UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

A STRUCTURAL MODEL OF HOSPITAL AND INSURER COMPETITION

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN ECONOMÍA APLICADA

BENJAMÍN VATTER JUNEMANN

PROFESOR GUÍA:
CARLOS NOTON NORAMBUENA

MIEMBROS DE LA COMISIÓN:
JUAN ESCOBAR CASTRO
MARCELO OLIVARES ACUÑA

SANTIAGO DE CHILE
2016

RESUMEN DE LA TESIS PARA OPTAR
AL TÍTULO DE MAGÍSTER EN ECONOMÍA APLICADA
POR: BENJAMÍN VATTER JUNEMANN
FECHA: 2016
PROF. GUÍA: SR. CARLOS NOTON NORAMBUENA

A STRUCTURAL MODEL OF HOSPITAL AND INSURER COMPETITION

Este trabajo estudia el mercado de la salud privada en Chile, enfocándose principalmente en los efectos de la integración vertical entre clínicas e isapres sobre los precios del mercado y el bienestar de los consumidores. Se formula un modelo estructural del mercado, incluyendo la negociación entre clínicas e isapres y la competencia por clientes en ambos lados del mercado. Utilizando datos detallados provenientes de la Superintendencia de Salud de Chile, el modelo es estimado con el fin de recuperar los parámetros que gobiernan la demanda y los costos marginales de las clínicas principales de Santiago. Utilizando esto, ejercicios contrafactuales son realizados para evaluar el impacto de la integración vertical y horizontal en el mercado.

El principal aporte del trabajo está en extender la literatura sobre mercados verticales y oligopolios bilaterales, incluyendo un modelamiento explicito de la integración vertical. Una simple caracterización del equilibrio resultante de la competencia entre isapres es presentada, mostrando que puede ser resuelto mediante métodos de punto fijo, reduciendo el costo computacional de este modelo y facilitando el análisis económico de las dinámicas de mercado. Adicionalmente, el trabajo vincula los precios de atenciones médicas y costos hospitalarios mediante una ecuación que incorpora las discrepancias entre las partes negociantes y que permite resolver el modelo numéricamente.

Los resultados muestran que los consumidores son substancialmente más sensibles a las primas de los planes de salud que a los precios de las atenciones médicas. Dada la alta concentración del mercado de las isapres y su capacidad para optimizar sus primas posterior a las negociaciones de precios, el modelo encuentra que las isapres son más flexibles que las clínicas a la hora de optimizar sus ingresos, entregándoles una posición ventajosa a la hora de negociar precios. Los ejercicios contrafactuales muestran que estos dos resultados implican que las clínicas integradas utilizan la flexibilidad de sus isapres para negociar margenes más altos, permitiendo a las isapres aumentar sus ingresos aprovechando la demanda más inelástica por atención médica. Los resultados preliminares indican que de prohibirse la integración vertical, los precios de atenciones médicas bajarían en aproximadamente 25% en promedio, sin modificar significativamente las primas de planes de salud. Eliminar adicionalmente la integración horizontal presenta pocos beneficios en términos de precios promedio y bienestar.

RESUMEN DE LA TESIS PARA OPTAR
AL TÍTULO DE MAGÍSTER EN ECONOMÍA APLICADA
POR: BENJAMÍN VATTER JUNEMANN
FECHA: 2016
PROF. GUÍA: SR. CARLOS NOTON NORAMBUENA

A STRUCTURAL MODEL OF HOSPITAL AND INSURER COMPETITION

This work studies the private health care market in Chile, focusing mainly on the pricing and welfare effects of vertical integration. A structural model is formulated, which incorporates the bargaining between insurers and hospitals for service prices and the competition for clients on both sides of the market. The model is estimated using detailed data from the Chilean regulatory agency (Superintendencia de Salud) in order to identify the parameters of the demand and the marginal costs of hospitals. Using this, a counterfactual analysis is done in order to assess the effects of vertical and horizontal integration on the market.

The main contribution of this work is in extending the literature on vertical markets and bilateral oligopolies, including an explicit model of vertical integration. A simple characterization of the equilibrium of insurer competition is presented showing that it can be resolved using fixed-point methods, reducing the computational cost of the model and facilitating the economic analysis of the market's dynamics. Additionally, this work links medical prices and costs using a familiar equation that incorporates the discrepancies resulting from negotiation, allowing the model to be solved numerically.

The results show that consumer as substantially more sensitive to premiums than to medical prices. Given the high concentration of the insurers market and their capacity of optimizing premiums after prices have been negotiated, the model finds that insurers are more flexible than clinics when optimizing profits, giving them an advantageous position in the bargaining process. The counterfactual analysis indicates that these two results imply that vertically integrated hospitals exploit their insurer's flexibility to negotiate higher markups and allows integrated insurers to increase profits using the more inelastic demand for medical care. Preliminary results show that eliminating vertical integration from the market would results in an average reduction of medical prices of approximately 25%, without substantially changing premiums. Further restricting horizontal integration among hospitals appears to have little effect on overall prices and welfare.

A Naomi.

Agradecimientos

Agradezco el apoyo y consejo de Carlos Noton, Gautam Gowrisankaran, Sebastian Fleitas, Kenneth Train, Roberto Cominetti y a los miembros de la comisión Juan Escobar y Marcelo Olivares.
Se agradece el apoyo financiero del Millennium Institute for Market Imperfections & public Policy (MIPP), ICM IS130002. Además se agradece el apoyo del National Laboratory for High Performance Computing (NLHPC) [1].

[1]University of Chile. Powered@NLHPC: This research was partially supported by the supercomputing infrastructure of the NLHPC (ECM-02)

Contents

Chapter 1

Introduction

This paper studies a model of the health insurance market, in which insurers act as intermediaries between consumers and health care providers. Health insurers negotiate service prices with providers and later compete by setting premiums to a collection of health insurance plans. Building upon recent work by Gowrisankaran et al. (2015) and Ho and Lee (2013), I formulate a framework to study the interactions between upstream and downstream competition in the health care market. The model allows the analysis of both vertical and horizontal mergers in the market, while accounting for the downstream competition in premiums. This structural model is estimated using individual level data from the Chilean private health insurance market. The results are used to conduct a counter-factual analysis to study the effects of upstream vertical integration.

The main contribution of this work is in providing the first extension to this class of models that allows Bertrand-Nash competition in premiums to take place after upstream negotiations. Additionally, it clarifies how equilibrium negotiated prices and plan premiums are affected by vertical integration. Gowrisankaran et al. (2015) and Ho and Lee (2013) recognize that the problem at hand is prohibitively computationally intensive. I propose a handy simplifying assumption which greatly reduces computation time. The assumption relies on the fact that health insurers in Chile are allowed to price insurance plans on age and gender. Together with coverage denial, the market effectively segregates consumers into smaller sub-markets in which downstream competition is held. Given the scarce variation in consumer attributes within sub-markets, I assume that insurers consider only the mean consumer attributes in each sub-market. This assumption allows to solve the equilibrium premium using Lambert's W function [1], which provides a close expression for solving premium derivatives as an in inversion problem. Although the assumption might not fit every scenario, the application of the Lambert W is (to my knowledge) new to this literature and proves to be helpful both for analysis and computation. Based on Gowrisankaran et al. (2015), I believe this model can be generally extended and applied to study markets in which intermediaries negotiate base prices with providers and later compete on prices over a series of sub-markets,

[1] The Lambert W functions are the solutions to the equation $x\exp(x) = k$ for x. Though multivalued, the function has a unique real branch for $x >= 0$, often called W_0 or *main branch*. The Lambert W function is standard in Physics and Chemistry and efficient routines for its calculations are provided with all major mathematical libraries, including Matlab, GSL (C) and Python's Scipy.

homogeneous markets or markets in which little information regarding consumer attributes is available to the intermediaries.

The health care market has been the focus of large reforms during the last decade (2011 reforms to the NHS in the UK or the Affordable Care Act in the US). Given the large private and public expenditure in this market, along with its significant effect on people's lives, it is important to understand the way market participants would react to policy changes. As pointed out by Ho and Lee (2013), the mechanics of the market can sometimes lead to counter-intuitive results. For example, increasing downstream competition can reduce insurers leverage while negotiating with providers, which will result in increased service prices and in some cases further lead to increased premiums to consumers, decreasing both insurer and consumer surplus. Hence to guide policy reforms appropriate models of the market are required. In order to gain tractability some aspects of the market are simplified in this work. First of all, the model is static and therefore omits dynamic effects such as lock-in (Atal, 2015) or reclassification risk (Handel et al., 2015). Also, plans are exogenous with coverage and network given throughout the paper. Hence aspects such as coinsurance rates or preferential access to providers are not optimized by insurers during downstream competition and are assumed to be a result of long run equilibrium [2]. Finally, equilibrium insurer-provider networks are considered to be fixed, hence network formations dynamics such as studied by Lee and Fong (2013) are also omitted.

The model is a four stage game. In the first stage, upstream negotiations take place. Insurers and hospital systems (i.e, companies that administrate one or more health care providers) bargain over service prices for the insurer in each of the system's hospitals. Offers follow the sequential *take it or leave it* protocol, with the solution given by standard Nash Bargaining (Horn and Wolinsky (1988); Crawford and Yurukoglu (2012); Collard-Wexler et al. (2015)). All negotiations take place simultaneously. Vertically integrated insurers and hospital systems also optimize service prices during this round. The bargained service price is a scalar for each insurer-hospital pair which defines the base price of service for all medical conditions. This service price is then multiplied by a condition specific weight in order to determine the final service price. Condition weights are identical across insurers and represent common knowledge on the resource intensity of each medical condition[3].

Downstream competition in prices takes place in the second stage. Insurers compete by simultaneously setting premiums for a collection of plans, seeking to maximize their profit. Each plan is offered in a particular sub-market, based on age and gender. Because sub-markets are exhaustive and exclusive, consumers can not substitute into plans offered in a different sub-market and therefore the downstream competition is effectively split into several smaller markets. Under this segmentation, I will refer to sub-markets simply as markets from now on, understanding that each market is defined by it's consumers gender, age group, and whether they seek a family or individual plan [4]. In each of the markets, insurers optimize premiums taking into account only the attributes of the mean consumer which, given the

[2] To my knowledge there is no paper that endogenizes plan attributes within this setting.

[3] Although in this work weight are directly calculated from the data, similar weights are used by Medicaid to price inpatient payment as explained by Gowrisankaran et al. (2015)

[4] Given that family plan are far more convenient than purchasing separate insurance for each member, I consider this also segments the market in the same way as age or gender does.

large number of markets, can be seen as a discrete approximation to the distribution of the population.

In the third stage, consumers choose health insurance plans knowing prices and premiums. Each consumer chooses among the available plans, taking into account the expected utility from each plan for the entire household and considering each member's respective risk of illness. Payers are obliged to purchase insurance worth at least 7% of their income (up to a high threshold) and therefore consider only the difference between each plan's premium and that legal amount when evaluating a plan's cost. Public insurance at the mandatory amount is always available for the consumer, which has a limited network restricted to the public hospital system.

Finally, in the fourth stage, each consumer receives a health draw. If ill, consumers must choose which hospital to attend based on the coverage and network provided by their plan. Out-of-pocket payment is determined as a function of the price negotiated between the insurer covering the consumer and the chosen hospital, multiplied by the condition weight and the plan's coinsurance rate, which is hospital-condition specific. Additional utility shocks are drawn from the location of the hospital, whether it is the plan's preferential provider (if the plan has one) and other provider specific unobservables, such as number of emergency beds or medical staff. The choices determined in this round define the expected utility of each plan to each consumer, becoming part of the demand for plans, which in turn defines the demand equation for the downstream premium equilibrium.

By explicitly solving the Bertrand pricing game of the second stage, conditional on the demand and negotiated prices, I find an expression for the premiums setting. This solution shows that all premiums are set based on a common factor accounting for the general price elasticity of consumers. On top of this value, plans are set to cover their expected cost and their owners opportunity cost of including them in the market. This opportunity cost extends to vertically integrated firms and represents the profit each firm would obtain in the absence of the given plan in equilibrium. An additional adjustment term is added to this expression that holds the information regarding the substitution patterns implied by the logit demand. This solution is one of the key results of the model as it allows for quick evaluation of equilibrium premiums, disagreement premiums and premium derivatives with respect to negotiated prices.

Next, by using the first order condition of the Nash Bargaining problem and including the optimal price set by vertically integrated firms, I generalize the marginal cost equation specified in Gowrisankaran et al. (2015) to allow for vertical integration, consumer changes of plans and disagreement adjustments to premiums. This equation is the bargaining analog of the standard techniques used in Berry et al. (1995) and Bresnahan (1987) for backing-out marginal cost from equilibrium behavior.

The model is estimated in three stages. First, I estimate hospital demand conditional on plan and medical status using Train (2015) method's. This estimation routine is similar in spirit to normal mixed logit estimation, in which a flexible mixing distribution that approaches a Gaussian form is used to provide the joint distribution of all parameters. This method allows for heterogeneity in preferences and substitution patterns, while permitting the code to be parallelized easily and solved more efficiently than standard mixed logit.

Train's method, to which I will refer to as *flexible mixed logit*, is also used to estimate plan demand. This demand is set conditional to each consumer's market and restricted to the plans that cost no less than her mandatory amount, corresponding to 7% of her income. The demand estimations use very large and detailed datasets and are considered as Big Data problems [5], demanding this departure from the traditional methods of random coefficient logit estimation.

Finally, the third step consists in recovering the hospital's marginal costs of treating patients from a given insurer. Additionally, Nash bargaining weights are estimated in order to fully recover the pricing problem. I estimate this model using moments conditions based on orthogonality restrictions on marginal cost. This estimation is self consistent in all of the model's variables, namely premiums, prices, shares, plan expected utility to consumers and plan expected costs.

I find that consumers are significantly more sensitive to premiums than to hospital prices, allowing vertically integrated firms to increase their profit through upstream prices to consumers without incurring the losses of higher costs for the insurer. A counterfactual analysis is done in order to evaluate the effect of vertical integration in the market, showing that hospital prices would decrease in about 25% if vertical integration were prohibited. However, as premiums vary insignificantly with the change and consumers make scarce use of their insurance, consumers would be willing to pay only 0.04 UF per month on average (or $18 USD per year) in order to prohibit vertical integration in the market.

The main contribution of the paper is to the literature on vertical markets and bilateral oligopoly with bargaining. The work of Ho and Lee (2013) is probably the closest in spirit to this paper, with both studying the health care market and giving a particular focus to the effects through which prices and premiums are set. However,the models are different since Ho and Lee (2013) assumes simultaneous premium setting and service price bargaining, which forbids adjustments to premiums under disagreement and removes the reaction of optimal premiums to marginal increases of negotiated prices. The estimation procedure also differs, with their paper estimating demand and bargaining simultaneously, and adding a difference between the real price elasticity of consumers and the one perceived by insurers and hospitals. Gowrisankaran et al. (2015) also estimate a structural model of bargaining in the health insurance market with downstream premium competition. However, in their work insurers don't compete for consumer in the downstream market, hence hospitals can't recapture consumers under disagreement and countervailing power effects are omitted. Both of these papers are related to the work by Crawford and Yurukoglu (2012), which was the first to estimate a structural model of bargaining based on Horn and Wolinsky (1988) with downstream premium competition. In particular, they study the bargaining between content providers and cable companies. Much of the theoretical modeling in this papers follows their work.

Additionally, this paper is related to the literature that estimates bargaining models in

[5]Though there is no consensus regarding the definition of what is Big Data, I stick with the common definition that states that it's data large enough so that standard methods are limited by system limitations. In this case, standard mixed logit requires storing multiple large matrices for algebraic operations that, given the size of the data, can not fit in contiguous memory in a modern high performance computer.

which hospitals negotiate with insurers for inclusion in their network. Capps et al. (2003) use a model which is consistent with bargaining without estimating a structural specification of bargaining. Ho (2009) studies how insurers determine their provider network by focusing on the role of different networks on downstream competition. In contrast, this paper takes networks as given and studies how service price negotiations and premium competitions are connected.

Finally, this work builds on a recent literature that provides method for estimating flexible demand patterns with large data. I use the method developed by Train (2015) which I find to closely approximate standard normally distributed mixed logit under a series of Monte Carlo experiments. Additionally the method allows for distributed memory parallel programming and also moves most of the computational cost to the first iteration in which a grid of possible consumer utilities is formed, making the following iterations quick and reducing communication overheads between nodes. Related are the works of Bajari et al. (2007) and Fox et al. (2011) which also provide a mixing distribution approach to estimation, with the last one being linear in estimated parameters. However, Train's method can be seen as a generalization of these papers and allows for more flexibility when setting up the mixing distribution, granting the researcher the possibility of easily changing it when evaluating alternative demand specifications.

The rest of the paper is structured as follows. Section 2 provides details regarding the Chilean institutional framework. Section 3 presents the theoretical model. Section 4 discusses the data used for estimation. Section 5 shows the estimation procedure and estimated parameters. Section 6 presents the counterfactual analysis and results. Section 7 concludes.

Chapter 2

Institutional framework

The Chilean law makes the enrollment in health insurance mandatory for all individuals [1]. In fact, workers and retirees have the obligation to spend at least 7 percent of their wages (upto a cap of approximately 186 USD per month) into the purchase of an insurance plan for themselves and their dependents. The insurance system is divided into a public and a private regime. The first, called FONASA, is a pay-as-you-go system mostly financed by public resources and is set as the default coverage for all otherwise uninsured individuals. FONASA charges no more than the mandatory percentage, regardless of income, and provides copayment that ranges between 0 and 20 percent, depending on income and family composition. However, FONASA enrollees are mostly restricted to the public hospital system, which usually has longer waits for service and is known to be drastically under-financed in comparison to private hospitals. In contrast, private insurers, called ISAPRES, provide greater access to private providers at a varying copayment rate, and often charge additional fees on top of the mandatory 7 percent. Although the law allows private insurers to price plans exclusively on age and gender, about 40% of plans have a single insured individual and the average number of enrollees per plans is about 28, effectively allowing for price discrimination at an individual level. Furthermore, private insurers are allowed to deny coverage to individuals that are not currently enrolled with them and are not obliged to offer the same plans to similar individuals [2]. Given these aspects, its not surprising that FONASA covers the less wealthy, older and riskier fraction of the population, with about two thirds of the population enrolled in one of the four levels of public insurance. On the other hand, private insurers cover around 17 percent of the population, with the remainder being affiliated with special health-care systems such as those of the Armed Forces.

Private insurers are limited in their ability to change premiums. A 2005 reform to the health insurance system ("ley larga de ISAPRES") restricted the increment of a plans premium to be no larger than 1.3 times the average price increase for all enrollees of the given insurer, and no smaller than 0.7 times the same [3]. Hence, although single-consumer plans are

[1]For more details on the Chilean health care market refer to Atal (2015),Duarte (2012) and Palmucci and Dague (2015). This section is follows their description of the market.

[2]Most insurers ask for a "health declaration" before signing a new contract. The law allows ISAPRES to deny coverage of preexisting conditions to insured individuals during their first 18 month of enrollment.

[3]The law is formulated as weighted average price increase for all of the insurers plans, with the weight set

very common in the market, they are effective for discriminating consumers mostly at entry. Additionally, the Chilean law provides guaranteed renewability in all plans and allows for premium increases only once per year, informed prior to the plans renewal and applied only afterwards. Consumers are usually contractually obliged to stay in their plans only during the first year of enrollment, being free to change plans freely afterwards (subject to the risk of premium reclassification).

The plans provided by private insurers mostly differ on their provider network and copayment rates. Restricted network plans (such as HMO in the U.S.) are very uncommon, with most plans either offering a preferred provider network (similar to PPO in the U.S.) or a free choice of providers. Under both regimes consumers are allowed to attend all hospitals, paying higher prices for providers outside of their insurer's network. Preferred provider plans usually have a small set of private hospitals in which the insured receives lower copayment rates for ambulatory and inpatient care, and a high copayment for non-preferential providers. Free choice plans usually have lower average copayments for each provider, but higher than what preferential providers get under the previous type of plans. In addition to the percentage of medical expenses the plan covers, most contracts stipulate caps on each claim, which don't accumulate over visits. Furthermore, some plans include a cap on yearly payments, which limits the amount the insurer will cover in a year. Once any of these caps are reached, the consumer pays the entirety of the remaining expenses until either the next visit to a health provider in the first case, or the next year in the second.

Finally, the Chilean private health market is characterized by a high level of concentration [4]. Only 5 insurance companies control about 97% of the market in terms of enrolled individuals. Furthermore, four of the five largest insurers (in terms of enrollees) are known to be vertically integrated, with integrated hospitals accounting (as of 2012) for more than 42% of the total value of medical treatments in the market. Table 2.1 shows the list of insurers and hospitals that operate within the metropolitan region of Santiago and that are known to be integrated. Additionally, on average 32% of an insurer's costs is due to claims originating from its own hospitals, with this number increasing when focusing on inpatient care and when grouping horizontally integrated clinics.

by the number of consumer insured in each plan

[4]The following information is detailed in article provided by the regulatory agency, *Superintendencia de Salud*, dated December 2013 and publicly available online.

Table 2.1: Vertical integrations in the market (2012)

Insurer	**Hospital**
Banmédica - Vida Tres	Clínca Santa María
Banmédica - Vida Tres	Clínica Davila
Banmédica - Vida Tres	Clínica Vespucio
Banmédica - Vida Tres	Vidaintegra
Consalud	Red Megasalud
Consalud	Clínica Tabancura
Consalud	Clínica Avansalud
Consalud	Clínica Arauco Salud
Consalud	Clínica Bicentenario
Cruz Blanca	Integramédica
Masvida	Clínica Las Lilas
Masvida	Clínica Providencia

Chapter 3

Model

3.1 Overview

The model describes the interactions between upstream health care providers, downstream insurers and private households. Consumers purchase private insurance plans for themselves and their dependents in order to cover their potential risk of illness, taking into account their expected costs from hospital care. These costs depend on the service prices negotiated between the insurer currently representing the consumer and the hospital of choice. Hence health insurance plans are defined by their provider network , their monthly premium and their different copayments across medical conditions and hospitals.

The market is modeled through a four stage game. In the first stage insurers and hospitals negotiate or directly optimize service prices, depending if they are vertically integrated or not. In the second stage insurers optimize their plan premiums given the previous set of prices. Thirdly, households choose among the available plans in their market, knowing networks, premiums and service prices. Lastly, consumers receive a series of health draws and choose to which hospital to attend, conditional on their selected plan and medical condition. In both consumer stages there is an outside option available. In the plan-choice stage, consumers can always acquire the public health insurance which costs exactly 7% of their income and gives exclusive access to public hospitals [1], and in the hospital-choice stage they may choose to attend a public hospital, which may have a positive cost but is often cheaper than private hospitals. I assume, given the evidence available, that private insurers don't negotiate prices with the public system [2]. Though prices, premiums and plan choices are set once per year, as is common in the literature and with which the data seems to agree, consumers can attend hospitals multiple times every year. As in Gowrisankaran et al. (2015), I assume that the price of treating condition d at hospital j can be written as $w_{\mathrm{d}} p_{mj}$, where m is the insurer covering the consumer. This allows for the bargaining between the insurer and the hospital to be over a scalar p_{mj}, which corresponds to the base price of treatment. The multiplier w_{d}

[1] Although under some special cases publicly insures individuals might access private hospitals.

[2] Public hospital prices are directly posted by the public system and are the same for all insurers, including the public insurers. Although coverage might differ among insurance plans.

adjusts this price according to the resource intensity of condition d and is therefore referred to as the condition's difficulty weight.

3.2 Demand

I start by describing the model that governs the demand for plans and hospitals for a given individual i. I assume that consumers have a linear utility, with heterogeneous preferences and type-1 extreme value distributed unobserved shock ε, as is standard in most logit-based consumer models. Consider the case of consumer i, who has signed a yearly contract with insurer m by enrolling in plan k. The utility of this consumer of visiting hospital j to be treated for condition d is given by

$$u_{\mathrm{i}j\mathrm{d}} = X_{\mathrm{i}j\mathrm{d}}\beta_{\mathrm{i}} - \alpha_{2\mathrm{i}} c^{k}_{\mathrm{d}j} w_{\mathrm{d}} p_{mj} + \varepsilon_{\mathrm{i}j} \tag{3.1}$$

where $c^{k}_{\mathrm{d}j}$ is the coinsurance rate of plan k when going to j for treatment of d, w_{d} is the difficulty weight of condition d and p_{mj} is the base service price of insurer m at hospital j, which is either the result of the negotiation process or the direct optimization of a vertically integrated system. $X_{\mathrm{i}j\mathrm{d}}$ are consumer-hospital-condition attributes and interactions that might affect the consumer choice, such as distance between the consumers home and the hospital or the number of Labor and Delivery rooms that the provider has.

The outside good utility is given by the option of going to one of the public hospitals in the area, which has a possibly non-zero cost and is therefore described by

$$u_{\mathrm{i0d}} = -\alpha_{2\mathrm{i}} w_{\mathrm{d}} p_{0} + \varepsilon_{\mathrm{i0}}$$

where $p_{0\mathrm{d}}$ is the expected cost of treatment of condition d in the outside option. The value of the non-price attributes of the public outside option is normalized to zero, making the coefficients β_{i} relative to this normalization.

Using this formulation we can solve for the probability with which consumer i chooses hospital j when suffering condition d using the standard logit formula:

$$\sigma_{\mathrm{i}}(j|\mathrm{d}, k) = \frac{\exp(X_{\mathrm{i}j\mathrm{d}}\beta_{\mathrm{i}} - \alpha_{2\mathrm{i}} c^{k}_{\mathrm{d}j} w_{\mathrm{d}} p_{mj})}{\sum_{h\in\mathcal{N}_k} \exp(X_{\mathrm{i}h\mathrm{d}}\beta_{\mathrm{i}} - \alpha_{2\mathrm{i}} c^{k}_{\mathrm{d}h} w_{\mathrm{d}} p_{mh})} \tag{3.2}$$

Where $\mathcal{N}_k$ is the provider network of plan k, which includes the public hospital.

Thus, the expected utility of plan k (which belongs to insurer m) for consumer i is given by the sum over medical conditions $\mathrm{d} \in D$ and their respective risks f_{id}, of the expected maximum utility choice of hospital, given by the well known formula

$$W_{\mathrm{i}k} = \sum_{\mathrm{d}\in D} f_{\mathrm{id}} \left[\ln\Big(\sum_{j\in H_k} \exp(X_{\mathrm{i}j\mathrm{d}}\beta_{\mathrm{i}} - \alpha_{2\mathrm{i}} c^{k}_{\mathrm{d}j} w_{\mathrm{d}} p_{mj})\Big) - \gamma \right] \tag{3.3}$$

Where γ is the Euler constant which will be omitted from the following calculations as it cancels out in the choice probability equation.

A useful simplification of the previous equation can be achieved by multiplying and dividing within the logarithm by $\exp(-\alpha_2 p_{p0\mathrm{d}})$, allowing it to be written exclusively in terms of the choice probability of the outside good hospital which we observe in the data and whose price is not subject to the negotiation process:

$$W_{\mathrm{i}k} = -\sum_{\mathrm{d}\in D} f_{\mathrm{id}}(\alpha_{2\mathrm{i}} p_{0\mathrm{d}} + \ln(\sigma_{\mathrm{i}}(0|\mathrm{d}, k))$$

Denoting plan k's premium as ϕ_k and consumer i's income as y_{i}, and given the assumption made regarding the consumer's utility functional form and the consumer obligation of purchasing a plan covering at least 7% of his income, the total utility of subscribing to plan k for patient i is given by

$$U_{\mathrm{i}k} = \alpha_{1\mathrm{i}}(\phi_k - 0.07 * y_{\mathrm{i}}) + \alpha_{3\mathrm{i}} \sum_{j\in F} W_{jk} + \alpha_{4\mathrm{i}} z_k + \xi_k + \varepsilon_{\mathrm{i}k} \tag{3.4}$$

Where z_k are observed attributes of the health plan that are unrelated to coverage, such as telephone assistance or a plan administrator. ξ_k are unobserved attributes of the plan, which will be the econometric residual, and $\varepsilon_{\mathrm{i}k}$ is an i.i.d type 1 extreme value error that accounts for transitory and uncorrelated unobservable shocks to the consumer's demand for the given plan, such as being contacted by a sales representative (considering the contact random conditional on the individual). F denotes the set of family members that the plan covers and for which the payer i makes the decision.

I assume that the outside good plan costs exactly the mandatory insurance and includes the outside option of going to a public hospital within its network of providers [3]. Given the previously mentioned assumption that the choice probability of the public hospital for a consumer in a public insurance plan is equal to one we can write the utility of the public plan as:

$$U_{\mathrm{i}0} = -\alpha_{2\mathrm{i}} \sum_{\mathrm{d}\in D} f_{\mathrm{id}} p_{0\mathrm{d}} + \varepsilon_{\mathrm{i}0}$$

Hence we can simplify the first part of $W_{\mathrm{i}k}$ within each consumer choice, and write the probability with which consumer i chooses plan k as:

$$s_{\mathrm{i}k} = \frac{\exp(\delta_{\mathrm{i}k})}{1 + \sum_{r\in K_{\mathrm{i}}} \exp(\delta_{\mathrm{i}r})}$$

With

$$\delta_{\mathrm{i}k} = \alpha_{1\mathrm{i}}(\phi_k - 0.07 * y_{\mathrm{i}}) - \alpha_3 \sum_{\mathrm{d}\in D} f_{\mathrm{id}} \ln(\sigma_{\mathrm{i}}(0|\mathrm{d}, k)) + \alpha_4 z_k + \xi_k$$

And where K_I is the set of plans available to the consumer who belong to market I, given by his gender, age group and income.

For the following analysis, it will often be useful to write $s_{\mathrm{i}k|r}$ denoting the choice probability of plan k by consumer i if plan r were removed from the market but premiums kept the same, with the obvious $s_{\mathrm{i}k|k} = 0$. Also, in order to simplify the notation I will often rewrite $\delta_{\mathrm{i}k} = \alpha_{1\mathrm{i}}\phi_k + x_{\mathrm{i}k}$.

[3]Note: It is natural to assume that most consumers would be in FONASA group D, hence the normalizing utility is that of a 20% coinsurance rate at certain providers but at a price not available in the data.

3.3 Insurers

I turn now to the modeling of insurer's profit maximization and competition. The insurer's Nash-Bertrand equilibrium is at the core of this research and the extent to which it is covered is unique within the current literature regarding competition in the health insurance market. Thanks to the highly segregated market structure describe above, I can equip the model with the simplifying assumption of full information and mean consumer consideration in each market. Hence from this point onward I will drop the consumer index i, linking every consumer related variable to the mean consumer of the given market being described.

Each insurer m optimizes the premiums of a set of plans K_{mI} in each market $I \in \mathcal{I}$. Denoting $|I|$ as the number of individuals in market I, Φ_m as the vector of premium controlled by insurer m and ec_k as the expected cost of plan k, the insurer's profit maximization problem can be written as

$$\max_{\Phi_m} \pi^m(\Phi, \mathcal{P}) = \max_{\Phi_m} \sum_{I \in \mathcal{I}} |I| \sum_{k \in K_{mI}} s_k(\phi_k - \mathrm{ec}_k) \tag{3.5}$$

Where $\mathcal{P}$ and Φ denote the total set of service prices and premiums, respectively, over all markets, insurers and hospitals. The expected cost of each plan is given by

$$\mathrm{ec}_k = \sum_{\mathrm{d} \in D} f_{\mathrm{d}} w_{\mathrm{d}} \sum_{j \in \mathcal{N}_k} (1 - c^k_{\mathrm{d}j}) p_{mj} \sigma(j|\mathrm{d}, k) \tag{3.6}$$

Where the expectation is over the risk of medical conditions for the individual (f_{d}) and the consumer's choice of hospital ($\sigma(j|\mathrm{d}, k)$). The remaining $w_{\mathrm{d}}(1 - c^k_{\mathrm{d}j})p_{mj}$ is the fraction of the service price the insurer must pay in each visit.

As insurers compete in prices, equilibrium premiums must satisfy the Nash-Bertrand first order condition:

$$\Phi^*_m \in \arg\max \pi^m(\Phi_m, \Phi^*_{-m} | \mathcal{N}, \mathcal{P}) \quad \forall m \in M$$

Where M is total set of insurers, Φ_m as the vector of plan premiums of insurer m and Φ^*_{-m} is the vector of its competitors premiums. Given the assumptions made regarding the market segregation an analytic solution to the Bertrand-Nash equilibrium can be obtained by solving the non-linear system defined by the following equation.

$$\phi^*_k = \pi_{m|k} + \mathrm{ec}_k - \alpha_1^{-1} - \alpha_1^{-1} W(\bar{\lambda}_k) \tag{3.7}$$

Where $\pi_{m|k} = \sum_{r \in K_{mI(k)}} s_{r|k}(\phi_r - \mathrm{ec}_r)$, $W(\cdot)$ is the Lambert W function and $\bar{\lambda}_k$ is:

$$\bar{\lambda}_k = (\sum_{j \in K_I, j \neq k} exp(\delta_j))^{-1} \exp(\alpha_1 \pi_{m|k} + \alpha_1 \mathrm{ec}_k - 1 + x_k)$$

A proof of this solution is provided in the appendix A. The possibility of writing a direct solution to the problem not only reduces the computational cost[4], but also provides a clear depiction of the value premiums take when disagreements occur between hospitals and insurers. This solution also allows a clearer analysis of the way Bertrand prices are set under logit

[4] Although not proven mathematically, the estimation routine shows this problem can be solve very quickly using standard fixed point iterations. On the other hand, standard non-linear Bertrand pricing is usually solved using general non-linear root finders which are far more computationally expensive.

demand. It states that premiums must not only cover the costs of the plan (ec_k) but also the opportunity cost of the insurer of including it in the market ($\pi_{m|k}$). The third term provides the base level of premiums, as α_1^{-1} is the monetary value of one util to consumers. The last term capture the additional increase (given that $\alpha_1 < 0$ and $W(x) > 0$ for positive x) due to the consumers substitution behavior and the plans non-monetary utility. This term can be seen as an adjustment factor accounting for the logit shape of the demand, as the numeric analysis shows it is between one and two orders of magnitude smaller than the effect of the first two terms.

Furthermore, this formulation provides a clear expression for the way that premium react to changes in negotiated prices, a characteristic of the model that has been so far limited in previous work. It can be shown that for a given change in the negotiated prices between insurer m and hospital j, the vector of market premiums Φ_I changes as

$$\begin{aligned}\frac{\partial \Phi_I}{\partial p_{mj}} &= \big(I - (\bar{S}_I \,\mathrm{di}ag(\alpha_1(\Phi_I - EC_I) + \mathbf{1}) \circ \Theta_I) \\ &\quad + \alpha_1 \mathrm{di}ag(\bar{\pi}_I)\bar{S}_I + \mathrm{di}ag(W(\bar{\lambda}))(I - \bar{S}_I)\big)^{-1} \bar{\mathbf{R}}_I(m,j)\end{aligned} \tag{3.8}$$

Where $\mathrm{di}ag(\cdot)$ is the diagonalization operator, $W = [W(\bar{\lambda}_1), \ldots, W(\bar{\lambda}_n)]^t$, $\bar{\pi}_I = [\pi_{m(1)|1}, \ldots, \pi_{m(n)|n}]^t$, and $\bar{S}_I$ is the matrix of counter-factual shares resulting from removing the plan indexed by the row, i.e, $\bar{S}_I[\mathrm{i}, j] = s_{j|\mathrm{i}}$. Θ_I is an adjacency matrix that has each element (i, j) equal to 1 if plan i and j belong to the same insurer and 0 otherwise[5]. Finally $\bar{\mathbf{R}}_I(m,j)$ has the form:

$$\begin{aligned}\bar{\mathbf{R}}_I(m,j)[\mathrm{i}] = &-(\sum_{r \in K_m^I} \frac{\partial x_r}{\partial p_{mj}} s_{r|\mathrm{i}})(\sum_{r \in K_m(\mathrm{i})^I, r \neq \mathrm{i}} s_{r|\mathrm{i}}(\phi_r - \mathrm{ec}_r) - \alpha_1^{-1} W(\bar{\lambda}[\mathrm{i}])) \\ &+ \mathbb{1}_{\{m(\mathrm{i})=m\}}[\sum_{r \in K_{m(\mathrm{i})}, r \neq \mathrm{i}} s_{r|\mathrm{i}}((\phi_r - \mathrm{ec}_r)\frac{\partial x_r}{\partial p_{mj}} - \frac{\partial \mathrm{ec}_r}{\partial p_{mj}}) \\ &+ \frac{\partial \mathrm{ec}_\mathrm{i}}{\partial p_{mj}} - \alpha_1^{-1} W(\bar{\lambda}[\mathrm{i}])\frac{\partial x_\mathrm{i}}{\partial p_{mj}}]\end{aligned}$$

with [i] indicating the row index and $\mathbb{1}$ being the indicator function. The proof of this derivative is given the appendix.

3.4 Hospitals

The theoretical framework for hospitals is directly related to that of insurers. Each hospital system s owns a set of hospitals H_s for which it must negotiate base service prices with each

[5] I assume that the Hadamard product ($\circ$) has lower precedence than the normal matrix product. i.e, it applies after the normal product.

insurer. The negotiations are done for each hospital separately, but disagreements involves removing all of the system hospitals form the insurer, allowing it to leverage the full weight of the system in the bargaining. For each hospital $j \in H_s$ and insurer $m \in M$, the negotiated price is denoted as p_{mj} and the marginal cost of serving a consumer in that hospital for the given insurer is mc_{mj}. In each of these negotiations, the system is seeking to maximize his total expected profit, given by:

$$\pi^s(\Phi, \mathcal{P}) = \sum_{m \in M} \sum_{j \in H_s} q_{mj}(p_{mj} - mc_{mj}) \tag{3.9}$$

Where q_{mj} is the expected demand to hospital j originating from insurer m, weighted by the resource intensity of each visit:

$$q_{mj} = \sum_{I \in \mathcal{I}} |I| \sum_{k \in K_{mI}} s_k \sum_{\mathrm{d} \in D} f_{\mathrm{d}} w_{\mathrm{d}} \sigma(j|\mathrm{d}, k)$$

Although the equations above seem to indicate the all insurers are connected to the system, the model allows for incomplete networks by simply setting $\sigma(j|\mathrm{d}, k) = 0$ for every condition d and plan k that is not connected to hospital j. Furthermore, the estimation routines allow for partial connections, such as j being connected to k only for condition d and not for another condition d$'$.

As can be seen above, the base marginal cost mc_{mj} (being the effective marginal cost $w_{\mathrm{d}} mc_{mj}$) can differ for each insurer in each hospital. This is to allow differences arising from payment frequencies, better business-to-business relationship and more importantly reduced costs from vertical integration. However, I restrict this formulation by imposing that the effects from each insurer and vertical integration must decompose linearly. That is,

$$mc_{mj} = v'_{mj}\rho + \zeta \tag{3.10}$$

Where v is a vector of indicators for each insurer, provider and year being analyzed, with an additional indicator for vertically integrated insurers and systems, and ζ is an error term.

3.5 Vertically integrated

A precise description of the behavior of vertically integrated insurers and health care provider is required in order to capture the market's characteristic concentration. I assume vertically integrated insurers and hospital systems set prices and premiums at the same time as non-integrated firms, while jointly maximizing both profit equations. I show that the optimal premiums for an integrated insurer can be solved using the same approach as before, with the addition of a component that captures the way the insurer considers the plan premium's impact on his own hospitals profits.

Consider a firm a that own an insurance company m and as hospital system s. In the premium setting round the company must solve the profit maximization problem,

$$\max_{\Phi_m} \pi^m(\Phi_m, \Phi_{-m}, \mathcal{P}) + \pi^s(\Phi_m, \Phi_{-m}, \mathcal{P}) \tag{3.11}$$

In order to write the optimal premium I first define $\tilde{\pi}^s_{I|k}$ as the hospital system analog of $\pi_{m|k}$, where k is the relevant insurance plan, which belongs to insurer m, and I is the corresponding market in which plan k operates:

$$\tilde{\pi}^s_{I|k} = \sum_{l\in M_s}\sum_{j\in H_s}\left(\sum_{\mathrm{i}\in K^I_l} s_{\mathrm{i}|k}\sum_{\mathrm{d}\in D} f_{\mathrm{d}} w_{\mathrm{d}}\sigma(j|\mathrm{d},\mathrm{i})\right)(p_{lj}-mc_{lj}) \\ - \sum_{j\in H_s}\sum_{\mathrm{d}\in D} f_{\mathrm{d}} w_{\mathrm{d}}\sigma(j|\mathrm{d},k)(p_{mj}-mc_{mj}) \tag{3.12}$$

Additionally, I define $\tilde{\lambda}_k = \bar{\lambda}_k \exp(\alpha_1 \tilde{\pi}^s_{I|k})$.
With this the optimal premium for plan k belonging to firm a can be written as

$$\phi^*_k = \pi_{m|k} + \tilde{\pi}^s_{I|r} + \mathrm{ec}_k - \alpha_1^{-1}(1 + W(\tilde{\lambda}_k)) \tag{3.13}$$

The proof applies the same steps used to solve equation (3.7) as presented in the appendix. Additionally, this new formulation of premiums implies that the general formula for the derivative of premium with respect to negotiated cost can be found by solving the inversion:

$$\frac{\partial \Phi_I}{\partial p_{mj}} = \big(I - (\bar{S}_I\,\mathrm{diag}(\alpha_1(\Phi_I - EC_I) + \mathbf{1}) \circ \Theta_I) + \alpha_1 \mathrm{diag}(\bar{\pi}_I)\bar{S}_I \\ + \mathrm{diag}(W(\lambda))(I - \bar{S}_I) - \alpha_1 \mathrm{diag}(\mathbb{1}_\Omega)(\bar{Q} - \mathrm{diag}(\bar{Q}\mathbf{1})\bar{S}_I)\big)^{-1}\mathbf{R}_I(m,j) \tag{3.14}$$

Letting the vectors and matrices being indexed by the plans ids, Ω being the set of vertically integrated plans, Θ_I the manufacturer-plan adjacency matrix, and $s(k)$ being the hospital system integrated to plan k, the new components of the previous equation are $\lambda[\mathrm{i}] = \bar{\lambda}_{\mathrm{i}} \exp(\mathbb{1}_{\{\mathrm{i}\in\Omega\}}\alpha_1\tilde{\pi}^{s(\mathrm{i})}_{I|\mathrm{i}})$ and $\bar{Q}_I[\mathrm{i},j] = \mathbb{1}_{\{m(j)\in M_{s(\mathrm{i})}\}}\sum_{z\in H_{s(\mathrm{i})}} s_{j|\mathrm{i}}(p_{lz} - mc_{lz})(\sum_{\mathrm{d}\in D} f_{\mathrm{d}} w_{\mathrm{d}}\sigma(z|\mathrm{d},j))$.Finally $\mathbf{R}_I(m,j)$ has the extended form of:

$$\begin{aligned}
\mathbf{R}_I(m,j)[\mathrm{i}] = &-(\sum_{r\in K^I_m}\frac{\partial x_r}{\partial p_{mj}} s_{r|\mathrm{i}})(\sum_{r\in K_m(\mathrm{i})^I, r\neq \mathrm{i}} s_{r|\mathrm{i}}(\phi_r - \mathrm{ec}_r) - \alpha_1^{-1}W(\lambda[\mathrm{i}])) \\
&+ \mathbb{1}_{\{m(\mathrm{i})=m\}}[\sum_{r\in K_{m(\mathrm{i})}, r\neq \mathrm{i}} s_{r|\mathrm{i}}((\phi_r - \mathrm{ec}_r)\frac{\partial x_r}{\partial p_{mj}} - \frac{\partial \mathrm{ec}_r}{\partial p_{mj}}) + \frac{\partial \mathrm{ec}_{\mathrm{i}}}{\partial p_{mj}} - \alpha_1^{-1}W(\lambda[\mathrm{i}])\frac{\partial x_{\mathrm{i}}}{\partial p_{mj}}] \\
&+ \mathbb{1}_{\{\mathrm{i}\in\Omega\}}\Bigg[\sum_{l\in M_{s(\mathrm{i})}}\sum_{z\in H_{s(\mathrm{i})}}\sum_{r\in K^I_l}\{(\sum_{a\in K^I_m, a\neq \mathrm{i}}\frac{\partial x_a}{\partial p_{mj}} s_{a|\mathrm{i}} \\
&+ \mathbb{1}_{\{l=m\}}\frac{\partial x_r}{p_{mj}})s_{r|\mathrm{i}}(p_{lz} - mc_{lz})\sum_{\mathrm{d}\in D} f_{\mathrm{d}} w_{\mathrm{d}}\sigma(z|\mathrm{d},r) \\
&+ \mathbb{1}_{\{l=m\}}(s_{r|\mathrm{i}}(p_{lz} - mc_{lz})\sum_{\mathrm{d}\in D} f_{\mathrm{d}} w_{\mathrm{d}}\frac{\partial \sigma(z|\mathrm{d},r)}{p_{mj}} + \mathbb{1}_{\{z=j\}} s_{r|\mathrm{i}}\sum_{\mathrm{d}\in D} f_{\mathrm{d}} w_{\mathrm{d}}\sigma(z|\mathrm{d},r))\} \\
&- \mathbb{1}_{\{m(\mathrm{i})=m\}}(\sum_{z\in H_{s(\mathrm{i})}}\sum_{\mathrm{d}\in D} f_{\mathrm{d}} w_{\mathrm{d}}\frac{\partial \sigma(z|\mathrm{d},\mathrm{i})}{\partial p_{mj}}(p_{m(\mathrm{i})z} - mc_{m(\mathrm{i})z}) + \sum_{\mathrm{d}\in D} f_{\mathrm{d}} w_{\mathrm{d}}\sigma(j|\mathrm{d},\mathrm{i}))\Bigg]
\end{aligned}$$

3.6 Bargaining and service price

Hospital systems and insurers that are not vertically integrated follow a sequential bargaining as in Collard-Wexler et al. (2015). They show that the solution to the bargaining process is given by the standard Nash Bargaining equation,

$$\max_{p_{mj}} \ (\pi_A^s - \pi_{NA}^s)^{b_{ms}}(\pi_A^m - \pi_{NA}^m)^{(1-b_{ms})} \tag{3.15}$$

Where the A and NA subindex stand for agreement and disagreement, which as mentioned before, involves removing the full system of hospitals from all of the insurer plans.
Although this equation is sufficient for describing the bargaining process between non-integrated insurers and systems, the vertical integration in the market adds four additional bargaining scenarios, being the 4 combinations of integrated to not-integrated and integrated to integrated bargaining. In order to extend the bargaining to this cases, consider the bargaining over price p_{mj} between two firms, a and b, with a owning hospital j and b representing insurer m. With slight abuse of notation, I will denote Ω as the set of integrated firms. Hence the general formulation of the bargaining problem is

$$\max_{p_{mj}} \ (\pi_A^a - \pi_{NA}^a)^{b_{ab}}(\pi_A^b - \pi_{NA}^b)^{(1-b_{ab})} \tag{3.16}$$

And by using the logarithmic transformation on the problem and imposing the first order condition, this problem's solution is given by

$$\frac{\partial \pi_A^a}{\partial p_{mj}} = -(\frac{1-b_{ab}}{b_{ab}})(\frac{\pi_A^a - \pi_{NA}^a}{\pi_A^b - \pi_{NA}^b})\frac{\partial \pi_A^b}{\partial p_{mj}} \tag{3.17}$$

separating possible integrations and defining $\rho_{ab} = \frac{1-b_{ab}}{b_{ab}}$, the general formulation is given by:

$$\frac{\partial \pi_A^{s(a)}}{\partial p_{mj}} = -\rho_{ab}(\mathbb{1}_{\{a\in\Omega\}}\frac{\pi_A^{m(a)} - \pi_{NA}^{m(a)}}{\pi_A^b - \pi_{NA}^b} + \frac{\pi_A^{s(a)} - \pi_{NA}^{s(a)}}{\pi_A^b - \pi_{NA}^b})\frac{\partial \pi_A^b}{\partial p_{mj}} - \mathbb{1}_{a\in\Omega}\frac{\partial \pi_A^{m(a)}}{\partial p_{mj}} \tag{3.18}$$

Note that this is general enough to include even the service price set by an integrated firm to its own hospital-insurer pairs. Simply setting $\rho_{aa} = 0$ for every firm a, the equation above is the first order condition to the problem of maximizing the firms equilibrium profit given by equation (3.11).
By expanding the terms relating to the system $s(a)$ the previous equation becomes

$$\sum_{r\in M_{s(a)}}\sum_{l\in H_{s(a)}}\frac{\partial q_{rl}^A}{\partial p_{mj}}(p_{rl} - mc_{rl}) + q_{mj} = -\mathbb{1}_{a\in\Omega}\rho_{ab}(\frac{\pi_A^{m(a)} - \pi_{NA}^{m(a)}}{\pi_A^b - \pi_{NA}^b})\frac{\partial \pi_A^b}{\partial p_{mj}} - \mathbb{1}_{a\in\Omega}\frac{\partial \pi_A^{m(a)}}{\partial p_{mj}}$$
$$- \rho_{ab}(\frac{\frac{\partial \pi_A^b}{\partial p_{mj}}}{\pi_A^b - \pi_{NA}^b})\sum_{r\in M_{s(a)}}\sum_{l\in H_{s(a)}}(q_{rl}^A - q_{rl}^{NA})(p_{rl} - mc_{rl})$$

Which can be stacked over for each system $s(a)$ giving place to the matrix equation

$$(\Psi_a + \Lambda^a)^{-1}(Q_a + \Gamma_a) + P_a = MC_a \tag{3.19}$$

Where Q_a is a vector of all q_{mj} with $m \in M_s(a)$ and $j \in H_s(a)$ and P_a, MC_a follow the same structure for service prices and marginal costs.
$\Psi_a[\mathrm{i}, j] = \frac{\partial Q_a[j]}{\partial P_a[\mathrm{i}]}$, $\Lambda^a[\mathrm{i}, j] = \rho_{ab[\mathrm{i}]} \frac{(Q_a[j] - Q_a^{NA(m[\mathrm{i}])}[j])}{\pi_A^{b[\mathrm{i}]} - \pi_{NA}^{b[\mathrm{i}]}} \frac{\partial \pi_A^{b[\mathrm{i}]}}{\partial P_a[\mathrm{i}]}$, where the square braces in $m[\mathrm{i}]$ and $j[\mathrm{i}]$ indicate that this are the insurer and hospital corresponding to observation i in vector Q, P and MC. Also $Q^{NA(m[\mathrm{i}])}$ corresponds to the vector of equilibrium quantities to system s given a disagreement with insurer $m[\mathrm{i}]$, defining the quantities provided by that same insurer as zero. Finally $\Gamma_a[\mathrm{i}] = \mathbb{1}_{\{a \in \Omega\}}(\rho_{ab}[\mathrm{i}](\frac{\pi_A^{m(a)} - \pi_{NA[\mathrm{i}]}^{m(a)}}{\pi_A^{b[\mathrm{i}]} - \pi_{NA}^{b[\mathrm{i}]}}) \frac{\partial \pi_A^{b[\mathrm{i}]}}{\partial P_a[\mathrm{i}]} + \frac{\partial \pi_A^{m(a)}}{\partial p_a[\mathrm{i}]})$.

It is in this last equation that the true complexity of the model can be observed. Note that given the timing of the game, for every disagreement, premiums must be readjusted and that for every increase in the negotiated price, all market premiums react. Therefore, in every step of the negotiation or vertical service price optimization, each firm must not only consider the way the prices affect the choice probabilities of their consumers, but also the way the rest of the market will adapt to this changes and the subsequent effect on the distribution of consumers in each market among plans and among hospitals. The complexity of this market dynamics can lead to completely distinct results for an increase in service price, such as a large loss of market share by the hospital or a full recapture (and possibly even an increase in profits), through consumers changing their plans in order to access the hospital.

Despite that, equation (3.19) is fairly intuitive to analyze. First of all, notice that $\Psi_a^{-1} Q_a + P_a = MC_a$ is the standard first order condition of a Bertrand pricing game with heterogeneous goods, which in this case would be plan-hospital pairs offered to consumers. This can be understood as the price the hospital system would set whether it had full control over the price (i.e, $b_{ab} = 1$). As Ψ captures the effect of increasing prices on profits when the hospital system is in control, Λ^a is the adjustment to this due to the bargaining process. Clearly $\frac{\partial \pi_A^{b[\mathrm{i}]}}{\partial P_a[\mathrm{i}]}$ holds the insurer's profit effect of increasing prices and note that we can rewrite the preceding components as the ratio between $\frac{1 - b_{ab}}{\pi_A^{b[\mathrm{i}]} - \pi_{NA}^{b[\mathrm{i}]}}$ and $\frac{b_{ab}}{Q_a[j] - Q^{NA(m[\mathrm{i}])}}$, which is the relative bargaining power of the insurer with respect to the hospital system, for each unit of hospital profit being negotiated (as all the components within the inverse represent marginal changes with respect to hospital price). Finally Γ_a includes the effect that the vertical integration has on the way the previous components are translated to markups. The last component of Γ_a, specifies the added profit to the firm due to consumers the would be recaptured through the integrated insurer if the prices were to increase. The preceding term is the negotiating insurer's attempt to represent his interests against this competing vertically integrated insurer, with the bargaining power ratio again indicating the way this two conflicting components are resolved under bargaining.

Chapter 4

Data

I use data on the universe of insured individuals in the metropolitan region of Chile during the years 2007-2014. The data consists of several separate datasets constructed by the regulatory agency (*Superintendencia de Salud*) using information sent by all private insurers. I combine five different datasets in order to construct the panels required for each of the three estimation routines executed.

The first and largest dataset contains information on all the visits each privately insured individual made to a health care facility, private or public, during the period specified. This data contains the patients id, the payer id, an identifier that can be used to obtain the ICD-10 code for the treated medical condition, the patients age and gender, the amount charged by the provider, the amount payed by the insurer, and other details regarding the type of claim and costs, including an indicator if the treatment was ambulatory or inpatient. The second dataset contains inpatient discharge data for all hospitalized privately insured costumers during the seven year period. I use this data mainly to link the ICD-10 codes to the identifier from the previous table. The third dataset contains detailed information on each payer during each year, describing the consumer age, gender, working status, currently held health insurance plan, the date of subscription and renewal, the amount payed for the given plan, and the mandatory minimum the insured must pay, which can be used to recover his taxable wage up to the threshold mentioned previously. The fourth dataset contains details regarding all valid plans during the given year, including plans that are no longer available but still have active clients. This data contains details regarding the plans coverage and other benefits, the number of active clients and the reference price of the plan. Finally, the last dataset I use contains information regarding dependents: insured individuals with a third party payer (usually spouse or parents). I use this data to obtain the same demographics included for the payer in the third dataset in order to construct risk and preferences for all insured individuals alike.

Given that I posses diagnosis code only for inpatient care I restrict the data to this subset and group medical conditions based on the ICD-10 Diagnosis Chapter. I reduce the number of diagnosis groups to 11, with the first 10 representing the most frequent in the data and the last grouping the remaining conditions. I additionally remove from the data small providers, like independent doctors, ophthalmology clinics and other providers without names or unclear identity. This data contains more then 1.2 million claims, from 682363 unique payers, in 86 different health care providers, covered by 7 insurers, summing up to an estimated average of $640 million USD per year in medical claims. Additionally, the plan demand data contains the choices of 395497 unique payers, choosing among 1431 plans, paying more $24 million USD pear year on average in premiums.

Recent literature has focused on the effects of foreclosure in vertical markets (Chipty (2001); Crawford et al. (2015)). However in this data there is no clear evidence of foreclosure in prices and the institutional framework does not allow providers to deny service or restrict access through means other than price. Table 4.1 shows the distribution of base service prices separated by the type of bargaining that takes place between each insurer and provider, for the 11 largest private hospitals in the data. In the presence of foreclosure, the prices of *integrated insurer only* should be higher than those of *vertically integrated*, in order to redirect consumer to their integrated insurer. However, the data seems to show the contrary, with prices being higher for integrated insurers. A simple t-test does not reject the hypothesis of different means at a 5% confidence, but does reject it at 1%. Nonetheless, the difference is very small and I don't consider it an important characteristic of the data when modeling the market.

The model is estimated in four separate steps, for each of which I create distinct datasets in an attempt to extract as much information available from the five master files provided by the regulatory agency. In the following section I describe the estimation procedure in each of these steps, including a description of the dataset used and the creation procedure.

Table 4.1: **Service price by bargaining type**

	Negotiated price				
Bargaining type	**min**	**mean**	**max**	**sd**	**N**
not integrated	11.4588	109.2144	428.8748	50.43025	488
integrated insurer only	6.946354	105.658	252.4554	40.59592	1330
integrated provider only	2.288709	75.03959	351.5181	44.45129	1000
vertically integrated	11.50994	75.32896	200.575	39.49844	659
separately integrated	.1405515	69.67581	331.8322	35.79255	2558

Not integrated marks the prices set between insurers and providers which are not vertically integrated with anyone, *integrated insurer only* marks those in which only the insurer is vertically integrated with some provider different from the current one, *integrated provider only* is the analogous for providers, *vertically integrated* marks the prices set through vertical integration and *separately integrated* prices are those set by an integrated provider with an integrated insurer, where both integration are different.

Chapter 5

Estimation

5.1 Overview

As the theoretical framework showed, the model is composed by several layers of interaction which, given the available data, can be used to identify the different parameters of the model. The three main estimations steps are those linked to the hospital demand, the insurance demand and service price bargaining or vertically integrated setting step. A fourth additional estimation is required to compute the risk (f_{id}) and condition weights (w_{d}) from the data.

Each of these steps use a particular combination of data to leverage as much information as is available. The two demand estimations use particularly large datasets and can be considered as *Big Data* problems; the matrix size is too large to be handled by standard algorithms as many of the intermediary products are too large to fit in contiguous memory blocks and too lengthy to write to disk. To solve this problem I use a new method developed by Kenneth Train in his paper "Mixed Logit with a flexible mixing distribution" (working paper) and adapt the algorithm to high-order parallelism using distributed memory. The code is written in C (std11) using MPI and has a structure of master-slave with an intermediary group leader that handles the communication with the master and propagates the information to the worker nodes using a binary tree to reduce communication overheads. Given the novelty of the new Train method (to which I will refer to as Flexible Mixed Logit) I provide a brief analysis of it along with Monte Carlo experiments in the appendix, contrasting its precision against the standard Mixed Logit estimation using maximum likelihood. Additionally, the C code is written generically and is provided along with this research to be used by others. It's is very simple to use despite it's significant power and is already being employed in a separate research without any modification.

One of the peculiar characteristics of the Chilean health insurance market is the wide offer of insurance plans. There are often thousands of plans being offered each month, with a vast majority only having a few, if not a single, active clients. The insurance companies create this large assortment of plans in order to circumvent laws that forbid risk reclassification but permit readjusting plan prices up to a certain level. It also allows them to further discriminate consumers, beyond the legal characteristics of age and gender, offering a particular premium to each individual. However, as far as the data reveals, most plans differ in very small aspects. For example, given that coverage and price are fixed per plan, two plans that differ only by a 5% reduction in coverage at a particular provider for a given condition will be identified by separate ids. Therefore it makes sense to group plans by similar attributes to reduce the assortment and make the estimates more significant. Using the plans and claims data I group plans by insurer, plan group name [1], the identity of the plan's preferential provider (when there is one), the existence of a coverage cap, whether the plan is individual or covers a group of additional family members, and the average coverage rate in public, private and preferential providers. These procedure groups the 100738 observed plans into 3473 groups. Given that I use this grouping in all of the following estimations, I will simply refer to each of these plan groups as a plan, obviating the underlying individual plans.

In all the estimations, all public hospitals are grouped into one large provider offering the mean service price of all public providers for the condition, plan and year. Hence creating the public service outside good option, as described in the theoretical model. Additionally I separate the 11 most well known private hospitals from the remaining 74, referring to the first group as central private providers, and to the later as non-central private providers. Central private providers account for about 87% of all private claims and collect approximately 93% of all private income. This division is useful when estimating the model, as the remaining private providers have little leverage in the bargaining phase and often have an insufficient number of observations for individual effects to be estimated with precision.

Finally, in order to account for inflation between the observed year, all prices are measured in UF instead of Chilean pesos. The UF being a real valued unit that is used in many large transactions in Chile (such as real-estate and wages), which has an inflation adjusted nominal value posted daily by the internal revenue service (*Servicio de impuestos internos*).

5.2 Risk

I use data on the total number of privately insured individuals during 2007 and the related claims data, to generate the individual risk of illness. I split the market according to the standard division used by insurers in Chile to define plan prices, which separates individuals by gender and age, grouping ages from zero to two, two to five and forwards by groups of five years, until sixty years old, which are grouped with all older beneficiaries. The measures of risk are standard frequency estimator of the number of observed claims for each condition within the population group. To avoid inducing errors on the measurements, I remove the 2007 data from the posterior estimations of plan demand and bargaining. However, given that hospital demand is made *ex-post*, conditional on plan and health status, these estimates are not affected by the risk calculated here. Table 5.1 shows the descriptive statistics of the risk estimates by condition and gender, measured by year.

[1] Most plans are have names such as "dali-hombre-12315" and "dali-hombre-675", clearly showing that "dali-hombre" is a name that groups several plans under some logic.

Table 5.1: Risk distribution (f_{id}) by gender and condition.

Gender	Condition	min	mean	max
Male	Neoplasms	.0000282	.0010879	.0038573
	circulatory	.000014	.0008982	.0031077
	digestive	.0002213	.0027004	.0096432
	genitourinary	.0001513	.0013347	.0053038
	injury and exter	.0001818	.001125	.0050627
	musculoskeletal	.0000337	.0022061	.0113308
	nervous system	.0000424	.0010616	.0036162
	others	.0003249	.0021788	.0062681
	perinatal	.0005725	.0005725	.0005725
	pregnancy and related	.0000209	.0000227	.0000246
	respiratory	.0001846	.0012577	.0033751
Female	Neoplasms	.0000146	.0010456	.0027395
	circulatory	.0000146	.0002774	.0009224
	digestive	.0003659	.0009358	.0019358
	genitourinary	.0000523	.0008845	.0026194
	injury and exter	.0001024	.0002894	.0007053
	musculoskeletal	.0000293	.0006765	.0015036
	nervous system	.0000146	.0002383	.0004603
	others	.000322	.0014303	.0032399
	perinatal	.0000261	.0002776	.000972
	pregnancy and related	.0000293	.0088889	.0343412
	respiratory	.0001793	.000649	.0018359
	Total	.0000146	.0011926	.0343412

Risk is measured in expected number of events per year.

Next, I calculate the condition weights using the 2007 data. I assume that the mean of the observed price of treatment for condition d for a patient insured by insurer m at hospital j can be written as

$$\rho_{mj\mathrm{d}} = w_{\mathrm{d}} p_{mj}$$

Where p_{mj} is the negotiated price, w_{d} is the diagnosis weight and $\rho_{mj\mathrm{d}}$ is the observed price. I then regress the following equation

$$\ln(\rho_{mj\mathrm{d}}) = \ln(w_{\mathrm{d}}) + \ln(p_{mj}) + \xi_{\mathrm{i}j\mathrm{d}}$$

With $\xi_{\mathrm{i}j\mathrm{d}}$ being the regression error term. I identify the first term as the resulting coefficient of diagnosis-group fixed effects and the log of prices as the resulting coefficients of the interaction of plan and provider fixed effects. I further decompose $w_{\mathrm{d}} = \omega_{\mathrm{d}}\bar{w}$ in order to form a more familiar scaling effect. Table 5.2 shows the resulting coefficients as estimated.

Table 5.2: Condition weights

Condition	w_{d}	ω_{d}
circulatory	91.414685	1.305323
perinatal	28.820359	.41153
pregnancy and related	65.799184	.9395558
respiratory	75.033642	1.071416
digestive	100.31874	1.432465
genitourinary	71.247983	1.01736
injury and external	75.081742	1.072103
musculoskeletal	94.467431	1.348914
neoplasms	62.823293	.8970627
nervous system	43.122712	.6157553
others	70.743007	1.010149

5.3 Hospital Demand

Using the claims, discharge and payer datasets I create a panel of hospital demand, conditional on plan and diagnosis. I use the full data period (2007-2014), keeping only plans for which I observe at least one public provider, one private provider, 10 diagnosis groups, and 100 beneficiaries, over the entire eight year period. Defining markets as year-plan-diagnosis groups, I further restrict the data to those market that have at least one public and one private provider choice. I also remove from the data all claims for less than $10000 Pesos ($14.5 USD) to avoid including ambulatory care that might have been coded improperly. I use Train's Flexible Mixed Logit with logit-normal distributed coefficients and 10000 Sobol sequence draws, which I map from the unitary hypercube to the bounds specified in each version. Given that the model estimates the distribution of heterogeneous preferences over individual level parameters, the 241329 choice observations included in the filtered data must be expanded to include all possible choices for each consumer in each market in which he makes a decision [2], which creates more than 2 millions rows in the final dataset.

Table 5.3 shows the estimated coefficients for three different specifications. The coefficients shown in this table are the mean and standard deviation of each of the model's coefficients as implied by the distribution fitted through Flexible Mixed Logit. The actual fitted coefficients which define the shape of the distribution are presented in table 9.1 in the second appendix, as they offer little interpretation. The numbers shown in parenthesis are the estimated standard errors, which are calculated using 20 bootstrap draws with replacement and using a random sub-sample of choices of the size specified in the last row of the table. Besides the price coefficient, the other estimated variables are: *same comuna*, which indicates if the payer and the provider are located in the same geographical area (i.e, comuna); *preferential*, which indicates if the given provider is the plan's preferential provider; *private*, indicating if the provider is a private hospital or clinic. Although Flexible Mixed Logit does not allow for fixed effects in the usual way (all the coefficients are jointly distributed and therefore have a

[2] Though computing the value of each option in the market for each consumer during run-time is feasible, it is highly inefficient and results in the same memory usage as if it is created before estimating.

variance associated) I include effects on the system of hospitals using system dummies, and on the provider using hospital dummies for the central private providers [3]. I set the bounds for the positive price coefficient (α_2) to be $[0.0, 5.0]$ and for all other variables $[-5.0, 5.0]$, which given the results are sufficiently wide. I optimize the maximum likelihood problem using BFGS with analytic gradient using on average 40 computing cores at the National Laboratory for High Performance Computing (NLHPC), taking about a day to solve the first version and almost three days to solve the last, including the preliminary grid calculations and the post-estimation bootstrap [4].

Table 5.3: Hospital demand estimation

			specification	
coefficient	statistic	(I)	(II)	(III)
service price (α_2)	mean	0.04095	0.02807	0.03882
		(0.00008)	(0.00000)	(0.00001)
	stdev	0.40223	0.22982	0.03621
		(0.01803)	(0.00220)	(0.00002)
same comuna (β_1)	mean	-0.10423	-0.062643	0.83755
		(0.00187)	(0.00229)	(0.08792)
	stdev	2.94473	2.990052	2.17033
		(0.00657)	(0.00875)	(0.02754)
preferential (β_2)	mean	1.18086	1.26232	2.26647
		(0.00115)	(0.00315)	(0.06836)
	stdev	3.62511	3.46386	2.47679
		(0.00340)	(0.00100)	(0.10362)
private (β_3)	mean	1.10173		0.29518
		(0.00471)		(0.05415)
	stdev	2.05122		2.88313
		(0.00289)		(0.04173)
System Effect		No	Yes	No
Provider Effect		No	No	Yes
Bootstrap choices		100000	100000	1000

[3] The second version does not have a private indicator because networks are defined for both central and non-central private providers.

[4] The NLHPC has a maximum running time of 3 days, hence for the model to be estimated within the time limits, I had to reduce the bootstrap draws on the last specification.

5.4 Plan demand

Using the results of the third specification of table 5.3, I compute the expected utility (W_{ik} in the model) for all observed plans for all payers and their dependents, as observed in the respective master datasets. This naturally limits the assortment of plans offered in each market to those observed in use during that year. Risks are adjusted for each payer and dependent demographic, for each available plan in the semester in which they change plan or complete a year with their current plan. Plans available for the consumer are restricted to those that cost at least 7% of her income, as mandatory. Plan network values are computed on an annual basis, while prices are evaluated each semester. This is done in order to minimize the loss of plans due to insufficient observations in the claims data, and at the same time to estimate the demand under a scenario more likely to resemble the one the consumer faced. I further restrict the data only to plans that cost more than 1UF, that have at least 30 clients each year and that target payers over 20 years old [5]. Markets in which there are less than 3 plans being offered are also removed. In order to compare plans properly, plans are restricted to those for which the 5 most common medical conditions (neoplasms, digestive, respiratory, circulatory and pregnancy) are observed. Accordingly, each plan's expected utility is evaluated only on these 5 conditions. The resulting dataset has more than 6 million rows, with 258753 unique choices of 181465 different payers. The data spans the years 2008 to 2014, with 123 different plans being offered, where the average contract is worth 3.5UF (or 130 USD) per month, with the average insurer signing about 81265.86 UF (approximately 35.8 million USD) worth of health insurance per year.

I estimate the model using the same Flexible Mixed Logit code used for the hospital demand. I use 5000 Sobol sequence draws to form the grid, with the the premium coefficient being bounded to $[-3.0, 0.0]$, the household plan value coefficient to $[10.0, 35.0]$ and all other coefficients to $[-3.0, 3.0]$. The first version estimates the model using premiums, plan values and indicator of whether the plan has a preferential provider. The estimation used 40 computing nodes and took almost two days to solve. The second version adds effect on the insurer and on the preferential provider (if there is one and only of its one of the central private providers). The procedure takes about four days to estimate on the same amount of cores [6]. Table 5.4 shows the resulting price and plan value coefficients estimates, while table 9.2 in the second appendix shows the distributional parameters estimated.

Despite the detail of the data, the panel used for this estimation is truncated, omitting consumers who choose the public insurance (FONASA). Given that public insurance is cheaper and provides less coverage for consumers, it is likely that this truncation affects more consumers who face tighter budget constraints or are more expensive for insurers. Therefore, wealthier segments of the market are probably better represented by this estimation, while

[5]Plans under 1UF are very rare and therefore are treated as irrelevant and probably an input error. The same logic applies for payers under 20 years old. Plans used here are the plan groups mentioned in the overview, which are a grouping of smaller commercial plans.

[6]Intermediary results had to be stored for this last version, as the maximum runtime at the NLHPC is three days. This also slows the estimation time.

Table 5.4: Plan demand estimation

coefficient	statistic	specification (I)	(II)
price (α_1)	mean	-0.29999	-0.38358
		(0.00030)	(0.00100)
	stdev	0.36719	0.42733
		(0.00030)	(0.00052)
plan value (α_3)	mean	26.54798	29.15160
		(0.876337)	(0.15889)
	stdev	10.56137	5.26874
		(0.154076)	(0.17843)
Preferential provider	mean	1.15706	0.91774
		(0.00224)	(0.00427)
	stdev	2.71597	2.24620
		(0.00043)	(0.00688)
Insurer effects		No	Yes
Preferential provider effects		No	Yes
Bootstrap choices		5000	5000

more price sensitive consumers are not. In order to compensate for this bias, I recover the premium coefficients by rewriting the downstream equilibrium condition for non-integrated insurers in terms of observable parameters (a proof of this formulation is provided in the appendix).

$$\phi_k^* = \text{ec}_k - \alpha_1^{-1} \frac{1}{1 - \sum_{r \in K_{mI}} s_r} \quad \forall k \in K_{mI} \tag{5.1}$$

Premiums and market shares are provided in the plans data while expected costs can be recovered form the claims data. However, given that there are plans for which only few claims are observed, I use the unbiased hospital demand and risk estimates in order to construct the expected costs for each plan. Denoting by $F(\alpha_2, \beta)$ the estimated probability distribution of hospital demand parameters, I recover premium coefficients for each market (age-group, gender, and dependent status) as

$$\alpha_1^{-1} = -\sum_{\text{i}=1}^{R} ((\phi_k^* - \text{ec}_k(\alpha_2^{\text{i}}, \beta^{\text{i}}))(1 - \sum_{r \in K_{mI}} s_r) F(\alpha_2^{\text{i}}, \beta^{\text{i}}) \tag{5.2}$$

With R being the number of coefficient vector in the hospital demand grid.

Table 5.5 shows the resulting mean parameters for each market, which are well within the bounds and standard deviations estimated above. The estimated coefficients indicate that consumers grow less sensitive to premiums as they age, which is likely due to increased wealth and to higher premium reclassification risks. The observation that female payers older than 40 are more premiums sensitive than male payers, and that male payers have a decrease in premium coefficients reaching their fifties, coincides with additional benefits provided by the public insurance system (AUGE) for conditions such as breast and prostate cancer that are more frequent in the older population and that affect women earlier than men. The reduced premium sensitivity of payers with dependents is likely due to higher costs of switching. As expected, table 5.6 shows that this adjustment increases estimated premium elasticity for most consumers, while indicating that payers with dependents where biased upwards. Figure 5.1 illustrates the evolution of premium coefficients and elasticity for a particular group. The estimated plan demand introduces variation both in premium elasticity through the adjusted premium coefficient and in network valuations through differences in risks, allowing for the theoretical model's assumption of within-market homogeneity and across-market heterogeneity.

Table 5.5: **Adjusted premium coefficient** (α_1)

	No dependents		With dependents	
Age group	**Male**	**Female**	**Male**	**Female**
20	-.9597096	-.9555035	-.2866724	-.2849887
25	-.6375663	-.5990327	-.3720486	-.3866599
30	-.614356	-.6369665	-.3138012	-.3203015
35	-.519389	-.612922	-.2339378	-.3049266
40	-.3731824	-.5487115	-.2653409	-.3147073
45	-.38773	-.5288834	-.2700835	-.3141856
50	-.5365676	-.5319779	-.2582539	-.2797625
55	-.5234088	-.4958592	-.2527101	-.3038548
60	-.4162695	-.4875786	-.2560229	-.3087475
Total	-.5407104	-.5906936	-.2781074	-.3133298

5.5 Bargaining and Service Price

Using the datasets created for the plan and hospital demand, with the addition of the estimated parameters in the last version of both demands, I proceed to solve the service price equation (3.19). I write a custom solver in C (std11) which estimates the marginal costs and bargaining weights (when not vertically integrated) by nesting two GMM routines. The solver reads the hospital, plan, and risk data mentioned above, and an additional bargaining dataset that contains the identity of each bargaining pair (hospital, insurer) together with their owners identity, the negotiated service price as observed in the data and the vertical integration status. Every iteration, the solver provides the algorithm with a guess of marginal costs for all the vertically integrated hospitals and bargaining weights for the non integrated systems and insurers. The algorithm then proceed to recalculate the equilibrium premiums and its derivatives given these parameters (as marginal costs affect the premiums through vertical integrations) using equations (3.13) and (3.14). These premiums are then used to solve the service price equation (3.19) in order to recover the equilibrium marginal costs, conditional on the starting guess. Though I provide no formal proof of contractancy, the numeric experiment shows that both equation (3.13) and (3.19) can be solved via standard Picard

Table 5.6: **Adjusted and Biased own-premium elasticity**

		No dependents		With dependents	
Age group	**stats**	**Male**	**Female**	**Male**	**Female**
20	Adjusted	-.9817464	-1.521845	-.5866011	-.7832114
	Biased	-.3717375	-.6109367	-.6609695	-.8877195
25	Adjusted	-.8218073	-1.271374	-.983108	-1.436038
	Biased	-.4084401	-.6725222	-1.013583	-1.362667
30	Adjusted	-1.107186	-1.772737	-1.150803	-1.452965
	Biased	-.6636355	-1.024842	-1.406709	-1.670419
35	Adjusted	-.9909385	-1.794437	-1.112792	-1.475559
	Biased	-.7318329	-1.123003	-1.824616	-1.856177
40	Adjusted	-.7053151	-1.466892	-1.263854	-1.439052
	Biased	-.579976	-.9434089	-1.827052	-1.753992
45	Adjusted	-.7458282	-1.437231	-1.234714	-1.393084
	Biased	-.5902791	-.9589863	-1.753583	-1.700784
50	Adjusted	-1.127319	-1.471217	-1.329662	-1.341762
	Biased	-.7736624	-1.018386	-1.974933	-1.766099
55	Adjusted	-1.356448	-1.618302	-1.322302	-1.495332
	Biased	-.994077	-1.25187	-2.007086	-1.887685
60	Adjusted	-1.637489	-1.874442	-1.476316	-1.536663
	Biased	-1.508905	-1.474638	-2.211864	-1.90912
Total	Adjusted	-1.052958	-1.582948	-1.178068	-1.386742
	Biased	-.7443061	-1.019181	-1.658754	-1.664308

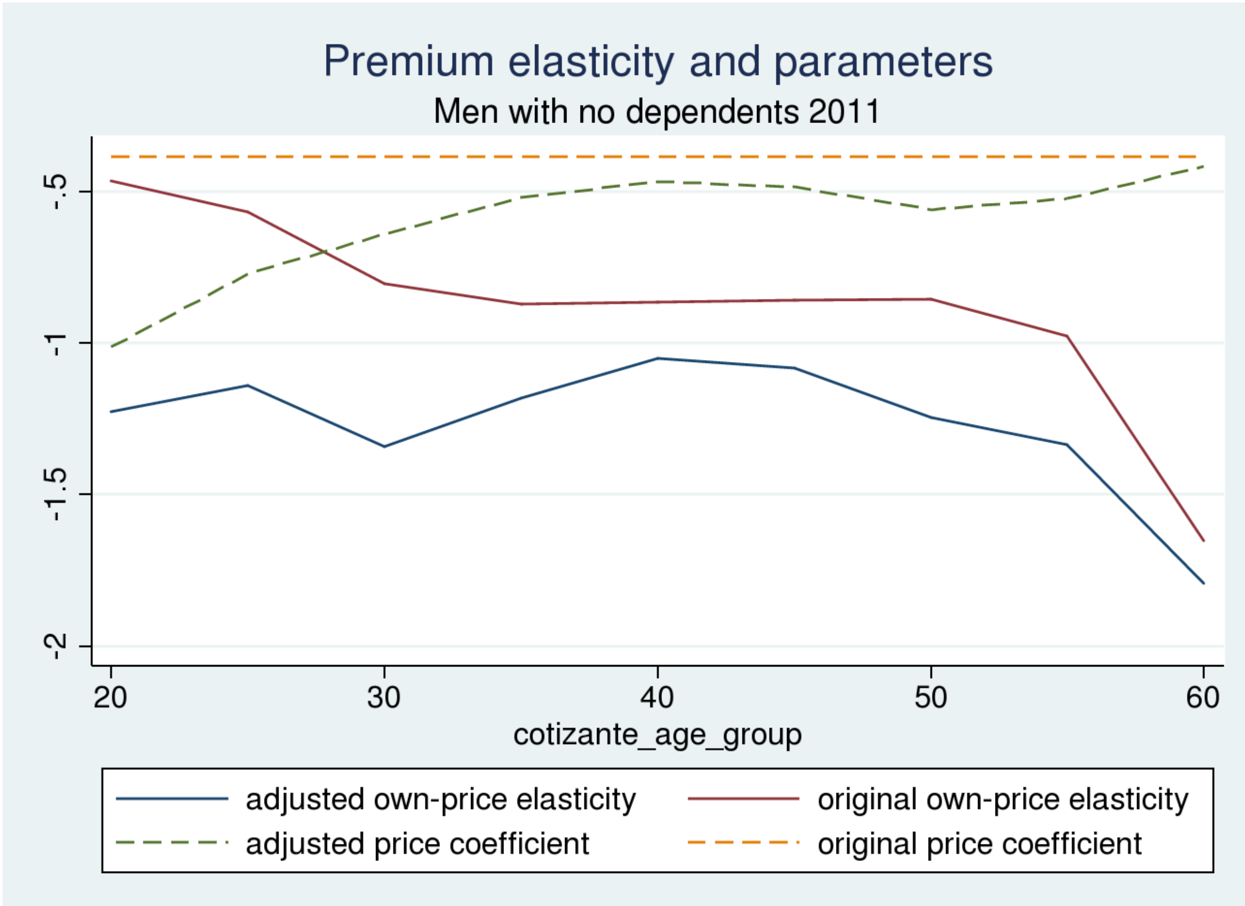

Figure 5.1: Premium elasticity evolution

fixed point iterations, in which the last resulting premium or marginal cost is used as starting point to compute the next until convergence is reached [7]. The solver then uses the resulting marginal costs and solves the marginal cost decomposition equation (3.10) via GMM estimation. Given that this is a linear problem it can be solved directly in one iteration. Then the residual of this estimation is used as the input for the non-linear outer GMM estimator which minimizes the moment conditions by adjusting the bargaining weights. Though the solver uses an underlying implementation of the Nelder and Mead Simplex algorithm [8] which searches over $\mathbb{R}^n$, I map the bargaining weights directly to $(0, 1)$ by applying the function $\frac{1}{1+\exp(-x)}$. It's worth noting that the non-linear GMM only has to solve for the bargaining weights given that the conditional equilibrium marginal costs can be calculated every round using equation (3.19). Hence, although the vertical integration makes every iteration of the solver significantly more computationally expensive, the number of parameters to estimate is reduced and with it so does the number of required iterations to convergence.

Given that the algorithm requires the use of all the previously created datasets, only the intersection of dates and observations between them can be used. Additionally, there are significant algorithmic advantages of using only markets in which all insurers and systems

[7] I evaluate the convergence of $f(mc) = mc$ using the euclidean distance between the vectors, i.e $\varepsilon = ||f(mc) - mc||_2$

[8] I use the GSL Nelder-Mead 2 algorithm which is $O(N)$ rather than the usual $O(N^2)$.

are present. Therefore the usable data for this last estimation is much smaller than before.

I execute the algorithm using data from the years 2008, 2011, 2013 and 2014 which provide the best match between the different datasets. I restrict the attention to the three biggest insurers (Banmedica, Colmena, Consalud) which together account for 84.51% of claims and 81.67% of all chosen health insurance plans. I remove providers that have a difference between their lowest and highest negotiated price of more than 400% in order to reduce noise resulting from input errors and payment methods that are not observed in the data and might affect the negotiated price estimate. This removes one of the central providers (Clinica UC) and a few smaller clinics. I also discard three very expensive clinics (Clinica Alemana, Clinica Las Condes and Clinica UC San Carlos) because I believe that demand for this clinics is strongly correlated with income, implying a covariance between income levels and provider fixed effects that are not estimated in this work. Bargaining is limited to the seven largest private providers which account for 82.97% of all inpatient private hospital care, and for approximately 87.41% of all the private sector's claims value [9]. In addition to these providers, the data includes the public system and 14 smaller non-integrated private providers, whose price is taken as exogenous. As mentioned previously, the public system plays an important role as a common outside option among insurers. The additional small private providers add variety to consumers outside option among plans and allows for more realistic substitution patterns between providers. Given that they represent less than 10% of the claims and about 5% of the private value of the market, whether their price is set through bargaining or taken as exogenous has little effect on the parameters of interest, namely the costs of the large private hospitals. The data is further restricted to include only the 30 more populous markets and the 3 largest plans for each insurer in each market, leaving a total of 581 plans. I omit market with dependents from this estimation in order to reduce the computational cost associated with keeping track of different family compositions and risks. As moment conditions I use the exogeneity of the indicators used for the marginal cost decomposition (3.10) and the average plan value of all other insurers in the same year. Table 5.7 shows the ownership structure of the data and sets the nomenclature for the five combinations of integration that define the price setting.

Table 5.8 shows the estimated marginal costs and bargaining weights, broken down by bargaining type. It also includes the estimated bargaining ratio, which are the quotients of agreement profits to disagreement profits for insurers and providers, and the observed prices of the negotiation. Table 5.9 shows the estimated marginal costs and observed prices by hospital. Table 5.10 contrasts the estimation's premiums with the observed.

[9] Before removing the outlier clinics, the seven largest providers account for 60.57% of all inpatient private hospital care and for 48.34% of all the private sector claim value.

Table 5.7: **Ownership and price types**.

System	Provider	Banmedica	Colmena	Consalud
1	**Davila**	Vertical	Hospital	Separate
1	**Santa Maria**	Vertical	Hospital	Separate
1	**Vespucio**	Vertical	Hospital	Separate
2	**Avansalud**	Separate	Hospital	Vertical
2	**Tabancura**	Separate	Hospital	Vertical
2	**Bicentenario**	Separate	Hospital	Vertical
3	**Indisa**	Insurer	None	Insurer

Vertical: Vertically integrated firm price setting.
Hospital: Integrated hospital bargaining with a non-integrated insurer.
Insurer: Integrated insurer bargaining with a non-integrated provider.
Separate: Both insurer and provider are integrated but separately.
None: Neither the insurer or the provider are integrated.

Table 5.8: **Estimated marginal costs and bargaining weights by bargaining type**

		marginal	service	hospital barg.	bargaining ratio	
Bargaining	**stats**	**cost**	**price**	**weight**	**insurer**	**provider**
None	min	59.72664	97.00546	.503179	1.000228	1.234192
	mean	73.10309	98.80271	.503179	1.000432	1.318673
	max	97.00656	100.5332	.503179	1.000553	1.364494
	N	4	4	4	4	4
Hospital	min	-9.683437	47.56427	.4979408	1.004067	1.001544
	mean	18.04907	80.41659	.5023519	1.006397	1.003636
	max	64.44559	121.0516	.5155442	1.011714	1.007772
	N	23	23	23	23	23
Insurer	min	25.87344	57.33937	.4985904	1.000122	1.29506
	mean	50.62006	81.07557	.5001038	1.000334	1.665488
	max	79.63468	104.954	.5016173	1.000548	2.071941
	N	8	8	8	8	8
Vertical	min	1.419893	22.18529	.	.	.
	mean	18.75004	61.55799	.	.	.
	max	48.65353	112.6702	.	.	.
	N	23	23	.	.	.
Separate	min	-35.10755	30.98623	.4890355	1.00423	1.004049
	mean	.3763779	62.35931	.5002004	1.005243	1.007355
	max	29.61468	100.3145	.5030982	1.008892	1.021511
	N	23	23	23	23	23

Table 5.9: **Estimated marginal costs by hospital**

hospital	stats	marginal cost	service price
Tabancura	min	9.625797	48.48587
	mean	24.48675	84.99908
	max	55.53538	110.8401
Bicentenario	min	-8.159	37.49257
	mean	5.512339	58.58128
	max	18.29834	70.64651
Avansalud	min	-9.683437	22.18529
	mean	1.242695	49.99243
	max	26.77224	75.24006
Indisa	min	25.87344	57.33937
	mean	58.11441	86.98462
	max	97.00656	104.954
Santa Maria	min	-4.937314	67.5334
	mean	29.54708	96.78509
	max	64.44559	121.0516
Davila	min	-25.66529	44.8206
	mean	6.565036	63.90128
	max	21.26764	89.25868
Vespucio	min	-35.10755	30.98623
	mean	5.277192	52.02614
	max	28.50003	77.56956

Table 5.10: **Estimated and observed premiums**

	premiums	
stats	**estimated**	**observed**
min	1.268384	1.086392
mean	2.897706	2.676648
max	5.548181	5.813891
sd	.6992822	1.005451
N	581	581

Chapter 6

Results

The demand estimations indicate that consumers are significantly more sensitive to premiums than to hospital prices, with similar results reported by Gowrisankaran et al. (2015). This is likely to be the result of a strong preference for quality in medical providers and the effect of reduced information regarding prices when changing providers.

The bargaining ratios estimated in table 5.8 show that insurers have little to loose under disagreement, as their agreement profits is less than a percent larger than their disagreement. This result is natural in a setting in which provider networks are very large and insurer are allowed to adapt their premiums after prices have been negotiated. This timing, together with the consumers large premium elasticity, enables an insurer that has lost a provider to adjust his premium configuration to compensate for almost all of the consumer's lost utility from the network. Given that there are only a few insurers, each of which offers plans with different preferential providers and coinsurance rates, which do not change under disagreement, it requires only a slight change in premiums to recover the loss generated by a disagreement. This characteristic of the market might be one of the reasons why vertical integration is widespread in the Chilean health care market.

Using the estimated marginal costs, I conduct a counterfactual analysis of the market in order to assess the effect of vertical integration on prices, premiums and consumer welfare. For each provider-insurer pair in which at least one member is integrated, I use the mean marginal cost for the provider to account for possible reductions in marginal costs due to integration. For the same no-longer-integrated pairs I set the bargaining weights to 0.5, which is the mean bargaining weight found for non-integrated pairs as shown in table 5.8. I restrict the analysis to 2014 in order to analyze the results on data that is closer to the current state of the market. The estimated model is used to recover negotiated prices from marginal cost and market composition using equation (3.19). Given that the vertically integrated insurers are associated to hospital systems I analyze an additional counterfactual scenario in which all holdings are prohibited in the market, breaking both vertical and horizontal mergers.

Table 6.1 contrasts the observed prices with the resulting negotiated prices under both counterfactual scenarios. The results indicate that hospital prices would drop in approximately 25% if vertical integration were prohibited, with integrated hospitals decreasing their prices by nearly 30% while the non-integrated hospital increases by 10%. This results are consistent with the estimated bargaining model, as integrated hospitals have access to their insurers leverage when negotiating prices. Given the estimated bargaining ratios, this additional leverage allows providers to negotiated higher prices with insurers, effect that disappears without vertical integration and that creates the reduction in prices. Furthermore, the scenario without holding shows that removing horizontal mergers as well does not have a significant effect in prices, besides making the decrease more homogeneous among providers. This indicates that hospital systems have little additional leverage when negotiating prices, but rather use their weight in order to obtain slightly better prices in one hospital in expense of others, hence improving its profit through demand redistribution.

Table 6.2 shows the variation in premiums between the different scenarios. The model indicates that removing vertical and horizontal mergers would result in a 0.05% reduction in premiums, on average. Table 6.3 show the related change in plan expected costs. These results indicate that although insurers costs reduce by about 3.5% under the counterfactual scenarios, only a small fraction is passed through to consumers premiums. As reduced prices also increase each plan's network value with respect to the outside good, downstream competition only reduces premiums slightly.

Finally, table 6.4 shows the expected utility from the market for each consumer, before knowing his preference shocks. The results show that the average consumer would be willing to pay about 0.04 UF per month (or $18 USD, per year) in order to prohibit vertical integration in the market, with older consumers benefiting more from the changes than younger consumers. Further removing horizontal integration has little effect on expected utility. The low value of this change in utility contrasts with the large reduction in service prices of table 6.1. However, as consumers are fairly insensitive to service prices and make scarce use of their insurance (as shown by their risk, table 5.1) this change is translated into small expected utility variation between scenarios.

Table 6.1: **Observed and counterfactual service prices by hospital (2014)**

			No vertical integration		No holdings	
hospital	**stats**	**Observed**	**Price**	**%ΔPrice**	**Price**	**%ΔPrice**
Tabancura	min	59.99378	62.24249	-42.28	56.82426	-47.31
	mean	89.38496	71.95004	-11.89	59.32709	-28.68
	max	107.8466	79.818	33.04	60.72169	.73
Bicentenario	min	44.11725	29.09902	-46.29	35.46137	-46.60
	mean	60.34236	34.27172	-42.11	39.96635	-30.01
	max	70.49408	37.86177	-34.04	43.36128	-1.71
Avansalud	min	43.96718	30.01095	-60.11	31.38297	-58.28
	mean	58.5008	33.31999	-38.67	35.87686	-34.43
	max	75.24006	39.8031	-9.47	39.24611	-10.73
Indisa	min	61.80044	90.48568	-13.28	90.50787	-13.28
	mean	89.09589	94.54094	11.57	94.57014	11.60
	max	104.954	102.1217	46.41	102.1967	46.45
Santa Maria	min	80.25912	68.25677	-42.05	62.22996	-47.17
	mean	102.4068	80.38025	-17.96	64.56143	-35.07
	max	117.7997	90.26297	12.46	66.18914	-18.68
Davila	min	50.11496	36.11205	-47.38	36.4598	-46.87
	mean	64.17444	43.69185	-31.80	41.24084	-33.88
	max	73.77632	58.61391	-20.55	45.15465	-15.97
Vespucio	min	40.13175	15.64643	-73.85	35.51608	-49.11
	mean	56.58815	27.82566	-47.06	39.18116	-26.12
	max	69.794	34.34198	-16.55	41.01849	2.18
Total	min	40.13175	15.64643	-73.85	31.38297	-58.28
	mean	74.35621	55.14007	-25.42	53.53198	-25.22
	max	117.7997	102.1217	46.41	102.1967	46.45

No holding implies no vertical or horizontal mergers.

Table 6.2: **Counterfactual premiums by insurer (2014)**

		Initial	No vertical integration		No holdings	
insurer	**stats**	**Premium**	**Premium**	**%ΔPremium**	**Premium**	**%ΔPremium**
Colmena	min	1.376124	1.374624	-.97	1.374658	-1.04
	mean	2.505657	2.501977	-.14	2.50179	-.15
	max	3.556799	3.542699	.06	3.542288	.07
Banmedica	min	1.622577	1.622135	-.41	1.622254	-.49
	mean	3.020949	3.023617	.07	3.022761	.05
	max	4.72376	4.723383	.51	4.720157	.48
Consalud	min	1.515241	1.512975	-1.02	1.512912	-.99
	mean	2.89583	2.893398	-.07	2.893533	-.07
	max	4.731433	4.709512	1.07	4.71529	1.06
Total	min	1.376124	1.374624	-1.02	1.374658	-1.04
	mean	2.807478	2.806331	-.049	2.806028	-.05
	max	4.731433	4.723383	1.07	4.720157	1.06

Initial premiums are the ones obtained from the bargaining estimation.

Table 6.3: **Counterfactual plan expected costs by insurer (2014)**

		Initial	No vertical integration		No holdings	
insurer	**stats**	**Cost**	**Cost**	**%ΔCost**	**Cost**	**%ΔCost**
Colmena	min	.0658484	.0624266	-5.51	.0623575	-5.78
	mean	.2626364	.2495737	-5.04	.2491916	-5.18
	max	.8884657	.8472264	-4.23	.8451995	-4.32
Banmedica	min	.087412	.0832871	-5.18	.0832807	-5.45
	mean	.3237658	.3095446	-4.46	.3087848	-4.67
	max	.8722713	.8413304	-3.54	.8388164	-3.83
Consalud	min	.061253	.0609356	-2.49	.0611171	-2.46
	mean	.2113054	.2094159	-.88	.2097033	-.76
	max	.5869221	.5783622	.28	.5854141	-.16
Total	min	.061253	.0609356	-5.51	.0611171	-5.78
	mean	.2659026	.2561781	-3.46	.2558932	-3.54
	max	.8884657	.8472264	.28	.8451995	-.16

Initial cost is estimated using the hospital demand and observed risks.

Table 6.4: **Mean ex-ante market expected utility by age group, in UF (2014)**

Age group	**Initial Utility**	**No vertical integration Utility**	**No vertical integration ΔUtility**	**No holdings Utility**	**No holdings ΔUtility**
20	1.774584	1.789972	.0153884	1.790652	.016068
25	2.639134	2.66156	.0224256	2.662607	.0234722
30	3.520963	3.554966	.0340032	3.556622	.0356588
35	3.185414	3.217741	.0323272	3.219393	.0339786
40	2.083331	2.112623	.0292926	2.11402	.0306893
45	2.151522	2.181747	.0302255	2.183147	.0316258
50	2.905189	2.940732	.035543	2.942132	.0369433
55	3.54015	3.584294	.0441448	3.586053	.0459032
60	5.126891	5.217888	.0909971	5.221337	.0944459
Population mean	3.539208	3.582812	.0436039	3.584698	.0454895

Initial expected utility is the one obtained from the bargaining estimation.

Chapter 7

conclusions

This work formulated a model for studying vertical oligopolistic markets, in which upstream prices are set through negotiations and downstream premiums are determined by Bertrand competition. It shows that downstream equilibrium premiums follow an intuitive expression that combines opportunity costs, expected insurance costs and price elasticity. Additionally, upstream prices are shown to generalize standard Bertrand pricing of differentiated products with heterogeneous consumers, where there is disagreement regarding the marginal effect of increasing prices and the optimal level of demand. The model makes explicit the effects of vertical and horizontal integration in the market and empirically shows the effect that different integration structures have on equilibrium prices.

The model is estimated using detailed data from the Chilean private health care market. I find that prohibiting vertical integration would reduce hospital prices by 25% in average, while keeping premiums almost unchanged. This indicates that policy seeking to limit the exercise of market power in the market through restrictions on premium variations have little effect when vertical integration is in place. The results show that integrated firms use their market power in order to increase profits from upstream rather than downstream demand, as consumers are more sensible to premiums than to service prices.

As many industries follow the structure of the market studied here, the model can be generally applied to study markets in which downstream competition takes place over a series of homogeneous markets. Work is left to be done in improving plan demand estimates, including consumers that chose the public insurance which might shift premium elasticities upwards. This would also allow to include more of the market when estimating marginal costs, providing more accurate counterfactual results. The possible effect of eliminating vertical integration on adverse selection is not studied here. As a fraction of the riskier publicly insured population might be willing to switch to the private insurance once service prices drop, this might make plans more costly to insurers and induce an increase in premiums, making healthier consumers to drop out of the private insurance system. However, given the estimated price and premium elasticities, it is unlikely that this effect be significant as premiums are the main determinant of the consumer's enrollment decision and they show little variation between the integrated and non-integrated regime.

Chapter 8

Bibliography

Atal, J. P. (2015). Lock-in in dynamic health insurance contracts: Evidence from chile.

Bajari, P., Benkard, C. L., and Levin, J. (2007). Estimating dynamic models of imperfect competition. *Econometrica*, 75(5):1331–1370.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, pages 841–890.

Bresnahan, T. F. (1987). Competition and collusion in the american automobile industry: The 1955 price war. *The Journal of Industrial Economics*, pages 457–482.

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

Chipty, T. (2001). Vertical integration, market foreclosure, and consumer welfare in the cable television industry. *American Economic Review*, pages 428–453.

Collard-Wexler, A., Gowrisankaran, G., and Lee, R. S. (2015). Nash-in-nash bargaining: A microfoundation for applied work. *NBER Working Paper*, 20641.

Crawford, G. S., Lee, R. S., Whinston, M. D., and Yurukoglu, A. (2015). The welfare effects of vertical integration in multichannel television markets.

Crawford, G. S. and Yurukoglu, A. (2012). The welfare effects of bundling in multichannel television markets. *The American Economic Review*, 102(2):643–685.

Duarte, F. (2012). Price elasticity of expenditure across health care services. *Journal of health economics*, 31(6):824–841.

Fox, J. T., Ryan, S. P., Bajari, P., et al. (2011). A simple estimator for the distribution of random coefficients. *Quantitative Economics*, 2(3):381–418.

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

Handel, B., Hendel, I., and Whinston, M. D. (2015). Equilibria in health exchanges: Adverse selection versus reclassification risk. *Econometrica*, 83(4):1261–1313.

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

Ho, K. and Lee, R. S. (2013). Insurer competition in health care markets.

Horn, H. and Wolinsky, A. (1988). Bilateral monopolies and incentives for merger. *The RAND Journal of Economics*, pages 408–419.

Lee, R. S. and Fong, K. (2013). Markov perfect network formation: An applied framework for bilateral oligopoly and bargaining in buyer-seller networks.

Palmucci, G. A. and Dague, L. (2015). The welfare effects of banning risk-rated pricing in health insurance markets: Evidence from chile. *Working Paper*.

Train, K. (2015). Mixed logit with a flexible mixing distribution. *Working Paper*.

Chapter 9

Appendix

9.1 Proof of equation (3.7)

Given that each market is independent and each plan is marketed in only one market (or can be considered as separate plans with different prices) it suffices to consider only the market to which the plan we wish to optimize belongs. Let that market be denoted as I. Denote k as the plan we wish to optimize and m denote its owner insurer. Hence $\phi_k^* \in \arg\max_{\phi_k} \sum_{r\in K_{mI}} s_r(\phi_r - \mathrm{ec}_r)$. The proofs builds from the first order condition as follow:

$$
\begin{aligned}
\sum_{r\in K_{mI}} \frac{\partial s_r}{\partial \phi_k}(\phi_r - \mathrm{ec}_r) + s_k = 0 \\
-\sum_{r\in K_{mI}, r\neq k} \alpha_1 s_r s_k(\phi_r - \mathrm{ec}_r) + \alpha_1(s_k - s_k^2)(\phi_k - \mathrm{ec}_k) + s_k = 0 \\
\sum_{r\in K_{mI}, r\neq k} s_r(\phi_r - \mathrm{ec}_r) - (1 - s_k)(\phi_k - \mathrm{ec}_k) - \alpha_1^{-1} = 0 \\
\sum_{r\in K_{mI}, r\neq k} \mathrm{e}^{\delta_r}(\phi_r - \mathrm{ec}_r) - (\sum_{r\in K_I, r\neq k} \mathrm{e}^{\delta_r})(\phi_k - \mathrm{ec}_k) - \alpha_1^{-1}\sum_{r\in K_I} \mathrm{e}^{\delta_r} = 0 \\
\sum_{r\in K_{mI}, r\neq k} \mathrm{e}^{\delta_r}(\phi_r - \mathrm{ec}_r) + (\sum_{r\in K_I, r\neq k} \mathrm{e}^{\delta_r})\mathrm{ec}_k - \alpha_1^{-1}\sum_{r\in K_I, r\neq k} \mathrm{e}^{\delta_r} = \sum_{r\in K_I, r\neq k} \mathrm{e}^{\delta_r}\phi_k + \alpha_1^{-1}\mathrm{e}^{\delta_k}
\end{aligned}
$$

Now define

$$
\begin{aligned}
\omega &= \alpha_1 \sum_{r\in K_{mI}, r\neq k} \mathrm{e}^{\delta_r}(\phi_r - \mathrm{ec}_r) + (\sum_{r\neq k, r\in K_I} \mathrm{e}^{\delta_r})(\alpha_1 \mathrm{ec}_k - 1) \\
\bar{\phi}_k &= \alpha_1 \phi_k \\
a &= \sum_{r\in K_I, r\neq k} \mathrm{e}^{\delta_r}
\end{aligned}
$$

With this the problem can be written simply as

$$\omega - a\bar{\phi}_k = \exp(\bar{\phi}_k + x_k)$$

Where $\delta_k = \bar{\phi}_k + x_k$. Now consider the change of variables $\bar{\phi}_k = -t + \frac{\omega}{a}$ with which the problem becomes

$$\begin{aligned} \omega + at - \omega &= \exp(-t + \frac{\omega}{a} + x_k) \\ t\exp(t) &= a^{-1}\exp(\frac{\omega}{a} + x_k) \end{aligned}$$

And by the definition of the Lambert W, this is implies

$$t = W[a^{-1}\exp(\frac{\omega}{a} + x_k)]$$

and by simply reordering and renaming terms we find the result:

$$\begin{aligned} -\alpha_1\phi_k + \frac{\omega}{a} &= W[a^{-1}\exp(\frac{\omega}{a} + x_k)] \\ \phi_k &= \pi_{m|k} + \mathrm{ec}_k - \alpha_1^{-1} - \alpha_1^{-1}W[\lambda_k] \end{aligned}$$

9.2 Proof of equation (3.8)

This results follow from the differentiating taking into account that the Lambert W derivative satisfies

$$\frac{\partial W(x)}{\partial x} = \frac{W(x)}{x(1+W(x))}$$

Hence when differentiating the equation

$$\phi_k = \pi_{m|k} + \mathrm{ec}_k - \alpha_1^{-1} - \alpha_1^{-1}W[\lambda_k]$$

we obtain

$$\begin{aligned} \frac{\partial\phi_k}{\partial p_{mj}} &= \frac{\partial\pi_{m(k)|k}}{\partial p_{mj}} + \mathbb{1}_{\{m(k)=m\}}\frac{\partial \mathrm{ec}_k}{p_{mj}} - \alpha_1^{-1}\frac{W(\lambda_k)\frac{\partial\lambda_k}{\partial\lambda_k}}{\lambda_k(1+W(\lambda_k))} \\ (1+W(\lambda_k))\frac{\partial\phi_k}{\partial p_{mj}} &= \frac{\partial\pi_{m(k)|k}}{\partial p_{mj}} + \mathbb{1}_{\{m(k)=m\}}(\frac{\partial \mathrm{ec}_k}{\partial p_{mj}} - \alpha_1^{-1}W(\lambda_k)\frac{\partial x_k}{p_{mj}}) \\ &\quad + \alpha_1^{-1}W(\lambda_k)(\sum_{r\in K_I, r\neq k}(\alpha_1\frac{\partial\phi_r}{\partial p_{mj}} + \mathbb{1}_{\{m(r)=m\}}\frac{\partial x_r}{\partial p_{mj}})s_{r|k}) \end{aligned}$$

By developing the derivatives we get

$$\begin{aligned}
\frac{\partial \phi_k}{\partial p_{mj}}(1+W(\bar{\lambda}_k)) = & \sum_{r\in K_{m(k)I}, r\neq k} (\alpha_1(\phi_r - \mathrm{ec}_r)+1)s_{r|k}\frac{\partial \phi_r}{\partial p_{mj}} \\
& - \pi_{m(k)|k}\sum_{r\in K_I, r\neq k}\alpha_1\frac{\partial \phi_r}{\partial p_{mj}}s_{r|k} + W(\bar{\lambda}_k)\sum_{r\in K_I, r\neq k}\frac{\partial \phi_r}{\partial p_{mj}}s_{r|k} \\
& + \mathbb{1}_{\{m(k)=m\}}\sum_{r\in K_{m(k)I}, r\neq k} s_{r|k}((\phi_r - \mathrm{ec}_r)\frac{\partial x_r}{\partial p_{mj}} - \frac{\partial \mathrm{ec}_r}{\partial p_{mj}}) \\
& - (\sum_{r\in K_{mI}, r\neq k}\frac{\partial x_r}{\partial p_{mj}}s_{r|k})(\sum_{r\in K_{m(k)I}, r\neq k} s_{r|k}(\phi_r - \mathrm{ec}_r) \\
& + \mathbb{1}_{\{m(k)=m\}}\frac{\partial \mathrm{ec}_k}{\partial p_{mj}} - \mathbb{1}_{\{m(k)=m\}}\alpha_1^{-1}W(\bar{\lambda}_k)\frac{\partial x_k}{\partial p_{mj}} \\
& + \alpha_1^{-1}W(\lambda_k)\sum_{r\in K_{mI}, r\neq k}\frac{\partial x_r}{\partial p_{mj}}s_{r|k} \\
= & \sum_{r\in K_{m(k)I}, r\neq k} (\alpha_1(\phi_r - \mathrm{ec}_r)+1)s_{r|k}\frac{\partial \phi_r}{\partial p_{mj}} - \pi_{m(k)|k}\sum_{r\in K_I, r\neq k}\alpha_1\frac{\partial \phi_r}{\partial p_{mj}}s_{r|k} \\
& + W(\bar{\lambda_k})\sum_{r\in K_I, r\neq k}\frac{\partial \phi_r}{\partial p_{mj}}s_{r|k} + \bar{\mathbf{R}}_I(m,j)
\end{aligned}$$

Which can be stacked over the market to get the result of equation (3.8).

9.3 Proof of equation (5.1)

In equation (3.7) we have that

$$\phi_k^* = \pi_{m|k} + \mathrm{ec}_k - \alpha_1^{-1}(1+W(\lambda_k))$$

With

$$\lambda_k = (\sum_{j\in K_I,\, j\neq k}\exp(\delta_j))^{-1}\exp(\alpha_1\pi_{m|k} + \alpha_1\mathrm{ec}_k - 1 + xk)$$

Multiplying the first equation by α_1 we find that $\alpha_1\phi_k + W(\lambda_k) = \alpha_1\pi_{m|k} + \alpha_1\mathrm{ec}_k - 1$. Substituting this in the second equation we get,

$$\lambda_k = (\sum_{j\in K_I,\, j\neq k}\exp(\delta_j))^{-1}\exp(\alpha_1\phi_k + W(\lambda_k) + xk)$$

The Lambert W satisfies the identity $\exp(W(x)) = \frac{x}{W(x)}$ making the previous equation equal to

$$\lambda_k = (\sum_{j\in K_I,\, j\neq k}\exp(\delta_j))^{-1}\exp(\alpha_1\phi_k + xk)\frac{\lambda_k}{W(\lambda_k)}$$

Hence

$$W(\lambda_k) = \frac{\exp(\alpha_1 \phi_k + x_k)}{\sum_{j \in K_I, j \neq k} \exp(\delta_j)}$$

Substituting this result in (3.7) we get

$$\phi_k - \mathrm{ec}_k = \pi_{m|k} - \alpha_1^{-1}(1 + \frac{\exp(\alpha_1 \phi_k + x_k)}{\sum_{j \in K_I, j \neq k} \exp(\delta_j)})$$

multiplying and dividing the last fraction by the sum of exponential utilities for the market in order to convert the observed market shares

$$\phi_k - \mathrm{ec}_k = \pi_{m|k} - \alpha_1^{-1} \frac{1}{1 - s_k}$$

$$(\phi_k - \mathrm{ec}_k)(1 - s_k) = \pi_{m|k}(1 - s_k) - \alpha_1^{-1}$$

$$\phi_k - \mathrm{ec}_k = \pi_{mI} - \alpha_1^{-1}$$

multiplying both sides by plan k's market share and summing this result over all of the insurers plans we obtain

$$\pi_{mI} = (\sum_{k \in K_{mI}} s_k)(\pi_{mI} - \alpha_1^{-1})$$

And substituting this result in the equation above we get

$$\phi_k = \mathrm{ec}_k - \alpha_1^{-1} \frac{1}{1 - \sum_{r \in K_{mI}} s_r}$$

9.4 Additional Tables

Table 9.1: Hospital demand estimation

		specification		
coefficient	statistic	(I)	(II)	(III)
price	mean	1.382377e+02	1.329385e+02	4.791137e+01
		(3.106822e+03)	(3.376525e+05)	(4.769911e+02)
	stdev	2.744543e+01	2.628451e+01	8.968789e+00
		(1.844393e+02)	(8.078197e+06)	(1.373203e+01)
same comuna	mean	-3.487545e-02	-1.887273e-01	7.832899e-02
		(6.078831e-04)	(2.652652e+00)	(5.283476e-01)
	stdev	-1.119754e-02	-3.141746e-02	-1.575875e-01
		(3.484369e-04)	(1.065826e-01)	(7.056639e-02)
preferential	mean	1.044452e-01	3.598100e-01	4.114035e-01
		(8.187793e-04)	(1.330287e-01)	(1.882332e-01)
	stdev	1.246922e-01	2.159292e-01	-4.816215e-02
		(2.103952e-04)	(4.034790e-01)	(9.750176e-02)
private	mean	1.736944e-01		2.301079e-01
		(1.302318e-03)		(1.919182e-01)
	stdev	-1.772474e-01		7.459028e-03
		(1.267047e-03)		(2.634368e-02)
System Effect		No	Yes	No
Provider Effect		No	No	Yes
Bootstrap choices		100000	100000	1000

Table 9.2: Plan demand estimation

		specification	
coefficient	statistic	(I)	(II)
price	mean	5.352271e+00	1.818962e+01
		(1.095570e-02)	(1.295172e-02)
	stdev	1.170387e+00	-7.377506e+00
		(3.119772e-03)	(8.960428e-02)
plan value	mean	6.670060e+00	-1.211190e+00
		(1.544544e-01)	(2.603475e-06)
	stdev	1.480524e-01	9.488028e-01
		(7.633269e-05)	(2.564991e-03)
Insurer Effects		No	Yes
Bootstrap choices		400000	1000