as total welfare standard. A further contrast to non-health care markets, i.e. a consequence of forgoing external effects is that simple CW can commit both type I and type II errors in merger control. If the externality is positive (e.g., owing to a cost reducing merger), TW enhancing mergers can be blocked, while if the externality is negative (e.g., owing to a cost increasing merger that raises premiums), TW reducing mergers that should be blocked can be waived. A CW standard that includes the external effects too (i.e., the extended CW) corrects for these disadvantages, and it is similar in results to CW in non-health care markets.

Third, the non-profit nature of hospitals increases the probability that quality-improving mergers are proposed since hospitals put much weight on the intrinsic value of quality improvement. If these improvements go hand in hand with high costs, TW can be negatively affected. The more weight hospitals put on quality compared to monetary profit in their merger decisions, the more important it becomes to apply a merger standard that involves the complete cost and quality effects. In conclusion, simple CW standard is not suitable for health care markets; extended CW should be used instead.

Finally, the distinction between simple and extended CW in the practice of merger control can be reflected in the exact definition of the 'consumer' used in the welfare analysis. The choice of definition is particularly important when a merger has both price and quality effects because the group of consumers directly affected by quality improvements (or deteriorations) usually does not correspond with the population bearing the financial consequences of the merger. The externality caused by insurance is reflected in the diverging valuation of quality improvements by different groups of consumers. The conclusions for merger control in health care markets is that it is essential to scrutinize which consumer groups are involved in the welfare analysis and how their benefits (or losses) from the merger are aggregated.

Acknoledgements

We would like to thank Jan Boone for his many valuable suggestions on the model set-up. We also thank Wolf Sauter, Rein Halbersma, Ramsis Croes, Frank Pellikaan and Krijn Schep for their helpful comments on drafts of this paper. Furthermore, we are grateful to the participants of the Competition Law and Economics European Network (CLEEN) workshop 2011, participants of the Centre for Competition Policy (CCP), University of East Anglia workshop 2011, and to colleagues at TILEC, Tilburg University and at the Dutch Healthcare Authority for their comments and suggestions.

Bibliography

Begg, D. K. and Vernasca, G. *Economics*. McGraw-Hill Higher Education, 9th edition, 2008.

Besanko, D. and Spulber, D. F. Contested mergers and equilibrium antitrust policy. *The Journal of Law, Economics, and Organization*, 9(1):1–29, 1993.

Calem, P. S., Dor, A., and Rizzo, J. A. The welfare effects of mergers in the hospital industry. *Journal of Economics and Business*, 51(3):197–213, 1999.

Capps, C., Dranove, D., and Satterthwaite, M. Competition and market power in option demand markets. *RAND Journal of Economics*, pages 737–763, 2003.

Chang, T. and Jacobson, M. What is the mission of a not-for-profit hospital? evidence from california's seismic retrofit mandate. Technical report, National Bureau of Economic Research, 2010.

Clement, J. P., White, K. R., and Valdmanis, V. Charity care: do not-for-profits influence for-profits? *Medical Care Research and Review*, 59(1):59–78, 2002.

Commission, E. et al. Guidelines on the assessment of horizontal mergers under the council regulation on the control of concentrations between undertakings. *Official Journal of the European Union C*, 31:5–18, 2004.

Deneffe, D. and Masson, R. T. What do not-for-profit hospitals maximize? *International Journal of Industrial Organization*, 20(4):461–492, 2002.

Diener, A., O'brien, B., and Gafni, A. Health care contingent valuation studies: a review and classification of the literature. *Health economics*, 7(4):313–326, 1998.

Fridolfsson, S.-O. A consumer surplus defense in merger control. Technical report, Research Institute of Industrial Economics, 2007.

Gaynor, M. and Vogt, W. B. Competition among hospitals. Technical report, National Bureau of Economic Research, 2003.

Hammer, P. J. and Sage, W. M. Antitrust, health care quality, and the courts. *Columbia Law Review*, pages 545–649, 2002.

Hart, O. and Moore, J. Property rights and the nature of the firm. *Journal of political economy*, 98(6):1119–1158, 1990.

Heyer, K. Welfare standards and merger analysis: Why not the best? Technical report, Department of Justice, Antitrust Division, 2006.

Horwitz, J. R. Making profits and providing care: comparing nonprofit, for-profit, and government hospitals. *Health affairs*, 24(3):790–801, 2005.

Horwitz, J. R. and Nichols, A. Hospital ownership and medical services: market mix, spillover effects, and nonprofit objectives. *Journal of health economics*, 28(5):924–937, 2009.

Katona, K. and Canoy, M. Welfare standards in hospital mergers. *The European Journal of Health Economics*, 14(4):573–586, 2013.

Lakdawalla, D. and Philipson, T. Nonprofit production and competition. Technical report, National Bureau of Economic Research, 1998.

Lyons, B. R. Could politicians be more right than economists? a theory of merger standards. Technical report, 2003.

Malani, A., Philipson, T., and David, G. Theories of firm behavior in the non-profit sector. a synthesis and empirical evaluation. In *The governance of not-for-profit organizations*, pages 181–216. University of Chicago Press, 2003.

McGuirk, M. A. and Porell, F. W. Spatial patterns of hospital utilization: the impact of distance and time. *Inquiry*, pages 84–95, 1984.

Motta, M. *Competition policy: theory and practice*. Cambridge University Press, 2004.

Neven, D. J. and Röller, L.-H. Consumer surplus vs. welfare standard in a political economy model of merger control. *International Journal of Industrial Organization*, 23(9-10):829–848, 2005.

Philipson, T. J. and Posner, R. A. Antitrust in the not-for-profit sector. Technical report, National Bureau of Economic Research, 2006.

Pittman, R. Consumer surplus as the appropriate standard for antitrust enforcement. Technical report, Department of Justice, Antitrust Division, 2007.

Prüfer, J. Competition and mergers among nonprofits. *Journal of Competition Law and Economics*, 7(1):69–92, 2010.

Renckens, A. Welfare standards, substantive tests, and efficiency considerations in merger policy: defining the efficiency defense. *Journal of Competition Law and Economics*, 3(2):149–179, 2007.

Richman, B. D. Antitrust and nonprofit hospital mergers: a return to basics. *University of Pennsylvania Law Review*, pages 121–150, 2007.

Sage, W. M., Hyman, D. A., and Greenberg, W. Why competition law matters to health care quality. *Health Affairs*, 22(2):31–44, 2003.

Salop, S. C. Question: What is the real and proper antitrust welfare standard? answer: The true consumer welfare standard. Technical report, Statement before the Antitrust Modernization Commission, 2005.

Smith, R. D. Use, option and externality values: are contingent valuation studies in health care mis-specified? *Health Economics*, 16(8):861–869, 2007.

Vita, M. G. and Sacher, S. The competitive effects of not-for-profit hospital mergers: a case study. *The Journal of Industrial Economics*, 49(1):63–84, 2001.