Network Formation and Bargaining in Vertical Markets: The Case of Narrow Networks in Health Insurance

Soheil Ghili *

Abstract

This paper examines the effects of regulations that force insurers to include more hospitals as in-network providers. Such regulations can help consumers by increasing access to hospitals, but can also increase hospitals bargaining power when negotiating with insurers, thereby leading to higher reimbursement rates and insurance premiums. I develop and structurally estimate a model that endogenously captures how insurers (i) form hospital networks, (ii) bargain over rates with hospitals, and (iii) set premiums. Crucially, my bargaining formulation allows insurers to threaten to replace an in-network hospital with an out-of-network one. I use the model to empirically analyze the effects of these proposed regulations on a health insurance exchange from Massachusetts. I find that tighter regulations, which force insurers to include more than 85% of the hospital systems in the market, raise the average reimbursement rates paid by some insurers by at least 28%.

*Yale University. Email: soheil.ghili@yale.edu

1 Introduction

Insurers increasingly use “narrow network” plan designs in the health insurance exchanges established by the Affordable Care Act (ACA). These plans offer their enrollees a small set of “in-network” hospitals to choose from, and if the enrollee decides to go “out of network,” s/he will have to pay all or most of the healthcare expenses out of pocket.[1,2] Narrow hospital networks have received vast and often unfavorable media attention. Some provider- as well as consumer-advocacy groups have filed lawsuits against both insurers and the federal government regarding narrow networks. They have also pressed the administrators of the healthcare exchanges, mainly Centers for Medicare and Medicaid Services (CMS), for “network adequacy regulations,” which are regulations that would force minimum mandated network sizes on insurers in order to expand narrower networks and increase consumers' access to hospitals.[3]

However, the impacts of network adequacy regulations are not confined to hospital access. They also can have *price consequences*. Under such regulations, insurers cannot drop a hospital out of network if that action would take them below the minimum mandated network size. Taking advantage of this restriction faced by insurers, hospitals may gain further bargaining leverage and successfully negotiate higher reimbursement rates with insurers for the care they provide to patients. Insurers can, in turn, respond to the increased rates by raising their monthly premiums and passing the extra cost on to consumers.

The objective of this paper is to quantify the price consequences of network adequacy regulations. I do this by developing an empirically estimable model of insurer-hospital markets with two crucial features. First, I endogenize how insurers (i) form hospital networks, (ii) bargain with hospitals over rates, and (iii) set premiums. Second, my formulation of the bargaining procedure allows insurers to try to negotiate lower rates with hospitals not

[1] According to McKinsey & Company, about 60% of the ACA exchange plans cover less than 70% of local hospitals where their plans are offered, and about 20% of the plans cover less than 30% of local hospitals. Networks are expected to further narrow in 2017.

[2] Narrow networks are not confined to hospitals. The same is true of other types of providers, such as physician groups and prescription drugs. My focus in this paper is on hospitals.

[3] In fact, CMS had planned on introducing such regulations for the ACA exchanges in 2017. But in its final rule, it chose not to do so.

only by threatening to drop them from the hospital network, but also by threatening to replace them with currently out-of-network hospitals. I leverage detailed medical claims and insurance plan enrollments data to structurally estimate the model for the CommCare market, a health insurance exchange in Massachusetts, predating the ACA but similar to the ACA exchanges in many respects. I then use the estimated model to simulate a range of network adequacy regulations and study how they impact hospital networks, negotiated rates, and premiums. I identify and analyze multiple economic forces that govern how narrow network plans respond to the regulation as well as how broader network plans, which are not directly affected by the regulation, respond to the responses of narrower network ones.

My model consists of a game with two main steps. In the first step, equilibrium hospital networks and reimbursement fees are simultaneously determined. In the second step, insurers set premiums la Bertrand. To model the first step, I impose *network stability conditions* and *bargaining conditions* on the equilibrium hospital networks and reimbursements. For network stability, I adapt pairwise stability conditions from Jackson and Wolinsky [1996]. I require that at the equilibrium, no insurer or hospital can strictly profit either from dropping a bilateral contract or from adding a new bilateral contract that the other party would be willing to sign.

My bargaining conditions offer a substantial advantage over standard approaches to the modeling of bargaining in vertical markets, by accounting for an economic force that arises in settings in which the bargaining does not take place on a fixed network, but rather affects and is affected by network formation. The standard model of bargaining on a network in the empirical industrial organization literature is called Nash-in-Nash (NiN henceforth), and is based on Horn and Wolinsky [1988]. NiN assumes that the reimbursement rate negotiated between hospital i and insurer j is the outcome of a Nash Bargaining procedure (la Binmore et al. [1986]) between the two, *taking as given* the rest of the network structure and reimbursements. The construction of the Nash Bargaining formulation roughly implies that hospital i can charge a high rate if insurer j would lose a substantial amount of profit from leaving hospital i out of its hospital network, but not modifying the network otherwise. This is the case under NiN *even* if insurer j would lose little or no profit, or would gain some profit, from leaving hospital i out and instead accepting a competing offer made by a currently-out-of-network hospital i' that is trying to outbid hospital i and *replace it* in the network, in a negotiation process *simultaneous* with the negotiation between i and j. In other words, the negotiated reimbursement rate predicted by NiN is *not affected at all* by which hospitals are outside the network of insurer j and how close substitutes they are to hospital i.

To capture insurers' ability to use such out-of-network hospitals as "threats of replacement" is critical in the context of studying network adequacy regulations. For an insurer that is just meeting the minimum network size mandated by a network adequacy regulation, dropping hospitals without replacement is not an option. Thus, for such insurers, threats of replacement are the only bargaining chip that can help to keep reimbursement rates low. Not accounting for these threats, NiN would predict unrealistically high charges by hospitals to insurers bound by such regulations. My bargaining formulation, instead, assumes that the rate negotiated between hospital i and insurer j is equal to the Nash Bargaining rate *only if* under that rate, the insurer cannot strictly profit from replacing i with a currently out-of-network i' at the lowest rate that i' would accept. If, however, under the Nash Bargaining rate, insurer j can strictly profit from such replacement, my bargaining formulation assumes that hospital i brings its rate down to a level that would make the insurer indifferent between keeping hospital i and replacing it with hospital i'.[4] I show that my formulation can be sustained as an outcome of a Subgame Perfect Nash Equilibrium of a non-cooperative extensive form game for a class of 2×1 (i.e., two upstream firms and one downstream) settings.

My empirical analysis has several steps, in which I estimate demand and cost functions for insurers and hospitals. First, I use detailed medical claims data on CommCare to estimate a model of hospital choice, backing out perceived hospital qualities by patients for different diagnoses as well as patients' disutility from travel. Based on these estimates, I construct network-expected-utility measures for each CommCare plan across different en-

[4] That is, hospital i will charge the highest rate that deters threats of replacement by the insurer. Of course, if such a rate is so low that hospital i would rather drop the contract, then pairwise stability is violated and current network structure cannot be part of the equilibrium configuration. Hence, the replacement must take place. See section 3 for more details.

rollee demographic groups. I then use the network-expected-utility measures along with data on other plan characteristics and on CommCare enrollments to estimate a demand model for insurance plans. I estimate marginal costs of inpatient care to hospitals using hospital cost reports, and to insurers, using the observed payments in my CommCare medical claims data. Finally, I combine all of these estimates and impose the structure of my model to back out relative bargaining powers for insurers and hospitals, as well as insurers' fixed and variable non-inpatient costs.[5]

I use the estimated model to simulate a range of network adequacy regulations. I focus on the Greater Boston Area and the 2011 fiscal year. I consider regulations in the form of forcing all insurers in the market to cover at least $X\%$ of the hospital systems present in the market. A higher X, hence, means a tighter regulation. I simulate the regulation over a range of values for X, and examine the responses of the hospital networks, reimbursement rates, and premiums.

My main finding is that under tight regulations, with $X > 85\%$, affected insurers (i.e., those that are forced by the regulation to expand their networks) experience large increases in the average reimbursement rates they pay to their in-network hospitals. Some of these insurers respond by raising their monthly premiums.[6] For instance, Celticare, a CommCare plan that covered four out of the 16 hospital systems in 2011, is predicted to pay about \$4,000 (or 28%) more to its in-network hospitals per average hospital admission, when it is forced to cover at least 14 hospital systems. Celticare responds by raising its premium from \$404/month to \$425/month.[7] The key driving force behind this result is a *sorting* effect: the additional hospitals that insurers include in response to the regulation are those with the lowest cost and/or in highest demand. In other words, they are the best hospitals. This would leave the insurer with a pool of out-of-network hospitals comprising the worst $(100 - X)\,\%$ of the hospitals in the market. Thus, with high X, the threats of replacement that the insurer can make using its weakened out-of-network pool only become credible

[5] Sources of non-inpatient costs can be outpatient care, doctor visits outside of hospitals, prescription drugs, and administrative costs.

[6] Others do not raise premiums due to a premium cap which was imposed by the Massachusetts Health Connector, which administered CommCare.

[7] The premium could rise even further if there were no regulated premium cap at \$425/month in CommCare. I will discuss these premium regulations in section 4.

when the in-network hospitals, in particular the best ones among them, charge high rates.[8]

Besides the main result, my simulations also point to other interesting predictions about the functioning of health insurance markets under network adequacy regulations. I find that for lower ranges of X, in which the sorting effect that weakens the out-of-network pools of insurers is less stringent, network adequacy regulations may in fact *lower* the average reimbursement rates paid by some insurers. This happens when the hospitals added in response to the regulation have lower marginal costs of providing care than the insurer's pre-regulation hospitals. I also find that even though they expand the narrower hospital networks, in some cases network adequacy regulations *shrink* the broader ones. This happens because some hospitals that get to charge higher rates to a narrow network plan that is constrained by the regulation, are encouraged to abandon some of the broader hospital networks in order to steer their more loyal patients to the plan that pays those hospitals more for treating the patients. In section 6, I discuss these findings in more detail and argue that the overall effect of the regulation on consumers depends on specific market conditions.[9]

The framework I develop in this paper applies beyond network adequacy regulations. It can be applied to problems in which the interaction between network formation and bargaining is key. For example, it may be used to study how hospital-networks would respond to increased hospital bargaining power due to hospital consolidation. It may also be applied to other two-sided markets such as the market between TV channels and Cable companies.

The rest of the paper is organized as follows. Section 2 reviews the related literature. Section 3 sets up the model. Section 4 describes the CommCare market. Section 5 explains the estimation procedure and estimation results. Section 6 presents and interprets the counterfactual simulation results for CommCare and discusses potential differences from

[8]On top of this sorting effect, there are also more complex mechanisms that lead to rapid accelerations in the average reimbursement rate per admission that Celticare pays to its in-network hospitals and in Celticare's premium. Those are discussed in section 6.2.1.

[9]In the case of CommCare, even stringent regulations improve consumer welfare due to the very market-specific fact that in CommCare, premiums faced a mandatory cap set by the government. Thus, the increased reimbursement rates resulting from network adequacy regulations did not fully pass through to customers.

the ACA exchanges in terms of consumer-welfare implications. Section 7 concludes.

2 Related Literature

This paper can be considered part of the expanding body of literature in economics on issues related to the implementation of the Affordable Care Act. Recent papers have examined issues like subsidy schemes (Tebaldi [2015]), tiered networks (Prager [2015]), market competition and premiums (Dafny et al. [2015a]), and selection (Hackmann et al. [2015], Shepard [2015]). Some papers have studied issues directly related to narrow networks by comparing the health outcomes between enrollees of narrow and broad network plans (Gruber and McKnight [2014]), examining the relationship between network breadth and premium (Dafny et al. [2015b]), and estimating how much different groups of consumers value network breadth (Ericson and Starc [2014]). My paper takes another step by studying the consequences of regulating the networks for insurer-hospital bargaining and insurers premiums.

From a methodological standpoint, this paper also makes two contributions to the empirical analysis of vertical markets (e.g., Crawford and Yurukoglu [2012], Crawford et al. [2015], Gowrisankaran et al. [2013], Ho and Lee [2017b], Prager [2015], Pakes [2010], Ho [2009, 2006], Lee and Fong [2013]). First, I develop and empirically estimate a model that fully endogenizes network formation, bargaining, and downstream price setting. Second, by adding threats of replacement to NiN bargaining, I allow the upstream firms that a downstream firm excludes to impact its bargaining with those that it includes.

Also close in spirit to this paper are two recent working papers by Ho and Lee [2017a] and Liebman [2017]. As in this paper, they combine the bargaining model with a model of network formation, to predict a network structure and a set of prices charged by upstream firms (hospitals) to downstream firms (insurers). They also extend NiN to capture the importance of an insurer's out-of-network hospitals for its bargaining position with the in-network ones. The bargaining formulation in Ho and Lee [2017a] is more similar to this paper. One major difference between this paper and Ho and Lee [2017a] is in the network formation models. I model network formation by adapting pairwise-stability conditions from Jackson and Wolinsky [1996]. Ho and Lee [2017a], in contrast, employ a two-step model. In the first step, the insurer *commits* to a network. In the second step, the insurer bargains with the hospitals in that network according to NiN with threats of replacement. The choice of network in the first step is made to maximize the anticipated profit from the second-step bargaining.

A main feature of Ho and Lee's model is that an insurer can attain substantially lower reimbursement rates by committing to a narrow network: the insurer commits to excluding some hospitals in the first step (even if it pays off to include them) in order to use them as threats of replacement against included hospitals in the bargaining process in the second step. However, it is not clear why insurer commitment to exclusion should have the consequence for negotiated rates that Ho and Lee [2017a] describe. Threatening hospital i to replace it with hospital i' means threatening to form a network that includes i' but excludes i. But the insurer can always make this type of threat even without first committing to exclude hospital i'. If the insurer has gains from trade with both i and i', it can include both of them and then threaten to drop i. This would still be a threat to i of forming a network that includes i' but excludes i. In contrast, in this paper, an insurer always signs a contract with a hospital if there are gains from trade to the two.[10]

Another major difference from Ho and Lee [2017a] is estimation. This paper not only provides a model of network formation and bargaining, but also delivers a general estimation procedure to rationalize the observed networks and bargaining outcomes and obtain an estimated model for policy analysis.

[10]Note that some form of insurer commitment has been mentioned in the non-academic literature as a reason for narrow networks. But the mechanism is different from those in Ho and Lee [2017a] and Liebman [2017], which are based on commiting to exclude a set (or a number) of hospitals in order to use them as threats of replacement. The mechanism is that insurers promise their in-network hospitals high volume (through narrow networks) in exchange for low rates (see for instance, McKinsey [2013]) Usually the commitment is implemented relationally: if the insurer promises high volume and signs a contract with a low rate, but does not deliver the promised volume due to including other hospitals, then the original hospital will not be willing to charge the low rate next year. This kind of commitment story can be modeled using two-part tariff contracts between hospitals and insurers, and is not directly related to threats of replacement.

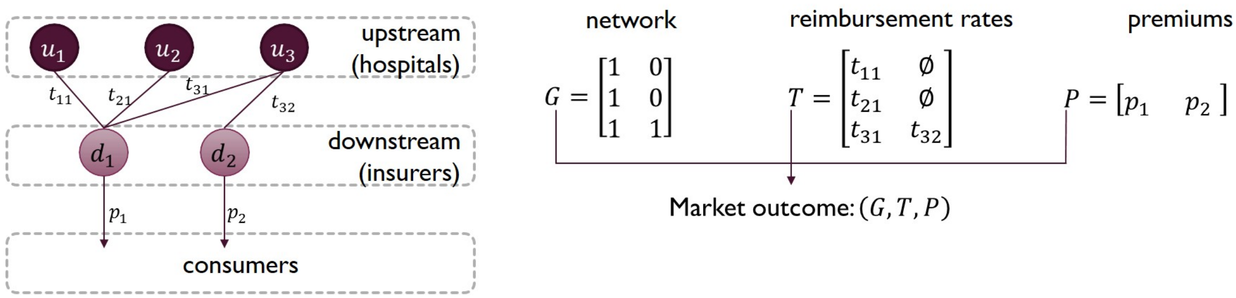

Figure 3.1: Schematic and matrix representations of the state of the market

3 Model

3.1 Basic Setup and Notations

My model captures interactions among three types of players in a vertical market: m hospitals (upstream firms, denoted u_i), n insurers (downstream firms, denoted d_j), and consumers (who are enrollees for insurers and patients for hospitals). The state of the market, or the "market outcome," is described by three important elements: network structure, denoted G; reimbursement rates, denoted T; and premiums, denoted P. Figure 3.1 exhibits the schematic and matrix representations of G, T and P through an example with three hospitals and two insurers.

Network G represents who contracts with whom, or equivalently, which hospitals are covered by each insurer. It can be represented graphically (as in figure 3.1) or using a matrix of zeros and ones. Reimbursements matrix T is also represented in a matrix form with elements t_{ij}. As is clear from the figure, for every inactive link $g_{ij} = 0$ in network G, the corresponding reimbursement rate is null: $t_{ij} = \emptyset$. For an active link $g_{ij} = 1$, the corresponding t_{ij} is interpreted as follows: insurer d_j reimburses hospital u_i with t_{ij} dollars for every unit of healthcare provided to an enrollee of d_j at hospital u_i. Also, each element p_j of P is the amount each enrollee of insurer d_j pays to that insurer per month.

I assume that the expected profit to each firm in the market is solely a function of the market outcome. That is, the profit functions are in the form of $\pi_{u_i}(G, T, P)$ and $\pi_{d_j}(G, T, P)$ for all u_i and d_j. The exact nature of the profit functions arises from demand and cost functions in the market, which I will turn to in later sections. My model of

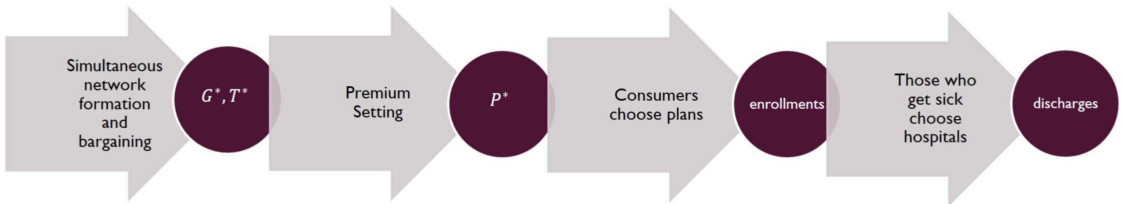

Figure 3.2: The four-step sequential game

the market takes the profit functions as primitives and predicts what market outcome(s) (G^*, T^*, P^*) will arise at the equilibrium.

This model consists of a sequential game with four main steps, as depicted in figure 3.2. In step 1, hospitals and insurers will engage in a "network-formation and bargaining" game. The outcome of this step is the equilibrium network and reimbursements (G^*, T^*). The structure of this step will be discussed in section 3.2. In step 2, insurers engage in a premium-setting game la Bertrand, taking into account the outcome of the previous step, (G^*, T^*). The outcome of this step is the set of equilibrium premiums P^*. In the third step, consumers decide which insurance plan to buy in the market. Finally, in the fourth step, if they get sick and need hospitalization, they decide which hospital within their plan's network of providers (as specified by G^*) to visit for treatment.

Step 2 in this game is fairly straightforward. Steps 3 and 4 are explained in more details in the estimation section. In the remainder of this section, I'll focus on step 1, where network formation and bargaining are endogenously captured.

3.2 Network-Formation and Bargaining Game

I combine two sets of conditions in order to characterize the equilibrium pair of network structure and reimbursements (G^*, T^*). The first set consists of *network-stability conditions*, which roughly require that no unilateral or joint deviation (among a pre specified set of deviations) from G^* be able to strictly pareto-improve the profits to the firms participating in the deviation. The second set of conditions consists of *bargaining conditions* which determine, for all ij with $g^*_{ij} = 1$, what t^*_{ij} the negotiation between hospital u_i and insurer d_j will result in.

The rest of this subsection is organized as follows. First I develop a base model with very simple network stability and bargaining conditions. I then demonstrate, using an example, that the base model is inadequate for capturing competition among hospitals for inclusion in insurers' hospital networks, which is crucial to my empirical analysis in this paper. I then modify the bargaining conditions to capture this force. I close this subsection by discussing my model's advantages over some other potential alternatives.

3.2.1 Base Model

I start by introducing the network stability conditions, which are adapted with small modifications from the notion of pairwise stability in Jackson and Wolinsky [1996].

Definition 1. The network-reimbursements pair (G^*, T^*) satisfies "pairwise network-stability" if the following hold:

(i) For $\forall g_{ij} = 1$, neither u_i nor d_j can strictly profit by unilaterally severing the link g_{ij}. That is:

$$\pi_{u_i}(G^*, T^*) \geq \pi_{u_i}\left(G^*_{-ij}, T^*_{-ij}\right) \tag{3.1}$$

$$\pi_{d_j}(G^*, T^*) \geq \pi_{d_j}\left(G^*_{-ij}, T^*_{-ij}\right) \tag{3.2}$$

(ii) For $\forall g_{ij} = 0$, there is no contract that u_i and d_j can sign that will yield a strict pareto improvement in their profits. That is:

$$\nexists t_{ij} \in \mathbb{R} \quad s.t. \quad \pi_{u_i}\left(G^*_{+ij}, T^*_{+t_{ij}}\right) \geq \pi_{u_i}(G^*, T^*) \ \& \ \pi_{d_j}\left(G^*_{+ij}, T^*_{+t_{ij}}\right) \geq \pi_{d_j}(G^*, T^*) \tag{3.3}$$

with at least one inequality holding strictly.

where $\left(G^*_{-ij}, T^*_{-ij}\right)$ is constructed from (G^*, T^*) by switching g_{ij} from 1 to 0 and t_{ij} to $\emptyset$; and $\left(G^*_{+ij}, T^*_{+t_{ij}}\right)$ is constructed by doing the exact inverse. Also, note that in this section the notation on P^* has been suppressed, as it has been assumed that P^* is anticipated by firms at the first stage of the game.

The bargaining conditions for the base model come from the NiN idea in Horn and Wolinsky [1988].

Definition 2. The network-reimbursement pairs (G^{*}, T^{*}) satisfies "NiN bargaining" relative to the "bargaining parameters matrix" $\gamma_{m\times n} \in [0,1]^{m\times n}$ if, for $\forall g_{ij} = 1$, we have $t^{*}_{ij} = t^{NB}_{ij}(G, T, \gamma_{ij})$, where $t^{NB}_{ij}(G, T, \gamma_{ij})$ is defined as:

$$\arg\max_{\tilde{t}\in\mathbb{R}} \left[\pi_{u_i}\left(G^{*}, \left(\tilde{t}, T^{*}_{-ij}\right)\right) - \pi_{u_i}\left(G^{*}_{-ij}, T^{*}_{-ij}\right)\right]^{\gamma_{ij}} \times \left[\pi_{d_j}\left(G^{*}, \left(\tilde{t}, T^{*}_{-ij}\right)\right) - \pi_{d_j}\left(G^{*}_{-ij}, T^{*}_{-ij}\right)\right]^{1-\gamma_{ij}} \tag{3.4}$$

where $\left(\tilde{t}, T^{*}_{-ij}\right)$ is constructed from T^{*} by substituting $\tilde{t}$ for its ij element.

The rough intuition behind the NiN conditions is that at the equilibrium, every ij pair is taking the rest of (G^{*}, T^{*}) as fixed, and negotiates over how to divide the total surplus created due to the presence of g_{ij}. The value t^{*}_{ij} is the one that divides this total surplus between u_i and d_j based on their respective bargaining parameters γ_{ij} and $1 - \gamma_{ij}$.[11]

Combining pairwise network stability and NiN bargaining, the base model gives us a solution concept that can predict equilibrium (G^{*}, T^{*}) pairs. So, thus far, we have a model that in principle does endogenously capture network formation and bargaining (and in the next stage, premium setting). Nevertheless, the particular way that the base model accomplishes this job may be a source of concern, which I turn to next.

3.2.2 Problems with the Base Model

It is well documented in the literature (Gowrisankaran et al. [2013], Lee and Fong [2013], etc.) that the implicit assumption in NiN bargaining (i.e., u_i and d_j taking the rest of (G, T) as fixed when bargaining with one another) may be a source of concern. I show, using a stylized example, that this concern is more pronounced when endogenous network formation is involved. My example focuses on a case in which an insurer is bound by a network adequacy regulation to cover at least one of the two hospitals (absent the regulation,

[11] Note that what the Nash Bargaining formulation formally does (which is to maximize a weighted product of the surpluses made by u_i and d_j, where the corresponding weights are γ_{ij} and $1 - \gamma_{ij}$), does not always lead to dividing the total surplus between the two according to shares of γ_{ij} and $1 - \gamma_{ij}$. Nevertheless, the surplus-division interpretation is useful for providing intuition about Nash Bargaining.

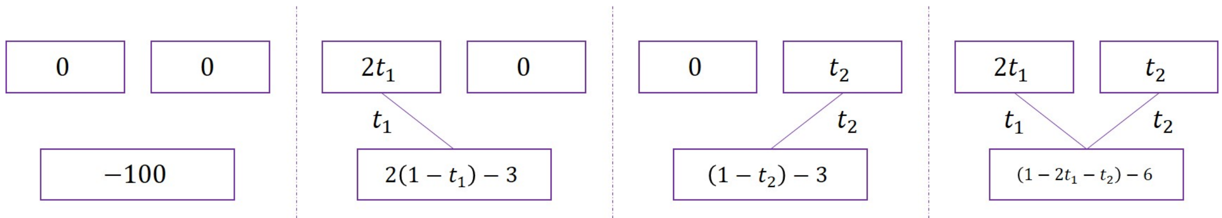

Figure 3.3: Profits as functions of network structure and reimbursements (G, T)

it would cover none[12]). I show that the base model fails to capture competition from out of network hospitals and, hence, will highly over-predict how much the hospital, that the insurer will cover to abide by the mandate, will be able to leverage the mandate and charge the insurer.

Setting of the example: There are two hospitals u_1 and u_2 and one insurer d. Bargaining parameters are symmetric: $\gamma = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$. The insurer's premium is equal to 1.[13] If u_1 is covered by d, it will bring 2 enrollees. u_2 brings 1 enrollee. Each enrollee consumes 1 unit of healthcare per month. There is no substitution between u_1 and u_2.[14] Covering each of the hospitals induces a large fixed cost of 3 to d. Finally, regulation would make d pay 100 in a fine if it did not cover any hospital. Figure 3.3 schematically presents the profits to all u_i and d as functions of the (G, T).[15]

Analysis of the example using the base model: it is fairly straightforward to verify that the base model predicts two equilibria: $(G^{1*}, T^{1*}) = \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 24.75 \\ \emptyset \end{bmatrix} \right)$ and $(G^{2*}, T^{2*}) = \left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} \emptyset \\ 49 \end{bmatrix} \right)$. Two features of this set of equilibria deserve further discussion. First, t_{11}^{1*} and t_{21}^{2*} both seem too high. The base model predicts these high

[12]This assumption is not realistic for a working insurance market. But it does capture the more general effect that I intend to highlight, while allowing for the setting to have only two hospitals, which significantly simplifies the discussion.

[13]For simplicity, I do not assume here that the insurer sets its premium optimally. This assumption is not crucial but simplifies a lot.

[14]That is, the profits that u_i and d make by signing a contract do not depend on whether u_{3-i} and d are also contracting or not. This assumption makes the exposition of the example simpler but is not crucial to the analysis and is not imposed on my empirical model.

[15]I continue to supress the notation on P, as it has been assumed constant here.

reimbursement rates because, for instance, at (G^{1*}, T^{1*}), hospital u_1 is taking advantage of the 100 that d would have to pay in a fine if the link g_{11}^{1*} broke off. What the base model does *not* capture is that, on the one hand, d can *replace* u_1 with u_2 for any $t_2 \geq 0$ (so that u_2 is willing to participate) and make a much higher profit than it makes under (G^{1*}, T^{1*}). The base model also does not capture that u_1 can foresee the aforementioned possibility and offer a much lower price than $t_{11}^{1*} = 24.75$ in order to deter this substitution threat. The second issue is that even though u_1 seems to dominate u_2 as a choice for d, we see that G^{2*} is sustained as part of an equilibrium. The reason is that at (G^{2*}, T^{2*}), neither dropping u_2 nor adding u_1 is beneficial to d, even though doing both (i.e., replacing u_2 with u_1) is. Next subsection introduces a model that addresses these issues.

3.2.3 Bargaining with threats of replacement

The objective of this subsection is to *minimally* expand on the bargaining conditions from the base model to deal with the issues illustrated above, while keeping the model *computationally tractable*. The basic intuition for the expansion is to allow the firm in danger of being substituted to anticipate this danger and bargain less aggressively. The definitions below formalize this idea.

Definition 3. Under (G, T), reimbursement rate t_{ij} is "safe for u_i" if $\nexists u_{i'}, t_{i'j}$ with $g_{i'j} = 0$ such that the following hold:

Profitability:

$$\pi_{d_j}\left(G^*_{-ij+i'j}, T^*_{-ij+t_{i'j}}\right) \geq \pi_{d_j}\left(G^*, T^*\right) \tag{3.5}$$

Incentive Compatibility:

$$\pi_{u_{i'}}\left(G^*_{-ij+i'j}, T^*_{-ij+t_{i'j}}\right) \geq \pi_{u_{i'}}\left(G^*, T^*\right) \tag{3.6}$$

No Commitment:

$$\pi_{u_{i'}}\left(G^*_{-ij+i'j}, T^*_{-ij+t_{i'j}}\right) \geq \pi_{u_{i'}}\left(G^*_{-ij}, T^*_{-ij}\right) \tag{3.7}$$

where either profitability (condition 3.5) or incentive-compatibility (condition 3.6) holds strictly.

In words, reimbursement rate t_{ij} is "safe for u_i" if insurer d_j cannot strictly profit from replacing hospital u_i with hospital $u_{i'}$ in such a way that $u_{i'}$ is willing to participate. Conditions (3.5) through (3.7) formalize this idea. The profitability condition (3.5) says that d_j prefers replacing its current contract with u_i with a contract of $t_{i'j}$ with $u_{i'}$ over the status quo. The incentive compatibility condition (3.6) says that $u_{i'}$ also prefers this move over the status quo. Finally, the no-commitment condition (3.7) says that after the substitution takes place, $u_{i'}$ would not prefer to drop its contract with d_j. Trivial and unnecessary as it may seem, condition (3.7) plays an important role in ensuring the existence of an equilibrium in which multiple insurers are competing and enrollees can spill over among them.[16]

Definition 4. Under (G, T), for a $g_{ij} = 1$, the "best safe reimbursement rate" for u_i is denoted $\hat{t}_{u_i}(G, T, d_j)$ and defined as the highest profit to u_i among all values of $\tilde{t}_{ij}$ that would be safe for u_i if charged to d_j.

Definition 4 provides us with the necessary notation for setting up the new bargaining conditions. The basic intuition is that at the equilibrium (G^*, T^*), each non-null t^*_{ij} is equal to the Nash-bargaining reimbursement rate $t^{NB}_{ij}(G, T, \gamma_{ij})$ unless $t^{NB}_{ij}(G, T, \gamma_{ij})$ is not "safe" for u_i, in which case, u_i will retreat to its best safe rate.

Definition 5. (G^*, T^*) satisfies "bargaining with threats of replacement" conditions if for $\forall g^*_{ij} = 1$ we have:

$$t^*_{ij} = \min\left(t^{NB}_{ij}\left(G, T, \gamma_{ij}\right), \hat{t}_{u_i}\left(G, T, d_j\right)\right) \tag{3.8}$$

My new solution concept for (G^*, T^*) combines pairwise stability conditions on network formation with bargaining-with-threats of replacement (rather than combining pairwise stability and Nash-in-Nash). Now let's revisit the above example to see the implications of the new concept. There is now only one equilibrium: $(G^*, T^*) = \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0.5 \\ \emptyset \end{bmatrix}\right)$. Because the insurer can now swap the hospitals, a network with only u_2 covered can no

[16]For more detail, see the online appendix.

longer be an equilibrium.[17] Also because now hospitals can anticipate and try to deter threats of replacement, u_1 does not charge unreasonably large amounts at the equilibrium just because the fine for not abiding by the mandate is very high. Rather, given that the replacement threat is binding, u_1 is only charging for how much better it is than u_2.

In the online appendix, I give more details behind my modeling choices. I discuss why I use other alternatives such as cooperative-based models, non-cooperative models such as Abreu and Manea [2012], vertical contracting models such as Segal [1999], Rey and Whinston [2013], Prat and Rustichini [2003], and bargaining models such as Stole and Zwiebel [1996a,b] among others. The rest of the paper is about the application.

4 The CommCare Market

In this section I introduce CommCare, the market to which I apply my framework in order to study the effects of network adequacy regulations. CommCare –or, more precisely, the "Commonwealth Care" market– was a subsidized health insurance exchange where people with low income (below 300% of the federal poverty line) who could not get insured by an employer or public programs were eligible to enroll. CommCare was operated by the state of Massachusetts as part of the Massachusetts Healthcare Reform (which preceded the ACA.) I chose CommCare for analysis because I believe that it is the most similar market to the ACA exchanges among those health insurance markets for which comprehensive data on claims, plan enrollments, network structures, and premiums are available. Both markets are subsidized health insurance exchanges for individuals in which private insurers compete by offering highly standardized plans.[18] Nevertheless there are some differences between the two markets.[19] In section 6, I will discuss how some of these differences might lead us

[17]Like with existence, I do not prove any uniqueness result. But so far, in all of my counterfactual analyses, my results have been unique.

[18]Other common features between CommCare and ACA exchanges include but are not limited to: (1) Mandatory participation for people who are eligible and cannot find insurance elsewhere; (2) Risk-adjustment programs to discourage insurers from competing for healthier enrollees.

[19]Some of the differences are: (1) CommCare included below poverty people but ACA exchanges leave that group for state Medicaid programs. (2) ACA exchanges also include people ineligible for subsidies but CommCare left them for another program named CommChoice, short for "Commonwealth Choice".

to expect different responses to network adequacy regulations between CommCare and the ACA exchanges.

CommCare was established in 2007 and had around 200,000 enrollees annually in fiscal years of 2011 through 2013. Enrollees were partitioned into three income-groups: below the Federal Poverty Line (FPL), between 100% and 200% of FPL, and between 200% and 300% of FPL. Each income group was subject to a different subsidy rate and would pay a different (uniform) co-pay when visiting an in-network hospital.[20] In particular, the below-FPL group was fully subsidized and paid no premium. This group also paid zero co-pay. Other groups were also heavily subsidized but paid positive premiums and co-payments. Given that the full subsidization of below-poverty enrollees rendered them price insensitive, CommCare imposed a new, auction-like, regulation from 2012 on in order to foster price competition among insurers. The regulation roughly stated that below-poverty consumers who had just joined the market were only allowed to choose between the two cheapest plans.

I examine the whole CommCare market for estimation and identification. But for the counterfactual policy I apply some restrictions. First, I focus on the 2011 fiscal year. I do this because the particular auction-like regulation for 2012 and 2013 complicates the premium-setting game.[21] Also, to save on computation, I restrict my analysis geographically by concentrating on the Greater Boston Area (which I define by all zip codes not more than 30 miles from zip code 02114 in downtown Boston). Finally, I focus on general acute care hospitals rather than all hospitals (throughout this paper, unless otherwise stated, by

(3) CommCare was more highly standardized than ACA exchanges are. In the ACA exchanges, plans can be offered in different "metal tiers" with different benefits. In CommCare, there was no deductible or co-insurance. There was only a small co-pay which was not to be decided by the insurer. It was the same for each income group across all insurers and all hospitals (MassHealth [2011]). (4) The subsidies for each income group are fixed in the ACA exchanges but were linked to premiums in CommCare, meaning in CommCare most consumers would receive more in subsidies if they bought a higher-premium plan.

[20] In fact, more precisely, the total number of income group was five rather than three: below FPL, 100%-150% of FPL, 150%-200% of FPL, 200%-250% of FPL, and 250%-300% of FPL. Unfortunately, in my data, I cannot distinguish between 100%-150% of FPL and 150%-200% of FPL or between 200%-250% of FPL and 250%-300% of FPL. Therefore, I work with three income groups.

[21] With this regulation, the Bertrand Nash model for premium-setting often leads to mixed-strategy Nash equilibria. Also, I found evidence for insurers' incomplete information about each others' pricing. For more detail on this evidence, see the online appendix.

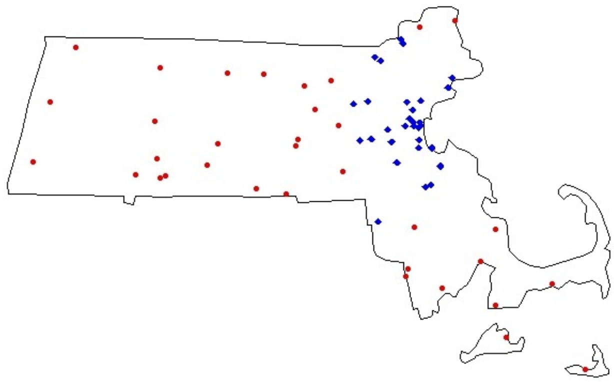

Figure 4.1: General Acute Care Hospitals in Massachusetts. Those in the Greater Boston Area are shown in blue rectangles.

"hospital" I mean a general acute care hospital.) There are 28 hospitals in the Greater Boston Area, which are owned by 16 hospital systems. Figure 4.1 locates these hospitals on a map.

In 2011 (and thereafter, until 2014 when CommCare was shut down and the ACA took over), five insurers competed in CommCare. One of them (Fallon) was inactive the Boston area. I focus, therefore, on the other four: BMC (Boston Medical Center), Celticare, NHP (Neighborhood Health Plan), and Network Health. Given that below poverty consumers in CommCare are insensitive to premiums, CommCare sets a cap on premiums in order to prevent arbitrary price increases. In 2011, the cap was $425/month per person. CommCare also had a mandated premium floor (likely to help prevent adverse selection by shifting the competition from the price domain to quality domain) which was at the level of $404/month in 2011. Table 1 summarizes the state of the market in the 2011 fiscal year. Celticare has the narrowest network among the four with only 4 systems covered out of the whole 16. It also charges the lowest premium. Nevertheless, my estimations indicate that Celticare pays, on average, a higher reimbursement to its in-network hospitals for a severity-adjusted admission than any of the other three insurers. The main reason is that Celticare covers the two most expensive hospitals in Massachusetts: Massachusetts General Hospital (MGH) and Brigham and Women's Hospital (BWH). Both MGH and BWH belong to the Partners Healthcare system ("Partners" henceforth). MGH and BWH are highly prestigious academic medical centers and attract a lot of patients in any network in which they are participate, and particularly if the network is a smaller one like Celticare's, leading to high average reimbursements as well as high average marginal cost of inpatient care. All of the other plans have broader networks, higher premiums, and lower estimated average reimbursements than Celticare.

Celticare has other major differences from the other insurers in the market. Celticare has a substantially lower market share than the rest of the plans. Also, unlike BMC, NHP, and Network Health, which are all based in Massachusetts and entered CommCare in fiscal year of 2007, Celticare is based in St. Louis, Missouri, and entered the market in 2010. In other words, Celticare is a *smaller* insurer compared to the other three. As section 5 will explain in more detail, Celticare's small enrollment size and little-known brand indeed have implications for its hospital network and premium. There is evidence from the ACA exchanges that smaller insurers have a harder time building hospital networks due in part to lack of established relationships with agents, as well as lack of sufficient number of enrollees to diversify risk (McKinsey [2015]). Celticare itself cites its "small size" as its biggest limitation to engage with providers [Health Policy Commission, 2013]. Also, according to McKinsey [2015], less well-known insurers like Celticare "may face greater pressure to be price competitive to attract members" to make up for their weak brand names. Section 5 shows how different features of my formulations for profit functions $\pi_{u_i}(G,T,P)$ and $\pi_{d_j}(G,T,P)$ capture these differences among plans (in particular Celticare's low economies of scale) as well as the implications of these differences for network formation, bargaining, and premiums.

5 Estimation

model's purpose was to predict the market outcome (G^*, T^*, P^*) given all profit functions $\pi_{u_i}(\cdot)$ and $\pi_{d_j}(\cdot)$ and bargaining parameters matrix $\gamma_{m\times n}$. Estimation is the reverse. We observe (G^*, T^*, P^*) plus some partial data about $\pi_{u_i}(\cdot)$ and $\pi_{d_j}(\cdot)$, and our objective is fully backing out estimates $\hat{\pi}_{u_i}(\cdot)$, $\hat{\pi}_{d_j}(\cdot)$, and $\hat{\gamma}_{ij}$ for all i, j so that we can then do counterfactual analysis. Of course, part of the identification comes from parametric assumptions on profit functions and bargaining powers. I start by specifying how profit functions depend on demand and cost functions. Later (in sections 5.2 through 5.5), I will further specify

	BMC	Celticare	NHP	Network Health
Systems Covered (out of 20)	11	4	15	14
Premium ($/month)	425	404	425	425
Average reimbursement ($/admission)	11,781	14,120	12,025	12,069
Average inpatient marginal cost ($/admission)	6721	10302	8335	8362
Market Share	32%	12%	27%	29%
Based in	MA	MO	MA	MA
First Fiscal Year in CommCare	2007	2010	2007	2007

Table 1: The CommCare Market in FY 2011, Greater Boston Area

the assumptions made on demand functions, cost functions, and the bargaining parameter matrix.

Hospital u_i's profit from each insurer d_j that covers u_i is how many units of healthcare per month care it provides to enrollees of d_j times the marginal profit u_i makes from each unit of care provided to d_j enrollees.[22] Different consumer groups κ may have different tastes for hospitals and insurers (I will discuss in more detail how consumers are binned into different κ groups in section 5.2). So, u_i's total profit is given by:

$$\pi_{u_i}(G,T,P) = \sum_{\kappa} \sum_{j} \left(D_j^{\kappa}(G,P) \times \sigma_{ij}^{\kappa}(G) \times (t_{ij} - c_i) \right) \tag{5.1}$$

In (5.1), c_i is the marginal cost of providing one unit of care to a patient. Function $D_j^{\kappa}(G,P)$ depicts how many enrollees from bin κ choose to enroll with insurer d_j. Function $\sigma_{ij}^{\kappa}(G)$ represents how many units of care per month the average d_j enrollee from bin κ receives from hospital u_i.[23,24]

Insurer d_j's profit is the sum of premiums it charges minus the different types of costs

[22]One "unit of healthcare" is one hospital admission with average severity. Severities are measured in DRG weights and discussed in section 5.2.

[23]Note that neither $D_j^{\kappa}(G,P)$ nor $\sigma_{ij}^{\kappa}(G)$ is a function of T. The reason is that in CommCare, there was no co-insurance. Where there is co-insurance, T needs to be included as an argument to both functions.

[24]As I will formally define soon, by construction, $\sigma_{ij}^{\kappa}(G) = 0$ for all ij such that $g_{ij} = 0$.

it incurs:

$$\pi_{d_j}(G,T,P) = \sum_{\kappa}\left(D_j^{\kappa}(G,P) \times \left(p_j - \xi_j - \sum_i \left(\sigma_{ij}^{\kappa}(G) \times t_{ij}\right)\right)\right) - \sum_{i\, s.t.\, g_{ij}=1} f_{ij} \qquad (5.2)$$

where p_j is the premium that d_j charges. The first element of cost to d_j is ξ_j, which denotes monthly "non-inpatient costs" to d_j per each enrollee. ξ_j can include costs from out-patient care by hospitals, pharmacies, non-hospital care, the variable component of administrative costs, etc. The second element is the amount that d_j pays to each in-network hospital u_i in reimbursements, which is equal to $\sigma_{ij}^{\kappa}(G) \times t_{ij}$ for an average enrollee of type κ.

The third component of costs to d_j is the set of fixed costs f_{ij} which d_j incurs for every hospital u_i that it covers. These fixed costs are an important part of what explains the differences in economies of scale across plans (the other key part is heterogeneity in plans' brand values, and is discussed in section 5.3). They explain why Celticare forms a narrow network including prestigeous hospitals like MGH and BWH but excluding a lot of cheap hospitals geographically close to them.[25] There are multiple sources for such fixed costs. First, the bargaining process itself is costly. Second, "risk-diversification" can be modeled as insurers facing fixed costs.[26]. Third, insurers sometimes leave hospitals out of the network if they fail to meet some quality standards that may not be directly observable to patients. I model such quality concerns as part of the f_{ij} costs.[27] Finally, the fourth source is insurer d_j's over- or under-estimation of how profitable signing a contract with

[25]For instance, in the absence of such fixed costs, one would not expect low-cost hospitals to be left out of hospital-networks that cover high-cost hospitals; because covering a hospital that would agree to join the network with a low price could steer patients away from more expensive hospitals, and, in the worst case of steering no one away, would do no harm to the insurer.

[26]Insurers with smaller sizes (e.g., Celticare in CommCare) often have a harder time contracting with hospitals since the expected numbers of their enrollees who would visit a hospital are less likely to be large enough to make the insurer feel more confident that those enrollees, who turn out to be more costly to the insurer, will be cancelled out by those who are less costly. Of course, the most accurate way to model this risk-diversification issue would be to (1) directly model the financial risk that insurers face regarding the healthcare costs of their enrollees, and (2) model insurers as risk-averse agents. This would substantially complicate the model. However, modeling this phenomenon as a fixed cost does not further complicate the model and, at the same time, captures the idea that a small insurer might have a harder time sign a contract with a hospital without facing risks.

[27]For instance, if a hospital has old equipment which the insurer believes might have serious side effects

Figure 5.1: The four steps of the estimation procedure

hospital u_i would be. This latter term plays the role of the structural errors that will be used in the estimation procedure to rationalize the data.

Given the specifications of $\pi_{u_i}(\cdot)$ and $\pi_{d_j}(\cdot)$ in (5.1) and (5.2), estimating the profit functions means estimating all $\hat{\sigma}^{\kappa}_{ij}(\cdot)$, $\hat{D}^{\kappa}_{j}(\cdot)$, $\hat{\xi}_j$, and $\hat{f}_{ij}$ for all i and j. The procedure that I develop for this estimation has four steps which correspond to the four steps of the sequential game (shown in figure 3.2) in reverse order. Figure 5.1 schematically represents the four steps of the estimation process. The first two steps are based heavily on methods developed by Capps et al. [2003] and Ho [2006]. In step 1, I estimate all $\hat{\sigma}^{\kappa}_{ij}(\cdot)$ using a logit model of hospital choice. In step 2, given the outcome of step 1, I estimate all $\hat{D}^{\kappa}_{j}(\cdot)$ using a model of plan choice. In step 3, given the full demand estimation outcome from steps 1 and 2, I impose the Nash Bertrand assumption on premium-setting to back out all non-inpatient costs $\hat{\xi}_j$. Finally, in step 4, I impose the network-formation and bargaining model developed in section 3 to back out all $\hat{f}_{ij}$ and $\hat{\gamma}_{ij}$. In the remainder of this section, I review the data and discuss each of the four steps of the estimation procedure in more details. For each step, after further parameterization of the object of interest if necessary, I (i) describe the estimation methodology, (ii) discuss the identification, and (iii) report and interpret the results.

for patients, the insurer may decide to leave the hospital out of network. Whether this comes from altruism towards customers or from concerns about future profits, it can be captured as a cost to the insurer for including that hospital in its network.

5.1 Data

Datasets used in my analysis consist of (1) Data on hospital discharges and medical claims, (2) data on insurance plan enrollments, and (3) data on hospital and insurance plan characteristics. Below, I discuss these different data in more details.

5.1.1 Data on Hospital Discharges and Medical Claims

I use data on hospital discharges and medical claims at two points. First, I use data on payments from insurers to hospitals in order to construct a measure of the reimbursements matrix T^*, the details of which are explained in the online appendix. Second, I use data on hospital discharges in step 1 of the estimation procedure (see figure 5.1) where I back out hospital choice functions $\hat{\sigma}^{\kappa}_{ij}(\cdot)$. The primary source of my data for discharges and medical claims is the Massachusetts All Payers Claims Database (MA-APCD) from the Center for Health Information and Analysis (CHIA). MA-APCD has a medical claims dataset which offers very rich and comprehensive information on medical claims and discharges from 2010 to 2014.

In the medical claims dataset, the unit of observation is claim-line, which pertains to an individual medical bill that a provider (e.g., physician, hospital, pharmacy) sends to an insurer. For each claim-line, MA-APCD's medical claims dataset contains information on patient demographics, diagnosis, type of claim (in particular whether or not the claim pertains to an inpatient hospital admission), date, payments, and identifiers for the provider, the patient, and the insurer. The demographic information consists of the patient's gender, age, and the 5-digit zip code of residence. The diagnosis information is reported in ICD-9-CM diagnosis codes.[28] The payments are broken down into payments by the patient and payment by the carrier (i.e., insurer). The provider ID is the National Provider Identifier (NPI) issued to providers by the Centers for Medicare and Medicaid Services (CMS). A hospital may have multiple NPIs. The patient identifier is an ID assigned by CHIA. This is an "APCD internal ID" which enables researchers to link together different claims and

[28] ICD-CM is short for International Classification of Diseases, Clinical Modification. The ICD-9-CM is an adaptation by the U.S. National Center for Health Statistics of ICD-9 diagnostic codes which are used internationally for diagnosis classification. There are about 14,000 ICD-9 codes.

enrollment records of the same patient within the APCD.[29]

I carry out several further processings on the medical claims data before I use it in the estimation procedure. I restrict the medical claims dataset to inpatient claims only. I also use data on the NPIs of general acute care hospitals in Massachusetts from CMS both to restrict the claims data to those from general hospitals in Massachusetts and to match each claim line in the data to a hospital name. Since a hospital may issue multiple bills to an insurer for different services provided during a single hospital stay, I aggregate the claims data from the "claim-line level" to the "admission-episode" level by lumping together all claims that have the same patient ID, insurer ID, hospital name (I do not use hospital NPI, since a hospital can have multiple NPIs), diagnosis, and service provision year and month. Finally, I link the ICD-9-CM diagnosis codes to a coarser category called Clinical Classifications Software (CCS) developed by the Agency of Healthcare Research and Quality (AHRQ). This linkage enables me to link diagnoses to MS-DRG severity indices for different diagnosis groups.

5.1.2 Data on Insurance Plan Enrollments

I use data on enrollments in CommCare insurance plans to back out plan demand functions $\hat{D}_j^{\kappa}(\cdot)$ from consumers' plan choice patterns in step 2 of the estimation procedure (see figure 5.1 and section 5.3). My enrollments data comes from the MA-APCD's enrollments dataset.

In the enrollments dataset, the unit of observation is enrollment record. An enrollment record is uniquely identified by its enrollee ID and fiscal year. Enrollee IDs are the same IDs as those assigned to patients in the medical claims file so the two files can be linked together. Fiscal year is a variable that I construct for each record from enrollment start year and month, as well as publicly available information on the timing of CommCare's operation. A CommCare fiscal year started on July with an open enrollment period, making, for instance, July of 2010 the first month of the 2011 fiscal year (henceforth, 2011FY). For each enrollment record, MA-APCD's enrollments dataset contains information on enrollee demographics and insurer ID. Demographic information in the enrollments dataset consists of the same demographic elements that were included in the medical claims file plus an

[29] This ID cannot be used to personally identify any patient or enrollee.

additional one: the enrollee's income group, which can be one of the following three groups: below FPL, within 100%-200% of FPL, and within 200%-300% of FPL.

5.1.3 Data on Hospital and Insurance Plan Characteristics

I supplement the MA-APCD medical claims data and enrollments data with several other datasets. I use public data from the Mass Connector on CommCare plans' hospital networks. I use this dataset in the hospital choice estimation (see step 1 in figure 5.1) to identify the choice set of each patient who visits a hospital. I also use it in the plan demand estimation (see step 2 in figure 5.1), in conjunction with the results of the hospital choice estimation, in order to construct measures of relative values of different plans' hospital networks to different consumer bins κ on CommCare. I also use data from the Mass Connector on CommCare plans' premiums as well as subsidy rates in CommCare. I use these data in the plan demand estimation process (see step 2 in figure 5.1).[30]

Finally, I use data on hospitals' costs from CMS's Healthcare Cost Report Information System (HCRIS). These data include annual costs reports by all Medicare-certified providers (which include all of the hospitals I study) to CMS, broken down into seven cost centers –such as general service, inpatient service, outpatient service, and ancillary service– and then, for each cost center, broken down in great details into many items. I then adopt the approach used in Schmitt [2015] to construct from these cost items a measure of c_i, hospital average inpatient costs per severity adjusted hospital admission. For details on how I construct these c_i measures, see the online appendix.[31]

[30] Even though data on hospital networks, plan premiums, and subsidies in CommCare was public, it was removed from the Mass Connector website before I could access it. I thank Mark Shepard for sharing these data with me.

[31] The original cost reports data from HCRIS is very raw and requires a great deal of processing before c_i measures can be constructed off of it. I thank Matt Schmitt for sharing with me a cleaned up version of the HCRIS data.

5.2 Step 1: Estimating Hospital Demand

In the first step of estimating the model, I back out the hospital choice function $\sigma_{ij}^{\kappa}(G)$. This function represents the total demand for hospital u_i from an *average consumer* who is in bin κ and has enrolled in insurance plan d_j. This is a weighted sum of demand levels $\sigma_{ij}^{\kappa,\psi}(G)$ for different diagnoses ψ, where the weight for each diagnosis is its DRG severity measure w_{ψ}. Formally:

$$\sigma_{ij}^{\kappa}(G) = \sum_{\psi \in \Psi} w_{\psi} \times \sigma_{ij}^{\kappa,\psi}(G) \tag{5.3}$$

Note that (5.3) has a direct implication for the interpretation of (5.1) and (5.2). It implies that the reimbursement made by insurer d_j to hospital u_i for a hospital admission with diagnosis ψ is equal to $w_{\psi} \times t_{ij}$. Similarly, the marginal cost to hospital u_i of providing care to an admitted patient with diagnosis ψ is implicitly assumed equal to $w_{\psi} \times c_i$. This linearity assumption is standard in the literature (see, for example, Gowrisankaran et al. [2013], Ho and Lee [2017b], Prager [2015]).

Diagnosis-specific hospital choice function $\sigma_{ij}^{\kappa,\psi}(G)$ is assumed to come from a multi-nomial logit model of hospital choice for different consumer bins κ. I bin the consumers based on two observables: the 5-digit zip code of residence location, and income group (below poverty, between 100% and 200% of poverty, and between 200% and 300% of poverty). Therefore, bin κ is a combination of location l and income group y. For each individual k, I use the notation $\kappa(k)$ (notation picked up from Ho and Lee [2017b]) to represent the consumer bin that consumer k belongs to. Notations $l(k)$ and $y(k)$ are also used in a similar manner.

Underlying my multi-nomial choice model is the following utility function for individual k with diagnosis ψ admitted to hospital u_i:

$$V_{ik\psi}^{H} = \delta d_{il(k)} + v_i \times \left(1 + \theta^{diag} \times w_{\psi}\right) + \varepsilon_{ik\psi} \tag{5.4}$$

This utility function assumes that when evaluating hospital u_i, consumer k pays attention to how far the hospital is located from where she lives as well as to the quality of the hospital. In (5.4), d_{ik} is the distance between hospital i and the residence of individual k

who belongs to bin $\kappa(k)$, and δ is the corresponding coefficient. Also v_i captures hospital i's quality as perceived by consumers. The utility function in (5.4) allows for the possibility that for more severe conditions, consumers care more about hospital quality and less about distance. Therefore, in (5.4), there is also an interaction term $v_i \times w_\psi$ is multiplied by the coefficient θ^{diag}, which measures the extent to which patients with more severe conditions (i.e., higher w_ψ) pay extra attention to hospital qualities v_i compared to patients with less severe conditions. The last term in (5.4), $\varepsilon_{ik\psi}$ is an idiosyncratic error term whose distribution is i.i.d Type 1 extreme value with variance of 1. The the distribution of $\varepsilon_{ik\psi}$ gives a closed-form representation for $\sigma_{ij}^{\kappa,\psi}(G)$. For $g_{ij} = 0$, we have $\sigma_{ij}^{\kappa,\psi}(G) = 0$. For $g_{ij} = 1$ we get:

$$\sigma_{ij}^{\kappa,\psi}(G) = \lambda_{\kappa,\psi} \frac{e^{\bar{V}^H_{i\kappa\psi}}}{\sum_{i' \ s.t. \ g_{i'j}=1} e^{\bar{V}^H_{i'\kappa\psi}}} \tag{5.5}$$

where $\bar{V}^H_{i\kappa\psi}$ (note that in the subscript, it is κ and not k) the average of $V^H_{ik\psi}$ over all consumers k of type κ. That is, $\bar{V}^H_{i\kappa\psi} = \delta d_{il} + v_i \times \left(1 + \theta^{diag} \times w_\psi\right)$, where l is the location element of κ. Also $\lambda_{\kappa,\psi}$ is the expected rate of hospital admission with diagnosis ψ per month per member for consumer bin κ.

Estimation Procedure: The parameters to estimate in this step are $\left(\hat{\delta}, \hat{v}_i, \hat{\theta}^{diag}, \hat{\lambda}_{\kappa,\psi}\right)$. I observe $\hat{\lambda}_{\kappa,\psi}$ directly from the data. To estimate the rest of the parameters, I estimate a multi-nomial logit model. I observe hospital choices $\sigma_{ij}^{\kappa,\psi}(G)$ for all i, j, κ, ψ in my data.[32] I estimate the parameters using a maximum likelihood approach, matching the observed $\sigma_{ij}^{\kappa,\psi}(G)$.

Identification: Hospital qualities v_i are identified by cross-hospital variation in discharge volumes. If hospital u_i has a higher discharge share than hospital $u_{i'}$ among consumer bins κ that live in locations equally far from u_i and $u_{i'}$, then v_i must be larger than $v_{i'}$. Distance coefficient δ is identified by within-hospital, cross-location variation in discharge volumes. The faster the discharge share of hospital u_i diminishes as we look at consumer bins κ farther from u_i, the more negative the distance coefficient δ must have been. The severity weight coefficient θ^{diag} is identified by the variation in discharge volumes within consumer bins κ but across hospitals and diagnoses. The more the discharge shares get

[32] I construct $\sigma_{ij}^{\kappa,\psi}(G)$ for each i, j, κ, ψ using the discharge data for choices made by patients and data on the network G for the choice set.

skewed towards higher quality hospitals as we look at more severe diagnoses ψ, the higher the implied θ^{diag} would be.

Results and Interpretation: Table 2 summarizes the results of the multi-nomial logit estimation of the hospital choice model. The distance coefficient is estimated to be $\hat{\delta} = -0.137$, which is consistent with the literature. It implies that an extra 10 miles of distance reduces the share of a hospital by an average of 29%. The DRG weight coefficient is estimated to be $\hat{\theta}^{diag} = 0.036$. To illustrate the interpretation of this number, suppose hospitals u_i and $u_{i'}$ are such that $\frac{\hat{v}_i - \hat{v}_{i'}}{\hat{\delta} \times \left(1+\hat{\theta}^{diag}\right)} = 10$. That is, the average patient with diagnosis ψ of severity $w_\psi = 1$ would choose $u_{i'}$ over u_i if $u_{i'}$ is no more than 10 miles farther from her than u_i. Then, a DRG coefficient of $\hat{\theta}^{diag} = 0.036$ would imply that the average patient with diagnosis ψ' of severity $w_{\psi'} = 2$ would be willing to travel $10 \times \frac{1+2\times 0.036}{1+0.036} \simeq 10.36$ more miles to visit hospital $u_{i'}$ than she would to visit u_i.

Hospital Choice Estimates

	coeff.	std. error
distance (miles)	-0.136***	(0.002)
DRG weight	0.036**	(0.017)
Hospital FEs	Yes	
Num. hospital admissions	40,247	
pseudo R^2	0.413	
std errors in parentheses, *: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$		

Table 2: Hospital Choice Model

5.3 Step 2: Estimating Insurance Plan Demand

Having estimated a model of hospital choice, I now turn to estimating the insurance plan demand functions $\hat{D}_j^\kappa(\cdot)$. I assume that consumer k's valuation of insurance plan j comes from the following utility function:

$$V_{jk}^I = \alpha \times EU_{j\kappa(k)} + \beta_{y(k)} p_{jk}^{sub} + \Delta_j + \epsilon_{jk} \tag{5.6}$$

where $EU_{j\kappa(k)}$ is the expected utility of insurer d_j's network of providers for consumer k

of type κ, and α represents how much consumers care about network expected utility when they compare plans. Network-of-providers utility EU_{jk} is given by:

$$EU_{j\kappa} = \frac{\sum_{\psi} \lambda_{\kappa,\psi} EU^{\psi}_{j\kappa}}{\sum_{\psi} \lambda_{\kappa,\psi}} \tag{5.7}$$

which means that $EU_{j\kappa}$ comes from a weighted average of $EU^{\psi}_{j\kappa}$, which are values of plan j's network of providers for different diagnoses ψ for a consumer of type κ. The weights come from the likelihoods of different conditions for type κ. Given the distribution of $\varepsilon_{ik\psi}$, there is a closed-form representation for how the values $\bar{V}^{H}_{i\kappa\psi}$ of hospitals covered by d_j contribute to EU^{ψ}_{jk}:

$$EU^{\psi}_{j\kappa} = \ln\left(\sum_{i\; s.t.\; g_{ij}=1} e^{\bar{V}^{H}_{i\kappa\psi}}\right) \tag{5.8}$$

Δ_j in (5.6) is insurer d_j's fixed effect, which can be interpreted as "brand effect". $\beta_{y(k)}$ is the sensitivity of consumer k, who belongs to income group $y(k)$, to the monthly premium p^{sub}_{jk} she has to pay for plan j. Due to subsidization in CommCare, there is a difference between the premium p_j each insurer charges for its plan and p^{sub}_{jk}, the one a consumer buying that plan pays. The relationship between p^{sub}_{jk} and p_j is linear. There is a fixed subsidy b_y and a pass-through rate a_y for each income group forming a linear relationship, as shown in (5.9).[33] All a_y and b_y are observed.[34]

$$p^{sub}_{jk} = a_{y(k)} \times p_j - b_{y(k)} \tag{5.9}$$

Finally, in (5.6), ϵ_{jk} is an i.i.d extreme value type I error term, which gives us the following closed-form solution for $D^{\kappa}_{j}(G,P)$:

$$D^{\kappa}_{j}(G,P) = \Lambda_{\kappa} \frac{e^{\bar{V}^{I}_{j\kappa}}}{\sum_{j'} e^{\bar{V}^{I}_{j'\kappa}}} \tag{5.10}$$

[33]There is one non-linearity in CommCare's subsidization policy. For the second poorest income group (i.e., between 100% and 150% of poverty), b_y is determined in such a way that there is at least one free option for income group y. That is, b_y should equate $\min_j p^{sub}_{jk}$ to zero for that group. In my counterfactual analysis, I do not incorporate this effect. Instead, I take all a_y and b_y remains constant in response to counterfactual scenarios. This simplifies the model greatly and does not substantially change the results.

[34]In 2011, an X% of the premiums charged by insurers in CommCare was subsidized through fixed subsidies (i.e., a_y) and another Y% through variable subsidies (i.e., b_y). Consumers paid the remaining 100-X-Y%,

where (similar to the construction of $\bar{V}^H_{i\kappa\psi}$) $\bar{V}^I_{j\kappa}$ is constructed by averaging V^I_{jk} over all consumers k of type κ (which gives the same formula as V^I_{jk} but without the error term ϵ_{jk}). Also, Λ_κ is the population of individual consumers of type κ.

Estimation Procedure: The parameters to estimate in this step are coefficients on network utility and monthly premium, as well as the brand effects: $\left(\hat{\alpha}, \hat{\beta}_y, \hat{\Delta}_j\right)$. Populations Λ_κ are observed, and network utilities $EU_{j\kappa}$ can be constructed from the results of the hospital choice estimation (and are hence, treated as data in this step). To back out $\left(\hat{\alpha}, \hat{\beta}_y, \hat{\Delta}_j\right)$, I estimate the multi-nomial logit model set up in equations (5.6) through (5.10) using a maximum likelihood approach.

Identification: Brand fixed effects Δ_j are identified by the variation in the enrollment volumes across plans. Network utility coefficient α is identified by within-plan, cross-location variation in enrollment volumes. To illustrate, if the ratio of Celticare enrollments to NHP enrollments for a certain income group in a certain year is constant across residents of different zip codes, then all of the difference is picked up by brand effects. But if this enrollments ratio is lower in locations where NHP's hospital network includes some nearby hospitals, but Celticare's does not, then this variation in ratios is explained by a non-zero α. Price coefficients β_y for different income groups are identified by within-plan, over time variation in subsidy rates and plan premiums. The traditional price-endogeneity issue is less of a concern in CommCare. Individuals with incomes below the federal poverty line accounted for more than 40% of the total CommCare enrollees every year. Health insurance for this group was always fully subsidized in CommCare. Therefore, the effective price sensitivity for below-poverty consumers is zero, and the underlying β_y for this group is not a coefficient that I estimate. But due to its large size and due to the particular institutional details of regulations regarding this income group on CommCare, I believe there is enough evidence that most (if not all) of the variation in insurer premiums in my data stems from insurers' strategies regarding this group. Therefore, there should be little concern about premium variation being endogenous to demand from *the other income groups*, for which I estimate premium coefficients β_y. For more detail on this, see the online appendix.

Results and Interpretation: Table 3 summarizes the results of the multi-nomial logit estimation of the plan demand model. An average CommCare enrollee with an income between 100% and 200% of the federal poverty line is estimated to be almost twice as sensitive to the after-subsidy premium paid as an average enrollee with an income between 200% and 300% of the federal poverty line. Brand effect estimates indicate that except for Celticare, the brand values for all the CommCare plans are close to one another. An average CommCare enrollee with an income between 100% and 300% of poverty would be willing to pay about $30/month less (in after-subsidy premium) for Celticare than she would for BMC, if the two plans offered the same network of hospitals. Similarly, the average above-poverty CommCare enrollee would value the NHP brand and Network Health brand, respectively, at $8/month and $3/month below that of BMC. Note that this is exactly where the model captures the idea that Celticare is a smaller insurer than the others on CommCare. A smaller brand fixed effect Δ_j implies that a plan would get a smaller market share even if it offered the same hospital network as the other insurers. This, as will be discussed later, hinders Celticare from adding hospitals (especially smaller ones) to its network since the number of new enrollees that those hospitals would bring Celticare is too few to justify the corresponding fixed cost.

Finally, the network utility coefficient α is estimated at 0.76. This magnitude for α implies that, an average CommCare enrollee with an income between 100% and 300% of poverty would be willing to pay almost 21$ more in after-subsidy premium for Celticare, if Celticare improved its network of hospitals to one similar to NHPs.[35]

5.4 Step 3: Estimating Insurers' Non-Inpatient Costs

Variable costs to insurers have two main components. The first component is inpatient costs, which are the reimbursements t_{ij} that the insurer pays to its in-network hospitals for

[35] A last note on my plan demand model is that it does not account for potential inertia in plan choice by consumers. Even though it would be reasonable to believe that such inertia does exist, I think abstracting away from the complexities that it would add to my model does not drastically change the results. This is because in the plan demand estimation, I use data only from fiscal years of 2012 and 2013. And when I use these estimates (which do not account for inertia) to project market shares for the fiscal year of 2011, my predictions are close to the observed shares. Given that substantial changes happened to some plans between 2011 and 2012 as well as between 2012 and 2013, I believe that the model doing well on predicting 2011 market shares suggests that abstracting from inertia is not a far from reasonable.

Plan Demand Estimates

	coeff.	std. error
network expected utility	0.76***	(0.14)
price coefficients		
0-100 % poverty (omitted, no premiums)	—	—
100-200 % poverty	-0.0413***	(0.0028)
200-300 % poverty	-0.0198***	(0.0021)
brand fixed effects		
BMC (omitted category)	0	—
Celticare	-1.137***	(0.041)
Fallon	0.012***	(0.004)
NHP	-0.296***	(0.002)
Network Health	-0.117**	(0.05)
Num. enrollments	381901	
pseudo R^2	0.704	
std errors in parentheses, *: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$		

Table 3: Plan Demand Model

the care that they provide to its enrollees. The second component is the non-inpatient cost ξ_j in (5.2). In the online appendix, I detail how I measure reimbursement rates t^*_{ij} using the medical claims file of the MA-APCD. In the main text, I treat them as data. In this section, I back out the non-inpatient costs ξ_j.

Estimation Procedure: The parameters to back out in this step are insurers' non-inpatient costs $\hat{\xi}_j$ for FY2011. This is done by imposing the Nash Bertrand equilibrium assumption on the premium-setting. That is, no insurer d_j should be able to do better by changing its premium:

$$\pi_{d_j}\left(G^*, T^*, P^*\right) \geq \pi_{d_j}\left(G^*, T^*, \left(p_j, p^*_{-j}\right)\right) \tag{5.11}$$

Insurer d_j's non-hospital cost per member per month is estimated by finding the ξ_j to make this optimality condition hold.

Identification: Unfortunately, the identification is only partial in this step of the

	BMC	Celticare	NHP	Network Health
Bound on ξ_j (\$/month/person)	284	275	281	276
Type of Bound	Lower	Upper	Lower	Lower

Table 4: Upper or lower bounds on non-inpatient costs ξ_j

estimation process. In FY2011, the before-subsidy premiums bid by all CommCare insurers were enforced by regulation to be no more than \$425/month and no less than \$404/month. Thus, all that the identifying assumption (i.e., equation (5.11)) implies is that for every insurer d_j, the observed premium p_j^* in the data would do weakly better than any other $p_j \in [404, 425]$. Given that all insurers' premiums were exactly at either of the two extremes of this continuum, equation (5.11) only gives us bounds on ξ_j values rather than point estimates.

Results and Interpretation: Table 4 presents the bounds backed out on non-inpatient costs ξ_j using (5.11).

Calibration: Given that I only have bounds on ξ_j, I need to calibrate the non-inpatient costs. I calibrate the vector $\hat{\xi}$ exactly at the boundaries of the inequalities. That is, $\hat{\xi} = (284, 275, 281, 276)$. I believe this is a reasonable calibration for multiple reasons. First, the fact that Celticare covers, on average, more expensive hospitals than the other three plans (see table 1) can also push Celticare's outpatient cost further upward compared to others. Therefore, given that, by table 4, the upper bound on $\xi_{Celticare}$ is below the lower bound on ξ_j for the other plans, calibration at the boundary seems not far from reasonable. Also, even though non-inpatient costs ξ_j may have varied over time, looking at the pricing decisions of insurers in fiscal years other than 2011 can still be suggestive regarding how reasonable the decision to calibrate $\hat{\xi}$ at the boundary is. In fiscal year 2010, the premium floor and cap were \$381/month and \$391/month respectively. The fact that BMC, NHP, and Network Health could all profitably charge no more than \$391/month (i.e., \$34/month below their prices in FY2011) suggests that calibrating their ξ_j at the lower bounds is a good working assumption. In addition, during fiscal years of 2012 and 2013 when CommCare had set the premium floor at the much lower levels of \$360/month and \$354/month respectively, and put in place a procurement process to encourage insurers to price lower, Celticare was not always among the two cheapest plan on CommCare.[36] This suggests that Celticare must not have had a ξ_j much lower than the other plans, which is consistent with my calibration.

5.5 Step 4:Estimating Insurer Fixed-Costs and Bargaining Parameters

Modifying the Model: Before I perform the estimation in this step, I make modifications to the model to avoid computational burden for the estimation of fixed costs and bargaining parameters and, more importantly, for the simulation of counterfactual scenarios. First, instead of the whole state of Massachusetts, I concentrate on the Greater Boston Area which I define by zip codes that are at most 30 miles far away from 02114, a zip code in downtown Boston. There are 28 hospitals in this area. Also, Fallon was not active in CommCare in this area. So, I am left with BMC, Celticare, NHP, and Network Health. Second, I coarsen the set of consumer bins κ by aggregating the zip codes into 38 different locations. I also aggregate all DRG weights into a single one.[37] Third, I simplify the way hospital-insurer pairs u_i and d_j anticipate how premiums respond to the outcome of their bargaining over reimbursement rate t_{ij}. I assume that when negotiating, the pair take as given the premiums set by *all insurers other than* d_j. That is, the pair assume that only d_j will adjust its premium optimally after the outcome of the bargaining between the two is determined. This assumption substantially simplifies the computation as it will not involve computing a complete Nash Bertrand equilibrium for insurer premium-setting for every candidate negotiated t_{ij}. In spite of the massive computational advantage, this assumption does capture the first-order effect of premium response to bargaining outcomes. It only leaves out how other insurers respond to the premium change by d_j.[38] This type of simplifying assumption has been made in the literature to ease the computation of models of bargaining on two sided markets. In fact, Ho and Lee [2017b] assume that when bargaining, hospital u_i and insurer d_j take as given *all of the premiums, including* p_j. The assumption in Ho and Lee [2017b] simplifies the computation much further than the assumption I impose.

[36] For more detail on CommCare procurement rules for fiscal years of 2012 and 2013 and the insurers' premiums during those years, see the online appendix.

[37] This aggregation abstracts away from heterogeneity in the severity of different admission events. I decided for this aggregation since trying it on the early versions of the model did little to change the results.

[38] Note that this assumption still retains the feature that at equilibium market outcome (G^*, T^*, P^*), premiums P^* are a Nash Bertrand equilibium among insurers, given G^* and T^*.

Nevertheless, as I will discuss in the online appendix, it would not be suitable to a model of bargaining with endogenous network formation such as the model in this paper, as it would lead to wide ranges of multiple equilibria in many cases.

The fourth modification to the model not only simplifies the computation, but is also rooted in the institutional details of the market. I assume hospitals do not bargain individually, but as systems. The participation of hospitals owned by the same hospital system across plans' hospital networks is highly correlated with one another with full correlation in most cases. Therefore, I assume that it is the whole system, denoted s, that negotiates with each insurer d_j over rates and that the outcome of the negotiation is the participation of either all of the system's hospitals u_i in d_j's network or none of them. This partitions the 28 hospitals in the Greater Boston Area into 16 systems[39]. I also assume that in the case of full participation, the prices negotiated between insurer d_j and all hospitals u_i in system s are fully correlated. That is, I construct a fixed "base-rate matrix" $T^{base} = \left[t_{ij}^{base}\right]_{m \times n}$ and assume that the negotiated prices between insurer d_j and all hospitals u_i in system s can only take the form of $t_{ij} = t_{ij}^{base} + z_s$.[40] In other words, I assume that the negotiation between the insurer and the hospital system is in fact over z_s. This assumption brings the dimensionality of each bargaining process from the size of system s to 1, without substantially affecting the outcomes.[41] After implementing all these four modifications, I turn to the estimation of fixed costs and bargaining parameters.

Estimation Procedure: The parameters to estimate in this step are insurers' system-specific fixed costs of inclusion $\hat{f}_{sj}$ and bargaining parameters $\hat{\gamma}_{sj}$. Note that given the

[39]If, within an actual system (i.e., joint ownership of hospitals), the network participation of some hospitals do not fully correlate with one another, I consider them separate systems. For instance, Faulkner Hospital is owned by the Partners system, which means that it is in the same system as MGH. Nevertheless, I consider them to be in separate systems, as they do not always appear in the same hospital networks.

[40]For every ij link that does appear in the network structure based on my FY2011 data, I set t_{ij}^{base} to the value of the estimated reimbursement rate $\hat{t}_{ij}^{*}$ corresponding to that link, which, as mentioned before, I measure in the online appendix. For all other links ij, I set t_{ij}^{base} to hospital u_i's marginal cost c_i.

[41]To see why this assumption does not have large effects on the outcome, observe that increasing some t_{ij} and decreasing some $t_{i'j}$ would have little net effect on either the insurer profit or the hospital system profit (which is the sum of the profits to the individual hospitals forming that system). Therefore, the bargaining here could be thought of as the insurer's attempt to get discounts from *all* of the individual hospitals in the system, and the system's attempt to charge the insurer more for the service of *all* of the hospitals.

aforementioned aggregation to the hospital-system level, index s for systems has replaced index i for hospitals. I combine the estimation results of the previous steps with network stability conditions in definition 1 and bargaining conditions in definition 5 to back out the set of $\hat{f}_{sj}$ and $\hat{\gamma}_{sj}$ that rationalizes the observed market outcome (G^*, T^*, P^*) in the data as an equilibrium of the sequential game schematically represented in figure 3.2, section 3.

Note that by only imposing network-stability conditions and bargaining conditions, we will not obtain point estimates for all $\hat{f}_{sj}$. For instance, bargaining conditions do not give any information on $\hat{f}_{sj}$ for any system-insurer pair who are not contracting with each other (i.e., when $g^*_{ij} = 0$ for all $i \in s$). For such pairs, network stability conditions also only give a lower bound on $\hat{f}_{sj}$. To obtain point estimates, I impose the assumption that fixed costs f_{sj} depend linearly on observable characteristics of hospital systems and insurance plans. This assumption allows for those f_{sj}s that are point identified from the structural assumptions to also carry some information about those for which the structure of the model only predicts bounds. Given that the number of data points is limited (there are a total of 64 f_{sj} values for 16 hospital systems and 4 insurers in the Greater Boston area), for hospital system characteristics, I only include total bed size, denoted χ_s, and for insurer characteristics, I only include insurer fixed effects. Formally:

$$f_{sj} = \mathit{\Omega} \times \chi_s + FE_j + \upsilon_{sj} \tag{5.12}$$

The procedure through which the linear regression model in (5.12) is combined with the structural equations of the model to estimate fixed costs and bargaining parameters is adopted from the estimation procedure in Gowrisankaran et al. [2013]. My estimation procedure solves a constrained optimization problem. It finds the $\left(\hat{f}_{sj}, \hat{\mathit{\Omega}}, \hat{FE}_j, \hat{\gamma}_{sj}\right)$ that minimize the variance of the residuals υ_{sj} as given by equation (5.12) (i.e., the moment condition), subject to two types of constraints on the profit functions: (i) (G^*, T^*, P^*) satisfies the inequalities imposed by the network stability conditions in definition 1, and (ii) (G^*, T^*, P^*) satisfies bargaining with threats of replacement in definition 5 relative to bargaining parameter matrix γ. Given the results of the previous steps, and given that hospital marginal costs c_i are observed, profit functions can now be fully computed up to the f_{sj} fixed costs and γ_{sj} bargaining parameters that are to be estimated in this step. This turns constraints (i) and (ii) on the profit functions into constraints on f_{sj} and γ_{sj}, making the optimization problem well-defined. For the formal characterization of the constrained optimization problem, see the online appendix.

Given that a high insurer fixed cost f_{sj} and a low hospital-system bargaining parameter γ_{sj} are expected to affect the bargaining outcomes in similar ways, I conservatively impose strong restrictions on bargaining parameters to ensure identification. I divide all of the hospital systems into two categories: "star" and "non-star." I consider two hospital systems to be stars. First, Partners, which includes Massachusetts General Hospital and Brigham and Women's Hospital. The second one is Tufts Medical Center. These systems are the highest ranked in Massachusetts according to U.S. News and World Report's 2016-2017 rankings of best hospitals in Massachusetts.[42] Another reason for considering these two systems to be star systems was that they seemed to have charged abnormally high to all insurers according to my measurement of T^*. I then assume that bargaining parameter γ_{sj} equals γ_{star} for all j when s is one of the two star systems, and it equals $\gamma_{nonStar}$ otherwise. Therefore, the problem of estimating $\left(\hat{f}_{sj}, \hat{\Omega}, \hat{FE}_j, \hat{\gamma}_{sj}\right)$ boils down to estimating $\left(\hat{f}_{sj}, \hat{\Omega}, \hat{FE}_j, \hat{\gamma}_{star}, \hat{\gamma}_{nonStar}\right)$.

In the online appendix I give more details on why I chose this estimation procedure based on Gowrisankaran et al. [2013] over other alternatives such as MLE or moment inequalities.

Identification: Identification of all of the parameters come jointly from the structural assumptions of the model described in section 3, as well as from the moment condition in (5.12). Fixed effects $\hat{FE}_j$ are roughly identified by average differences among different insurers d_j in the bounds on their respective f_{sj} by the structural assumptions. For instance, the imposition of no gains from trade by definition 1 on links sj with $g^*_{sj} = 0$ is expected to imply a lower bound $\underline{f}_{sj}$ on f_{sj}. If such implied lower bounds for insurer d_j tend to be larger than those implied for insurer $d_{j'}$, then FE_j can be larger than $FE_{j'}$. Also given the structure of the model, the bargaining formulation (3.8) implies an upper bound $\bar{f}_{sj}$ on f_{sj}.

[42] In that ranking, Massachusetts General Hospital, Brigham and Women's Hospital, and Tufts Medical Center are ranked first, second, and fourth respectively. Beth Israel Deaconess Medical Center was ranked third, above Tufts Medical Center. But I decided to not consider its parent system (i.e., Care Group) a star system since it also included many much lower ranked hospitals such as Beth Isreael Deaconess Needham and Beth Isreael Deaconess Milton, and especially Mount Auburn Hospital.

A similar argument applies when $\bar{f}_{sj}$ tends to be larger for d_j than for $d_{j'}$.[43] The intuition for the identification of $\hat{\Omega}$ is also similar, though the variation this time is within insurer but across hospital systems.

The bargaining parameter for star hospital systems γ_{star} is identified by searching for the γ_{star} value that makes the predicted fixed costs f_{sj} for the Partners and Tufts systems as close as possible to what a linear model of fixed costs with $\hat{\Omega}, \hat{FE}_j$ would predict. For instance, if the fixed effects $\hat{FE}_j$ and size effect $\hat{\Omega}$ are all positive, but with $\gamma_{star} = \frac{1}{2}$ the implied fixed costs for all or most sj with $s \in \{Partners, Tufts\}$ and $g_{sj} = 1$ are negative, then we expect that the "true" γ_{star} must have been larger than $\frac{1}{2}$. That is, the insurers must have accepted higher reimbursements from those star hospital systems, not because of very low fixed costs of inclusion, but because those systems have higher bargaining parameters. Bargaining parameters $\gamma_{nonStar}$ are also identified based on a similar logic.

A particular feature of my model that further helps to separately identify fixed-costs of inclusion from bargaining parameters is endogenous network formation. Some of the equilibrium conditions of the model only restrict fixed costs f_{sj} and *are invariant to bargaining parameters* γ_{sj}. For instance, lower bounds $\underline{f}_{sj}$ for those sj with $g^*_{sj} = 0$ are completely independent of the whole bargaining-parameters matrix. So, very small values of γ_{star} and $\gamma_{nonStar}$, which would imply very low and sometimes negative upper-bounds $\bar{f}_{sj}$ that are far below $\underline{f}_{sj}$, will not do well in minimizing the error in regression equation (5.12), and hence are not chosen by the estimation algorithm.[44]

Results and Interpretation: Table 5 summarizes the results. The two hospital systems I consider star are estimated to have full bargaining power $\hat{\gamma}_{star} = 1$ when negotiating with insurers.[45] I estimate that non-star systems have a bargaining power of

[43] Also, given that networks of different plans can be different, sometimes $\bar{f}_{sj}$ can be compared to $\underline{f}_{sj'}$. Similar arguments apply here.

[44] The analysis of these lower bounds from the network formation structure was also the reason why I chose how to "split" the effects between the bargaining model and the fixed costs model. In the early versions of the model, I also tried specifications where the fixed costs model included a star dummy but no size effect and the bargaining model had a size effect. That approach led, on average, to a larger diversion in $\hat{f}_{sj} - \hat{FE}_j$ for a fixed j across s than my chosen specification does, which means it does worse on matching the moments.

[45] An estimated bargaining parameter of $\hat{\gamma}_{star} = 1$ might seem too high since it seems to give all the

$\hat{\gamma}_{nonStar} = 0.72$. The fixed costs of inclusion f_{sj} are estimated to be increasing with the total bed size of the hospital system covered. The slope of this increasing relationship is estimated at $\hat{\Omega} =$ \$223K/year/100 beds. This magnitude, along with the fixed effects $\hat{FE}_j$ reported in the table, implies that such fixed costs sum up to as large as about a third of insurer profits (or, put differently, a fourth of insurer profits before accounting for the fixed costs themselves). Observe that the fixed costs estimated for Celticare are not higher than those for the other insurers. So, the rationzliation of Celticare's narrwer network than the other plan is not taking place by simply assigning higher fixed costs f_{sj} to Celticare. As mentioned before, an important role here is played by Celticare's low brand fixed effect Δ_j in the estimated plan demand model.[46]

A final observation is that the magnitude of the estimated fixed costs seem rather high. One reason for this can be that the Nash Bertrand assumption on the premiums tends to imply overly high insurer markups, thereby underestimating the non-inpatient costs. Such underestimation would overstate the value of new enrollees and, hence, the value of including an additional hospital system, to insurers. This makes higher fixed costs necessary to rationalize insurers' choice of network.[47]

surplus from the contract to the hospital. Two notes about this issue are worth making. First, note that all the surplus does not necessarily go to the hospital since the insurer is also allowed to make threats of replacement. Second, and more importantly, standard plan demand models in the literature (which I also use in this paper) do not capture consumer risk aversion as a source of bargaining leverage for star hospitals. Therefore, this leverage in my estimation process is picked up by an extremely high estimated bargaining parameter $\hat{\gamma}_{Star}$. See the online appendix for more detail on this.

[46] In fact, I ran a counterfactual simulation using the estimated parameters of the model, except that I set all of the brand effects Δ_j to zero to asume away brand heterogeneity. In the simulated equilibrium, Celticare and Network Health ended up having 10 hospital systems (out of the total 16) covered in their networks.

[47] The issue of Nash Bertrand conditions overstating markups has been mentioned in the literature before. Ho and Lee [2017b] address it by assuming that insurers perceive demand to be more elastic than it actually is. The estimates in this paper can be thought of as another alternative: Insurers failing to perceive the full relevance of their non-inpatient costs when setting premiums but do take them into account when forming networks, or equivalently, insurers implicitly taking part of their variable costs to be fixed. I, hence, do not expect the magnidute of such fixed costs to drastically affect the results that I get in my counterfactual analysis as long as they help to capture the heterogeneity in profitabilities of different hospitals to the insurer that are not otherwise captured, since these differential profitabilities are key to both which hospitals are included in the network and how strong they can be against each other as threats of replacement.

Fixed Costs and Bargaining Parameters

	coeff.	std. error
insurer fixed effects (K\$/year)		
BMC	1146***	(180)
Celticare	788***	(177)
NHP	415**	(181)
Network Health	636***	(177)
effect of size (K\$/year/100 beds)		
Ω	223***	(91)
bargaining parameters		
γ_{star}	1	
$\gamma_{nonStar}$	0.72	
std errors in parentheses, *: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$		

Table 5: Estimation of fixed costs of inclusion and bargaining parameters

6 Counterfactual Analysis

In this section, I use the estimated model to run counterfactual simulations that would help us better understand the effects of Network Adequacy Regulations on the functioning of the CommCare market. In section 6.1, I discuss the main economic forces that I believe underlie how the market responds to network adequacy regulations. In section 6.2, I present and interpret my simulation results on the effects of a range of network adequacy regulations on CommCare. In section 6.3, I discuss some ways in which one might expect the response of the ACA exchanges to such regulations to be different from CommCare's.

6.1 Main Economic Forces

The hospital network of an insurance plan can affect the reimbursements it pays to its in-network hospitals, as well as the premium it charges, through multiple channels. Among these channels, two are of special importance when it comes to understanding the effects of network adequacy regulations. The first channel has to do with the features of the in-network hospitals of each insurance plan. Hospitals with higher costs c_i and/or higher quality v_i tend to charge more for their services. So, a hospital network consisting mainly of such hospitals is expected to be more expensive to the insurer. The second channel is has more to do with how in-network hospitals of an insurance plan compare to its out-of-network hospitals. If the out of network hospitals are very close, in terms of cost and quality, to the in-network ones, the insurer can use them as a strong bargaining chip and make credible threats of replacement to its in-network hospitals in order to keep reimbursement rates low.

A network adequacy regulation can affect reimbursements through both of these channels. On the one hand, a network adequacy regulation forces an insurer with a narrow hospital network to add hospitals. Depending on where the joining hospitals stand in terms of cost c_i and quality v_i, their addition to the network can bring the average reimbursement rate –and hence the premium– up or down. On the other hand, a network adequacy regulation also *weakens* the insurer's "out-of-network pool" of hospitals. As the insurer is forced to cover more hospitals, it first appends to its hospital network those currently out-of-network hospitals that are the most profitable for it, leaving out of network the least profitable ones. Therefore, the out-of-network pool becomes a weaker replacement threat tool for the insurer when it is negotiating rates with in-network hospitals.

The extent to which network adequacy regulations weaken the ability of an insurer to make strong threats of replacement depends on how close to one another hospitals are in terms of their profitability to the insurer. To illustrate, if higher-quality hospitals (i.e., those with higher v_i, which can contribute more to network expected utility $EU_{j\kappa}$ of the insurer) tend to also be the hospitals with higher marginal costs c_j and/or fixed costs of inclusion f_{sj}, then one would expect the hospitals to be fairly similarly profitable to the insurer. Hence, once the best out-of-network hospitals go in-network due to network adequacy regulation, the remaining out-of-network hospitals may still provide almost as strong threats of replacement as the insurer could make before the regulation. However, if, for instance, higher quality-hospitals are also the cheaper ones, the insurer's ability to make threats of replacement should weaken more rapidly as tighter network adequacy regulations are imposed.

Before I turn to examining the effects of the regulation on CommCare through simulation, I would like to emphasize that the two channels highlighed in this subsection only provide a basis for thinking about the effects of the regulation. The intuition behind some of the simulation results can, in some cases, be much more complex due to the numerous moving parts of the model. An important instance of this will be discussed in section 6.2.1, where I explain in detail the large reactions of Celticare's premium and reimbursement rates to stringent network adequacy regulations. The ability to point us to such complex mechanisms in counterfactual simulations is one of the main advantages of a structural analysis of network-formation, bargaining, and premium setting in this market.

6.2 Effects of Network Adequacy Regulations on CommCare

I now simulate the effects of a range of network adequacy regulations using my estimated model on CommCare for the Greater Boston Area and FY2011. I examine network adequacy regulations in the form of mandating all insurers in the market to cover at least $X\%$ of the hospital systems in the Greater Boston Area, for a range of values of X.[48] To model this mandate, I just include an additional term in the insurer-profit-function in (5.2), representing a fine that the insurer has to pay if it falls below the $X\%$ requirement:

$$\pi_{d_j}^{NA}(G,T,P) = \pi_{d_j}(G,T,P) - \eta \mathit{\Gamma}_j(G,X) \tag{6.1}$$

where $\pi_{d_j}(G,T,P)$ is insurer d_j's profit function from equation (5.2), and $\Gamma_j(G,X)$ is the number of hospital systems by which insurer d_j is short of fulfilling the $X\%$ requirement (e.g., $\Gamma_{Celticare}(G^*, 0.5) = 16 \times 0.5 - 4$ because Celticare covers 4 hospital systems and is,

[48] Other types of network adequacy regulations to consider could be based on hospital counts or bed counts rather than hospital-system counts. Examining those regulations can be done using my model, but it would require first modifying the bargaining formula in (3.8) to capture threats of replacing multiple in-network hospital systems by multiple out-of-network ones. For instance, if an in-network hospital system u_s has five hospitals and no out-of-network system is as large, the insurer needs to be able to threat u_s to replace it, for instance, by two other systems $u_{s'}$ and $u_{s''}$ consisting of two and three hospitals respectively. I believe the main difference between the hospital-system based regulation in my model and, say, a total-beds-based one would be that a system based regulation will give more bargaining leverage to hospital systems with smaller total bed sizes (and, thereby, less leverage to systems with more beds) compared to a bed-based regulation, since a small system counts just as much as a larger system in terms of helping the insurer abide by the minimum mandated network size.

hence, 4 hospital systems short of 8 which would be 50% of the total 16 systems). Also η is the amount of fine an insurer has to pay per each of those missing hospital systems.[49] The exact amount of the fine η does not play a role in determining the result of the counterfactual simulation.[50] It only has to be large enough to ensure that no insurer would prefer going below $X\%$ to staying weakly above it. I then simulate the new equilibrium of the model again, using the parameters estimated in section 5 and using (6.1) instead of (5.2) for insurer profit.

Covering only four hospital system out of the total 16, Celticare has by far the narrowest network among the four plans in CommCare. Thus, it is the only plan directly affected by the regulation for most values of X. I hence, describe the simulated response of Celticare to network adequacy regulations, and I then turn to other insurers.

As a last point before detailing the simulation results, I would like to reiterate the main point made in section 3.2. As I argued in that section, using the NiN formulation for the bargaining component of the model would result in extremely high and unrealistic predicted increase in the reimbursement rates in response to network adequacy regulations. That is, the predicted reimbursement rates by NiN in response to regulation would not just be empirically biased. They would, in fact, be logically incoherent. In other words, part of the question that my counterfactual analysis asks is "if an insurer who has expanded its hospital network is no longer allowed to make a threat of dropping when negotiating with hospitals, to what extent can threats of replacement keep rates from surging?" If we use NiN which only allows for threats of dropping, it is straightforward that we will not get realistic predictions about situations where this threat cannot be used and insurers attempt to keep prices down in other ways. Therefore, in what follows, I analyze the counterfactual simulations only for the bargaining model with threats of replacements. I *do not* do the

[49] A more natural way of modeling the fine would be having a constant fine for being below 50% no matter how much below. With that formulation, however, my model would not capture Celticare's incentive to expand its network of hospitals. The reason is that the deviations allowed in my model are dropping, adding, or replacing one hospital system. No single such deviation can take Celticare from covering 4 systems to 8 systems. Therefore, if the fines for any degree of non-compliance with the requirement is the same, my model will not predict a change in Celticare's behavior.

[50] If the bargaining model was NiN instead of bargaining with threats of replacement, then η would play a significant role, with $\eta \to \infty$ implying $t^*_{ij} \to \infty$, which I do not find reasonable.

analysis for NiN and compare results.[51]

6.2.1 Effects on Celticare

Figures 6.1 through 6.3 depict the simulated effects on Celticare of a range of network adequacy regulations with $X = 18.75\%$ (i.e., must cover at least 3 out of the 16 hospital systems) through $X = 93.75\%$ (i.e., must cover at least 15 systems). Absent any regulation, Celticare covers four systems, which amount to 25% of those in the Greater Boston Area.

As figures 6.1 and 6.2 show, even though a regulation of $X = 25\%$ does *not* force Celticare to expand its hospital network, it does raise Celticare's average reimbursement rate (see the vertical dashed lines in the two figures). This is because this regulation prohibits Celticare from going below $X = 25\%$ by dropping any of its currently in-network hospital systems without replacement. Hospital systems then exploit this prohibition by raising their reimbursement rates exactly to their respective best safe reimbursements according to definition 4. Celticare, in turn, responds by raising its premium.

For $X = 25\%$ through $X = 81.25\%$, the average reimbursement rate declines gradually but fairly consistently. This is because the hospitals that Celticare brings in-network tend to have lower marginal costs c_i on average than the ones already in-network, which

[51] Liebman [2017] and Ho and Lee [2017a] do discuss the consequences of network adequacy regulations under NiN. That is because their approach to modeling such regulation is different from this paper which uses the negative profit shock Γ. I believe their approach can be thought of as a model of regulations that mandate to *negotiate with* a certain number of hospitals (with no restriction on whether the negotiation should lead to inclusion), rather than mandating to *include* them. To illustrate, those papers (which are based on the two-step approach with commitment explained in section 2) model a regulation that mandates a full network by assuming, in the first step, that when the insurer announces a subset of hospitals to negotiate with, that subset has to be the full list of hospitals. The second step (i.e., bargaining) proceeds as in the no-regulation case. This modeling approach abstracts away from the idea that hospitals get higher bargaining positions when the insurer is constrained by the regulation; because, in this approach, the insurer can still credibly threaten to drop hospitals in the second step of the game (i.e., the bargaining stage), even though the regulation mandates a full network. Additionally, this approach also assumes the regulation to be non-binding. For instance, if an insurer does not have gains from trade with 50% of the hospital systems but is required to negotiate with 70% of them, it will negotiate with 70%, but will still exclude some of them. In my approach, however, the penalty for going below 70% effectively creates gains from trade to ensure at least 70% of the systems are included.

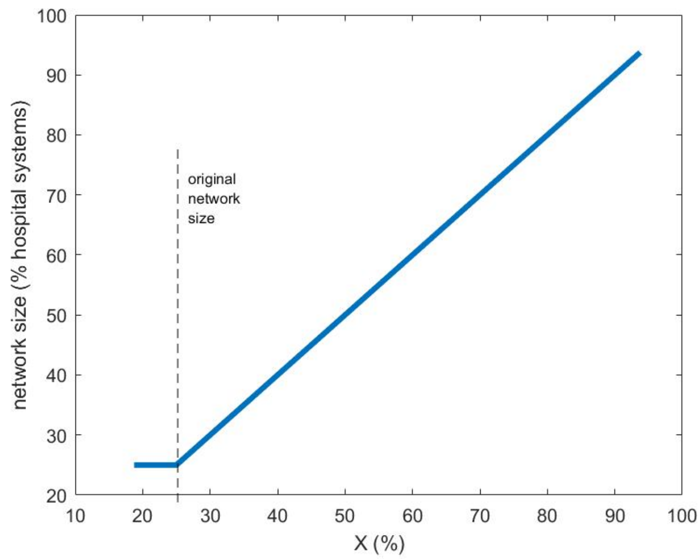

Figure 6.1: Effects of network adequacy regulations on Celticare's network size

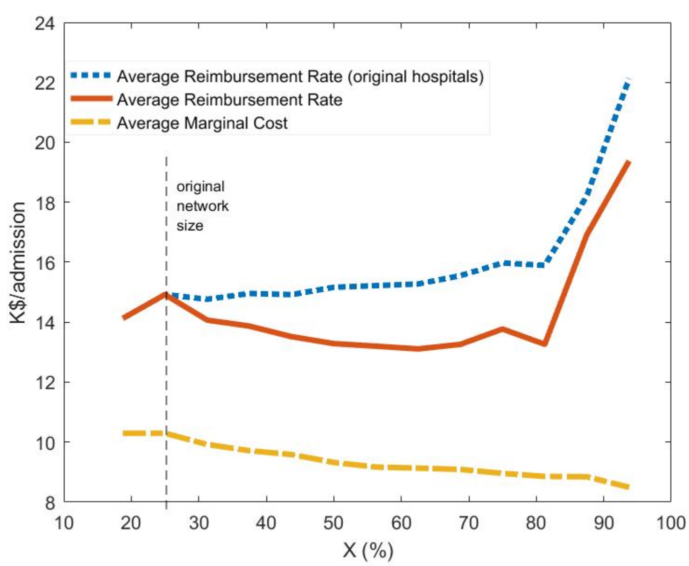

Figure 6.2: Effects of network adequacy regulations on Celticare's payments to in-network hospitals

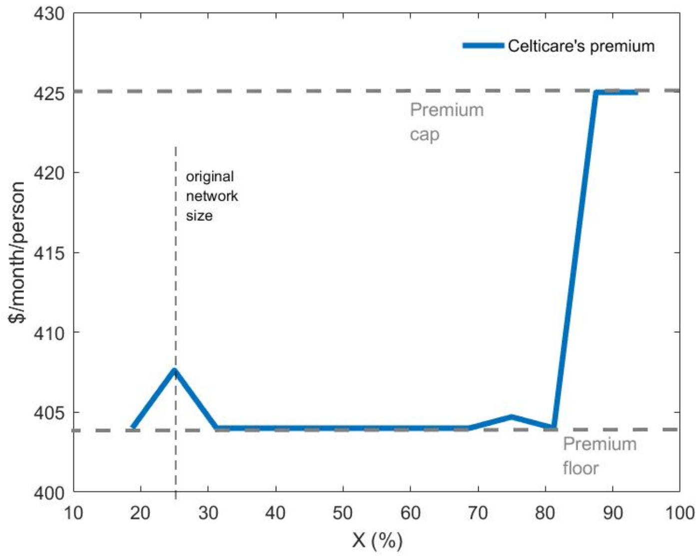

Figure 6.3: Effects of network adequacy regulations on Celticare's premium

means they can charge less. These hospitals also steer some patients away from Celticare's most expensive hospitals (in particular Brigham and Women's Hospital and Massachusetts General Hospital), further lowering the average reimbursement rate paid by Celticare. The orange dashed curve in figure 6.2 depicts how the *average marginal inpatient cost* of Celticare's hospital network changes with X, pushing Celticare's reimbursements downward.[52] This effect is partially offset by the second channel through which the regulation affects reimbursement rates: the weakening threats of replacement. As the blue dotted curve in figure 6.2 shows, the average reimbursement rate charged to Celticare by only those hospitals that were in Celticare's network before the regulation does indeed increase steadily. Nevertheless, this latter effect is not strong enough to dominate the effect of lowered average marginal costs. Hence, the average reimbursement goes down. The reason why the effect of the decreasing average marginal cost dominates that of the weakening threats of replacement is that hospital systems are indeed close enough to one another in terms of their profitability for Celticare. So, as Celticare adds the best out-of-network hospital system to its network in response to tighter regulations, the second best systems are good enough

[52] Average marginal inpatient cost is calculated by taking a weighted average of marginal costs c_i of Celticare's in-network hospitals, where the weights come from predicted discharge-volume shares of those hospitals for Celticare enrollees, based on my estimated demand model.

alternatives for Celticare to use as threats of replacement and contain the reimbursements. Over the $X = 25\%$ through $X = 81.25\%$ range of regulations, Celticare's premium is \$404/month/person, which is the lowest premium allowed by CommCare to charge. Absent this premium floor, Celticare's premium would be a little below \$404/month/person due to reduced average reimbursement.

From $X = 81.25\%$ up, the increasing trend in average reimbursement rates and the premium accelerates. Underlying this acceleration is an important self-reinforcing loop through which premiums and reimbursement rates both increase rapidly. The loop works in the following way: Once reimbursement rates increase due to depletion of Celticare's out-of-network pool, Celticare responds by raising its premium. The increased premium lowers Celticare's market share. This decreases the inpatient volumes that Celticare's in-network hospital systems get through Celticare, roughly proportionally to Celticare's market share decrease. Now, note that the rate that each system charges to Celticare under a binding network adequacy regulation is the highest safe reimbursement rate. That is, each in-network hospital system charges Celticare for the differential profitability between that system and Celticare's best replacement option. If an in-network hospital system has a lower *total* volume through Celticare (due to lower Celticare market share), that system will get to charge a larger rate paid *per admission* in order to cover the difference in profitabilities.[53] Therefore, an increase in Celticare's premium can feed back into increased reimbursement rates. But higher reimbursement rates in turn will lead Celticare to further raise its premium, which again further increases the rates, and so on. Another reason for the large price increase towards the end of the spectrum of the regulations is that with high X, other plans are also bound by the regulation, which means their in-network hospitals charge higher to them. Thus, it is more likely that Celticare's out-of-network hospital systems are now in-network for another plan affected by the regulation. This raises the reimbursement rates that such hospitals would be willing to accept in order to replace a hospital system currently in Celticare's network. Therefore, Celticare will have weaker

[53]One might think that the differences in the profitabilities of hospital systems for Celticare might also decrease proportinally to Celticare's market share. This is *not* true, since for Celticare itself, the decrease in market share due to increased premium is (at least partly) compensated for by the higher margin from the increased premium.

threats of replacement, which leads to increased reimbursement rates and premium.

6.2.2 Effects on other insurers:

The other three insurers in the market were BMC, NHP, and Network Health with hospital networks that included 11, 15, and 14 hospital systems, respectively. With X levels such that the regulation is binding for Celticare only, and the other three are only responding to Celticare's response to the regulation, their responses are not very large. The most notable observation for such X levels is that each of the other three insurers drops a hospital system or two out of their hospital network. The reason is, for X levels high enough to substantially increase the charges by some hospitals to Celticare, but low enough to not directly affect the other insurers, those hospitals that charge high to Celticare are more likely to drop out of the other insurers' networks. This arises from the fact that those hospitals now have higher incentive to try to steer their more loyal patients from other insurers to Celticare. In order to prevent these hospital systems from leaving, other insurers must also raise their reimbursements which is not always optimal, especially when the hospital system threatening to leave the network was already only marginally profitable to the insurer. Along the way, NHP drops the North Shore Health System and Emerson Hospital, and Network Health drops Cambridge Health Alliance.[54]

When the regulation becomes binding, BMC, NHP, and Network Health have similar reactions to that of Celticare. The average reimbursement increases abruptly for any insurer d_j once X hits the network size of d_j. Also, with higher levels of X, the average reimbursement also increases, as the out-of-network pool of hospitals weakens.

6.2.3 Summary of Consumer Welfare Implications

As the analysis in this section shows, lower levels of X can have good consumer-welfare effects. They expand Celticare's hospital network. Also, since Celticare adds hospitals with

[54]Network Health drops Cambridge Health Alliance after X passes the level that makes the regulation binding for BMC. So, it could be that dropping Cambridge Health Alliance is Network Health's response to BMC's response to the regulation, and not to Celticare's.

lower marginal costs c_i, they bring Celticare's average reimbursement down.[55] So, they do not lead to increased premium, thereby making Celticare a better plan for consumers. with mid-range values of X, however, the shrinkage of the other networks starts to happen which hurts consumers. With high values of X, Celticare's premium rapidly rises but the overall effect of the regulation is still positive for consumers, according to my welfare analysis. This is because all of the plans expands their networks and, besides Celticare, none of them can raise its premium in response to the higher reimbursement rates, due to the premium cap of \$425/month set by CommCare.

Given that there are substantial differences between CommCare and the ACA exchanges, they might have different responses to network adequacy regulations, potentially with different welfare implications. Thus, I do not discuss the consumer-welfare analysis of CommCare in further detail here. Instead, I allocate the remaining space to the comparing CommCare to the ACA exchanges on consumer welfare-relevant features.[56]

6.3 Potential Differences between CommCare and the ACA Exchanges

There are at least two major differences between CommCare and the ACA exchanges that could trigger different responses to regulation. The first difference is in subsidy structures and the second is in cost-sharing features of the plans.

6.3.1 Subsidy Structures

In CommCare, the amount p_{jk}^{sub}, which consumer k pays out of pocket for insurer d_j's premium, is a subsidization of p_j with a fixed and variable component, as given by (5.9). The value of the pass-through rate a_y was zero for the below-poverty income group. This renders a large portion of CommCare consumers effectively price insensitive, thereby considerably reducing the price sensitivity of an average CommCare consumer. ACA exchanges on the

[55]Note that those low-cost hospitals that Celticare appends to its network in response to regulations usually have high fixed costs f_{sj} and that is why Celticare did not already cover them in its network without regulation. Nevertheless, the f_{sj} costs are fixed and will not get passed on to consumers.

[56]Of course the fate of the ACA itself is actively in progress. As of now, the House of Representative has passed measure to repeal and replace the ACA. The matter will now be in the hands of the Senate.

other hand, follow a fixed subsidy policy. Each ACA exchange, enrollee pays the premium charged by the plan she's enrolled in minus the premium charged by the second cheapest "silver plan" offered in the enrollee's local exchange market.[57,58]

This difference in subsidy schemes can have implications for consumer-welfare impacts of network adequacy regulations. Lower effective premium sensitivity in CommCare may have led to high reimbursement rates in CommCare, both in the form of insurers covering hospitals with higher c_i and insurers paying a higher t_{ij} to each hospital u_i than they would have, had the subsidy scheme included only a fixed component. But in the ACA exchanges, plans with narrower hospital networks tend to cover, on average, hospitals that have lower c_i (Dafny et al. [2015b]).[59] Thus, in the ACA exchanges, including a new hospital in response to network adequacy regulation is not as likely to bring down the average cost. Therefore, the effect of the regulation, through the average marginal cost channel, on reimbursement rates, and hence on premiums, could be an increasing effect now, unlike what was the case with Celticare in CommCare. This difference, on the one hand, implies potentially steeper price increases in ACA exchanges, compared to CommCare, in response to the regulation. But on the other hand, given that hospitals with higher c_i tend to also be hospitals with higher quality v_i, a network adequacy regulation can improve the ACA exchanges' hospital networks much more than it does in CommCare. This leaves an overall consumer welfare comparison between the effects on CommCare and the exchanges somewhat ambiguous.

6.3.2 Cost-sharing Features

In CommCare, the only cost that patients paid out of pocket for hospital admission was a uniform co-pay, independent of the hospital chosen and the service received. This was mandated by regulation to be based only on the consumer's income group. In the ACA

[57] Bronze, silver, gold, and platinum are the four different categories, or "metal tiers" of the ACA exchange plans based on cost-sharing parameters. For a typical silver plan, the consumer pays, in expectation, 30% of her healthcare costs out of pocket and the rest is borne by the insurer. Corresponding numbers for bronze, gold, and platinum are, respectively, 40%, 20%, and 10%. CommCare's cost sharing parameters are more standardized and did not include this kind of tiering.

[58] If the price calculated by taking this difference is negative, then the consumer will pay zero.

[59] This is contrary to CommCare, in which Celticare, the plan with the narrowest hospital network, covers, on the average, more expensive hospitals than all other plans.

exchanges, there are four “metal tiers” of bronze, silver, gold, and platinum. For the bronze tier, the average consumer is expected to pay 40% of her healthcare costs out of pocket in the form of co-pays, co-insurance, and deductible. These numbers are 30%, 20%, and 10%, respectively for silver, gold, and platinum plans.

This difference in cost-sharing features may also lead to different responses by the two markets to network adequacy regulations, for at least two reasons. First, unlike CommCare, the cost-sharing structure in the ACA exchanges makes patients sensitive to reimbursement rates t_{ij}. This creates another control mechanism for reimbursement rates. In the ACA exchanges, even under network adequacy regulations by which insurers have very weak out-of-network pool to use as threats of replacement, in-network hospitals might still want to charge reasonable prices in order to better compete with other in-network hospitals for patients who do care about their out-of-pocket expenses when choosing which hospital to visit.

The second difference that cost-sharing parameters can make for the ACA exchanges is the network-expected-utility coefficient α introduced and estimated at 0.76 in section 5. With higher out of pocket costs in the ACA exchanges for in-network hospital use, it is possible that consumers place less value on how many and which hospitals a plan has in its network.[60] A lower α implies that a hospital's quality v_i matters less to insurers when evaluating the hospital for its profitability, and that cost parameters c_i and f_{sj} now matter relatively more. This effect can render the out-of-network hospitals in the exchanges, which seem to be on the high-cost side, weaker threats of replacement for the in-network ones, potentially amplifying the reimbursement rate increase in response to network adequacy regulations. Of course, when hospital qualities matter less to insurers, it is also possible that insurers spend more uniformly distributed resources on bargaining with different hospitals. This could potentially lower the heterogeneity in fixed costs f_{sj} of inclusion, likely mitigating the amplifying effect of a lower α on the response of the market to network adequacy

[60] This might be one reason why in Ericson and Starc [2014], the network-expected-utility coefficient is estimated at 0.375 for CommChoice, another market from the Massachusetts Healthcare Reform with cost-sharing tiering similar to the ACA exchanges. This value is almost half as large as what I estimate for CommCare. Given that similar normalizations are used in my paper and by Ericson and Starc [2014], the estimates are directly comparable.

regulations.

7 Conclusion and Future Research

In this paper, I developed a model of insurer-provider markets with three important features. First, it endogenously captures the formation of hospital networks, bargaining between hospitals and insurers over reimbursement rates, and premium setting by insurers. Second, in formulating the bargaining process, my model improves upon a standard model called Nash-in-Nash by allowing for the possibility that when bargaining with hospitals over rates, insurers not only threaten to drop hospitals from the network, but also to replace them with currently out-of-network hospitals. This helps capture the Bertrand-type competition among hospitals for network inclusion and the fact that insurers can play hospitals off against each other and get lower prices. The third feature is computational tractability. My model can be estimated and used for counterfactual simulations on relatively large markets. In this paper, I also develop an estimation procedure for the model and apply it to the CommCare market in the fiscal year of 2011 in the Greater Boston Area with 16 hospital systems and four insurers.

I use my model to study by how much network adequacy regulations in CommCare can undermine insurers' bargaining position against hospitals and, hence, lead to increased reimbursement rates as well as increased premiums. I find that milder regulations lead to only moderate increases and sometimes even reductions in reimbursement rates and premiums on CommCare. However, enforcing insurers to have almost-full networks can lead to drastic price hikes since it depletes the out-of-network hospital pools that insurers use as threats of replacement against in-network hospitals to keep rates down. I point out some differences between the ACA exchanges and CommCare that might cause ACA exchanges respond differently to such regulations than CommCare does.

This framework developed in this paper can applied beyond network adequacy regulations. An important example is merger analysis. One of the anti-competitive effects of a merger between two hospitals is that an insurer can no longer use them as threats of replacement against each other. Unlike my model, NiN abstracts away from this effect and, hence, under-predicts the magnitude of the anti-competitive effects. More generally, this framework can be applied beyong the insurer-provider market. It can be empirically applied to other two-sided markets (such as TV channels and cable companies, or manufacturers and retailers) to answer questions centered around how network formation and bargaining take place and affect each other.

References

Dilip Abreu and Mihai Manea. Bargaining and efficiency in networks. *Journal of Economic Theory*, 147(1):43–70, 2012.

Ken Binmore, Ariel Rubinstein, and Asher Wolinsky. The nash bargaining solution in economic modelling. *The RAND Journal of Economics*, pages 176–188, 1986.

Cory Capps, David Dranove, and Mark Satterthwaite. Competition and market power in option demand markets. *RAND Journal of Economics*, pages 737–763, 2003.

Allan Collard-Wexler, Gautam Gowrisankaran, and Robin S Lee. Bargaining in bilateral oligopoly: An alternating offers representation of the" nash-in-nash" solution. Technical report, National Bureau of Economic Research, 2014.

Health Policy Commission. Public hearings health care cost trends in the commonwealth: Celticare health plan of massachusetts, inc. (naic 13632), 2013. URL http://www.mass.gov/anf/budget-taxes-and-procurement/oversight-agencies/health-policy-commission/annual-cost-trends-hearing/2014/testimony/2014-pre-filed-testimony-celticare-health-exhibits-and-questions.pdf.

Gregory S Crawford and Ali Yurukoglu. The welfare effects of bundling in multichannel television markets. *American Economic Review*, 102:643–85, 2012.

Gregory S Crawford, Robin S Lee, Michael D Whinston, and Ali Yurukoglu. The welfare effects of vertical integration in multichannel television markets. Technical report, National Bureau of Economic Research, 2015.

Leemore Dafny, Jonathan Gruber, and Christopher Ody. More insurers lower premiums: Evidence from initial pricing in the health insurance marketplaces. *American Journal of Health Economics*, 2015a.

Leemore Dafny, Igal Hendel, and Nathan Wilson. Narrow networks on the health insurance exchanges: What do they look like and how do they affect pricing? a case study of texas. *The American Economic Review*, 105(5):110–114, 2015b.

David Dranove, Christopher Ody, and Mark Satterthwaite. Bargaining foresight in hospital markets. 2016.

Keith Marzilli Ericson and Amanda Starc. Measuring consumer valuation of limited provider networks. Technical report, National Bureau of Economic Research, 2014.

Gautam Gowrisankaran, Aviv Nevo, and Robert Town. Mergers when prices are negotiated: Evidence from the hospital industry. Technical report, National Bureau of Economic Research, 2013.

Jonathan Gruber and Robin McKnight. Controlling health care costs through limited network insurance plans: Evidence from massachusetts state employees. Technical report, National Bureau of Economic Research, 2014.

Martin B Hackmann, Jonathan T Kolstad, and Amanda E Kowalski. Adverse selection and an individual mandate: When theory meets practice. *The American economic review*, 105(3):1030–1066, 2015.

John William Hatfield, Scott Duke Kominers, Alexandru Nichifor, Michael Ostrovsky, and Alexander Westkamp. Stability and competitive equilibrium in trading networks. *Journal of Political Economy*, 121(5):966–1005, 2013.

Kate Ho and Robin S Lee. Equilibrium provider networks: Bargaining and exclusion in health care markets. 2017a.

Kate Ho and Robin S Lee. Insurer competition in health care markets. Technical Report 2, 2017b.

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

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

Henrick Horn and Asher Wolinsky. Bilateral monopolies and incentives for merger. *The RAND Journal of Economics*, pages 408–419, 1988.

Matthew O Jackson. Allocation rules for network games. *Games and Economic Behavior*, 51(1):128–154, 2005.

Matthew O Jackson and Asher Wolinsky. A strategic model of social and economic networks. *Journal of economic theory*, 71(1):44–74, 1996.

Robin S Lee and Kyna Fong. Markov-perfect network formation an applied framework for bilateral oligopoly and bargaining in buyer-seller networks. 2013.

Eli Liebman. Bargaining in markets with exclusion: An analysis of health insurance networks, 2017.

Eric Maskin and Jean Tirole. A theory of dynamic oligopoly, ii: Price competition, kinked demand curves, and edgeworth cycles. *Econometrica: Journal of the Econometric Society*, pages 571–599, 1988.

MassHealth. Report to the massachusetts legislature: Implementation of health care reform, 2011. URL https://betterhealthconnector.com/wp-content/uploads/annual-reports/ConnectorAnnualReport2011.pdf.

McKinsey. Winning strategies for participating in narrow-network exchange offerings. 2013. URL http://healthcare.mckinsey.com/winning-strategies-participating-narrow-network-exchange-offerings.

McKinsey. The emerging story on new entrants to the individual health insurance exchanges, 2015. URL http://healthcare.mckinsey.com/emerging-story-new-entrants.

Ariel Pakes. Alternative models for moment inequalities. *Econometrica*, 78(6):1783–1822, 2010.

Elena Prager. Tiered hospital networks, health care demand, and prices. 2015.

Andrea Prat and Aldo Rustichini. Games played through agents. *Econometrica*, 71(4):989–1026, 2003.

Patrick Rey and Michael D Whinston. Does retailer power lead to exclusion? *The RAND Journal of Economics*, 44(1):75–81, 2013.

Matt Schmitt. Do hospital mergers reduce costs? *Available at SSRN 2599000*, 2015.

Ilya Segal. Contracting with externalities. *Quarterly Journal of Economics*, pages 337–388, 1999.

Mark Shepard. Hospital network competition and adverse selection: Evidence from the massachusetts health insurance exchange, 2015.

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

Lars A Stole and Jeffrey Zwiebel. Organizational design and technology choice under intrafirm bargaining. *The American Economic Review*, pages 195–222, 1996b.

Pietro Tebaldi. Estimating equilibrium in health insurance exchanges: Price competition and subsidy design under the aca. *Available at SSRN 2688894*, 2015.

Appendices

A Non-cooperative Foundation for a class of 2×1 games

In this section, I provide a non-cooperative foundation for the solution concept introduced in model section of the main text. Note that the non-cooperative foundation provided here is for a class of 2×1 games (i.e. 2 upstream firms and 1 downstream firm) and are hence not as general as the application in the main text. Nevertheless, studying these 2×1 games is helpful. The reason is it demonstrates that when, on the one hand, we do not have a fully general non-cooperative foundation, but on the other hand, we do need to modify the Nash-in-Nash formulation to capture the effects of substitution threats, the formulation developed in this paper is a reasonable choice.[61] This appendix is organized as follows. First, I introduce the general class of 2×1 whose Subgame Perfect Nash Equilibria I intend to solve for. Second, I propose a strategy profile for the game. Finally, I show that my proposed strategy profile is in fact an SPNE of the game and that the network and reimbursements predicted by this SPNE matches what the model suggested in the main text of the paper would predict. Note that the proofs in this appendix are mostly at the technical level. For a more intuitive argument instead, skip the proofs and see the discussion after theorem 2.

A.1 A Class of 2×1 Network-Formation-and-Bargaining Games

The game that I set up to analyze the problem of network formation and bargaining involves two upstream firms u_1 and u_2 and a downstream firm d. This game will have infinitely many periods and different components of the payoffs to the firms can realize in different periods. In particular, similarly to Collard-Wexler et al. [2014], I make a distinction between two components of firm payoffs. The first one is the "pre-transfer" payoff to each firm, which is

[61] Note that even the Nash-in-Nash concept (for modeling bargaining *on an exogenously given network*) did not have a non-cooperative foundation until very recently when Collard-Wexler et al. [2014] developed one. Nevertheless it was widely used in applied work. Moreover, many of the assumptions made in the non-cooperative foundation provided by Collard-Wexler et al. [2014] are not made in the applied papers that use Nash-in-Nash.

../../../../Desktop/Yale/Talks/2017_10_26_FTC/Paper/ClassOf2x1Games.jpg

Figure A.1: Pre-transfer payoffs to the three firms as functions of the network

the payoff each firm makes if no reimbursement is made by d to either u_1 or u_2. Denote this by $\pi_{u_i}(G)$ and $\pi_d(G)$. The second component is, of course, the transfers t_i of money from d to u_1 and u_2. Note that, unlike the main text, here I am assuming that the reimbursements are in the form of lump-sum transfers. Restrictive as this assumption may be, it significantly reduces the complexity of the proof. Also the same assumption was made by Collard-Wexler et al. [2014] in the development of the non-cooperative model for Nash-in-Nash even though the price-per-unit could be thought of as the working assumption in the empirical literature using Nash-in-Nash. Therefore, for the rest of this appendix, I will call reimbursements t_i "lump-sum transfers" or simply "transfers."

Figure A.1 depicts the pre-transfer payoffs to the two upstream firms u_1 and u_2 and the one downstream firm d under different networks and reimbursements. When the network is empty, each firm makes a profit of 0. When d is connected to u_1 only, the pre-transfer profit to d is equal to $\pi_d\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \pi_1 > 0$ and that to u_1 is $\pi_{u_1}\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = 0$ (i.e. u_1 has a zero cost of production). Also because u_2 is not doing any trade, we have $\pi_{u_2}\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = 0$. When d is connected only to u_2, the pre-transfer profits to u_2 and d are equal to $\pi_{u_2}\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = 0$ and $\pi_d\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = \pi_2$ respectively, where $\pi_2 > 0$ and $\pi_2 < \pi_1$ (This latter assumption implies that, all else equal, d would rather have u_1 in its network than u_2). Finally, when the network is full, then the profits to u_1, u_2, and d are 0, 0, and π_3 respectively, where $\pi_3 < \pi_2$ (This latter assumption absent any transfers, then d would strictly prefer having either of u_1 and u_2 in its network to having both).

I now set up a non-cooperative game of network formation and bargaining among u_1, u_2, and d. This game is a natural extension of the classic bilateral bargaining game in Binmore et al. [1986] to 2 upstream firms and one downstream, and has a similar structure to the game in Collard-Wexler et al. [2014]. The extensive form of the game is as follows:

The game is a dynamic one with infinitely many periods $\tau = 0, 1, 2, 3, ...$ The time distance between every two consecutive periods is $\Delta > 0$ (not to be confused with the same notation in the estimation section of the main text. None of the estimation notations of the main text are used in this appendix). The "discount factors" that are used to discount firms' payoffs are δ_{u_1}, δ_{u_2}, and δ_d respectively for u_1, u_2, and d. These discount factors are given by: $\delta_{u_i} = e^{-r_{u_i}\Delta}$ and $\delta_d = e^{-r_d\Delta}$. I assume $r_{u_i}, r_d > 0$ which implies that δ_{u_i} and δ_d are all within the interval $(0, 1)$.

In period 0, no link has been "formed" yet. Every period has two stages. In the first stage of even periods, all of the upstream firms u_i who haven't yet formed a link with the downstream d simultaneously make offers of lump-sum transfers $t_i \in \mathbb{R}$. Then, in the second stage of the even periods, the downstream firm considers the offers made in the first stage (if any), and decides which one(s) to accept if any. Once the downstream accepts an offer t_i a corresponding link is "formed" between d and u_i. This link stays formed for the rest of the game. Every odd period also has two stages. In the first stage, downstream firm d makes offers t_i to all of the upstream firms with which it hasn't formed a link yet. In the second stage, all of those upstream firms, to whom d made offer(s) in the first stage, consider the offer(s) t_i and simultaneously decide whether or not to accept. For any offer t_i accepted by a u_i from d, a corresponding link will be formed for the rest of the game.

The payoff to each firm is the discounted sum of the payoffs the firm makes in all periods τ. The payoff to downstream firm d realized at the end of period τ has two components. First, d receives a flow payoff of $(1 - \delta_d)\pi_d(G_\tau)$ where G_τ is the network structure realized at the end of period τ and $\pi_d(G_\tau)$ is the pre-transfer payoff to d as specified in figure A.1. The idea behind assuming that the fellow payoff has a $1 - \delta_d$ term is that if the network structure is G throughout the whole game, the net present value of the pre-transfer payoff that it gives to d will be given by:

$$(1 - \delta_d)\pi_d(G) \times \sum_{\tau=1}^{\infty} \delta^\tau = \frac{(1 - \delta_d)\pi_d(G)}{1 - \delta_d} = \pi_d(G)$$

Therefore, the net present value of the pre-transfer payoff is exactly $\pi_d(G)$.[62] The

[62]This idea is adopted from Collard-Wexler et al. [2014].

pre-transfer components of the payoffs to u_1 and u_2 at every period τ are also formulated similarly, which, given figure A.1, means that this component is always zero for u_1 and u_2.

The second component of the payoff realized to each firm at every period τ is the transfer. Unlike the pre-transfer payoffs (which are "direct" payoffs from the network structure), I do not assume that payoffs from transfers are realized gradually and in perpetual cash flows. Rather, I assume, similarly to Collard-Wexler et al. [2014], that if at some period τ, an offer of t_i is made by u_i and accepted by d (or made by d and accepted by u_i), then u_i immediately realizes a flow payoff of t_i at period τ and never after. Similarly, d realizes an immediate flow payoff of $-t_i$ at period τ and never after from this transaction. Of course, from a net present value perspective, these one-time flow payoffs are evaluated at $\delta_{u_i}^{\tau} t_i$ and $-\delta_d^{\tau} t_i$ for u_i and d respectively. Therefore, for instance if in period $\tau = 0$ the downstream accepts an offer of t_1 from u_1 and never deals with u_2 throughout the game, the net present values of the payoffs to u_1, u_2, and d will be t_1,0, and $\pi_1 - t_1$ respectively.

A.2 A Proposed SPNE of the Game

In this section I specify a strategy profile as a candidate SPNE of the game laid out in section A.1. In next sections of this appendix, I prove that is in fact a SPNE and that the payoff it predicts to be realized on the equilibrium path are in line with the formulations used in the main text of this paper to analyze the problem of network formation and bargaining in the health insurance industry.

Denote:

$$\bar{t}_1 = \min\left(\pi_1 - \pi_2, \pi_1 \frac{1 - \delta_d}{1 - \delta_d \delta_{u_1}}\right) \tag{A.1}$$

$$\tilde{t}_1 = \delta_{u_1} \bar{t}_1 \tag{A.2}$$

My proposed suggests the following actions at different types of subgames of the game in section A.1:

I- Subgames S where u_1 and d have already formed a link: These subgames S are all subgames starting at the first or second stage of an even or odd period τ where there is already a link between u_1 and d, which by the specification of the rules of the game, we know cannot be broken. This, then, boils down to a bilateral Rubinstein bargaining game with no gains from trade since $\pi_3 < \pi_1$. Therefore, Any offer that u_2 would be willing to make would be rejected by d and vice versa. Thus, for such subgames, instead of fully specifying a set of strategies, we just observe that all equilibria of such subgames are offers being made and rejected, which means that we expect no transfers and no change in the network structure once we enter one of such subgames.

II- Subgames S where u_2 and d have already formed a link: These subgames behave in a similar way to subgames of type **I** since $\pi_3 < \pi_2$.

III- Subgames S where u_1 and u_2 simultaneously make offers to d: These subgames S are all subgames starting at the first stage of an even period τ where no link between d and any u_i has already been formed yet. Proposed strategies by u_1 and u_2 at all such S are:

$$t^*_{u_1}(S) = \bar{t}_1 \tag{A.3}$$

$$t^*_{u_2}(S) = 0 \tag{A.4}$$

The intuition behind these formulations for $t^*_{u_1}(S)$ and $t^*_{u_2}(S)$ is that u_2 tries its best (by offering 0) to undercut u_1; and u_1 offers a transfer that is consistent with Rubinstein bargaining *unless there is a danger of being undercut by u_2,* in which case u_1 offers the highest price that would just prevent u_2 from undercutting u_1 (this idea is in line with the concept of "highest safe reimbursement" introduced in the main text to formulate the network-formation and bargaining problem).

IV- Subgames S where d makes offers to both u_1 and u_2: These subgames S are all subgames starting at the first stage of an odd period τ where no link between d and any u_i has already been formed yet. Proposed strategy by d at all such S is as follows:

$$t^*_d(S) = \left(\tilde{t}_1, 0\right) \tag{A.5}$$

Note that since at these subgames, d has to make two offers, its strategy is a pair rather than one number.

V- Subgames S where u_1 and u_2 are simultaneously considering offers t_1 and t_2 from d: These are all subgames S starting at the second stage of an odd period of the game. Strategies here are denoted $A^*_{u_i}(S)$ which can take values of 1 (for "accept") and 0 (for "reject"). The proposed strategies are:

$$A^*_{u_2}(S) = 1\left[t_2 > 0\right] \tag{A.6}$$

$$A^*_{u_1}(S) = 1\left[\left(t_2 > 0 \,\&\, t_1 > 0\right)\ or\ \left(t_2 \leq 0 \,\&\, t_1 \geq \tilde{t}_1\right)\right] \tag{A.7}$$

The idea behind (A.6) is that u_2 (who expects to be excluded from the network and make a zero profit at the equilibrium) accept any offer that gives him anything more that 0. The idea behind (A.7) is that when $t_2 > 0$, upstream firm u_1 knows that u_2 will accept and if u_1 does not accept, the game will enter a subgame of type **II** which is anticipated to lead to no deal between u_1 and d, hence a profit of 0 to u_1. Thus, if $t_2 > 0$, then u_1 will accept any t_1 that would give u_1 a higher profit than zero, which means any $t_1 > 0$ will be accepted. But If $t_2 \leq 0$, then u_1 knows that it will have the opportunity to counter d's offer if it's not sufficiently profitable for u_1. Therefore, if $t_2 \leq 0$, upstream firm u_1 is expected to reject t_1 if $t_1 < \tilde{t}_1$. That is, as I will discuss in more details, if u_1 could do better by waiting and countering d's offer.

VI- Subgames S where d is considering offers t_1 and t_2 from u_1 and u_2 respectively: These are all subgames S starting at the second stage of an even period of the game. Strategy of d is denoted $A^*_d(S)$ which can be one of the pairs in the set $\{0,1\}^2$. The second element of the pair represents d's decision on whether to accept u_1's offer and the second element represents that for u_2's. In all such subgames S, the strategy of d works as follows: Denote $O = \left\{\delta_d\left(\pi_1 - \tilde{t}_1\right),\ \pi_1 - t_1,\ \pi_2 - t_2,\ \pi_3 - t_1 - t_2\right\}$.

$$A^*_d(S) = \begin{cases} if\ \pi_1 - t_1 \in \arg\max(O): & (1,0) \\ else\,if\ \pi_2 - t_2 \in \arg\max(O): & (0,1) \\ else\,if\ \delta_d\left(\pi_1 - \tilde{t}_1\right) \in \arg\max(O): & (0,0) \\ else & (1,1) \end{cases} \tag{A.8}$$

The idea behind this formulation is that d accepts any offer that gives it a higher net present value of profit from S on, and d breaks ties in favor of u_1.

A.3 Showing that the Proposed Strategy is a SPNE

Since this game has infinitely many periods, standard backwards induction cannot be used to analyze its subgame perfect equilibria. But since the discount factors are less than 1, I use the One-Shot-Deviation-Principle (OSDP). According to OSDP, no firm can strictly profit from changing its strategy only in one period of the subgame without changing its strategy elsewhere, where the strategies of the other firms have also been fixed. The following theorem show that this holds for the strategy profile introduced in section A.2.

Theorem 1. *The strategy profile described in section A.2 is a SPNE of the game set up in section A.1.*

Proof of theorem 1: I prove that there is no one-shot deviation by any firm at any subgame S. I study different subgames according to the types specified in section A.2.

Subgames S of types I and II: For these types, as mentioned before, we have a unique no-trade outcome from the bilateral Rubinstein bargaining. Therefore, for these subgames, I did not even specify strategies and just mentioned that no-trade is what we expect in continuation.

Subgames S of type III: In these subgames, we have $t^*_{u_1}(S) = \bar{t}_1$ and $t^*_{u_2}(S) = 0$. I first compute the profits that u_1 and u_2 make under these strategies, and then show that neither u_1 nor u_2 has a deviation that would strictly increase its profit. Lemma 1 below enables us to calculate payoffs under $t^*_{u_1}(S) = \bar{t}_1$ and $t^*_{u_2}(S) = 0$.

Lemma 1. *If at a subgame S of type III, u_1 and u_2 make offers according to (A.3) and (A.4), then in the second stage of that period (which is a subgame of type VI), downstream firm d would accept only u_1's offer if d acts according to (A.8).*

Proof of lemma 1: I first show that $\pi_1 - t^*_{u_1}(S) \geq \pi_2 - t^*_{u_2}(S)$:

$$\begin{aligned}
\pi_1 - t^*_{u_1}(S) = \pi_1 - \bar{t}_1 &= \pi_1 - \min\left(\pi_1 - \pi_2, \pi_1 \frac{1-\delta_d}{1-\delta_d\delta_{u_1}}\right) \\
&= \max\left(\pi_1 - (\pi_1 - \pi_2), \pi_1 - \left(\pi_1 \frac{1-\delta_d}{1-\delta_d\delta_{u_1}}\right)\right) \\
&= \max\left(\pi_2, \pi_1\left(1 - \frac{1-\delta_d}{1-\delta_d\delta_{u_1}}\right)\right) \qquad \text{(A.9)} \\
&\geq \pi_2 = \pi_2 - t^*_{u_2}(S)
\end{aligned}$$

Now I show $\pi_1 - t^*_{u_1}(S) \geq \pi_3 - t^*_{u_1}(S) - t^*_{u_2}(S)$. This is immediate from $t^*_{u_1}(S) \geq 0$, $\pi_1 - t^*_{u_1}(S) \geq \pi_2 - t^*_{u_2}(S)$, and $\pi_2 > \pi_3$.

Finally, I show $\pi_1 - t^*_{u_1}(S) \geq \delta_d\left(\pi_1 - \tilde{t}_1\right)$. First I rewrite the right-hand-side:

$$
\begin{aligned}
\delta_d\left(\pi_1 - \tilde{t}_1\right) = \delta_d\left(\pi_1 - \delta_{u_1}\bar{t}_1\right) &= \delta_d\left(\pi_1 - \delta_{u_1}\min\left(\pi_1 - \pi_2, \pi_1\frac{1-\delta_d}{1-\delta_d\delta_{u_1}}\right)\right) \\
&= \delta_d \max\left(\pi_1 - \delta_{u_1}\left(\pi_1 - \pi_2\right), \pi_1\left(1 - \frac{1-\delta_d}{1-\delta_d\delta_{u_1}}\right)\right) \\
&= \delta_d \max\left(\pi_1 - \delta_{u_1}\left(\pi_1 - \pi_2\right), \pi_1\frac{1-\delta_{u_1}}{1-\delta_d\delta_{u_1}}\right) \\
&= \max\left(\delta_d\left(\pi_1 - \delta_{u_1}\left(\pi_1 - \pi_2\right)\right), \pi_1\frac{1-\delta_{u_1}}{1-\delta_d\delta_{u_1}}\delta_d\right)
\end{aligned}
\tag{A.10}
$$

The fact that $\delta_d\left(\pi_1 - \tilde{t}_1\right)$ is the maximum between two terms gives us two natural cases to examine:

Case 1: $\delta_d\left(\pi_1 - \tilde{t}_1\right) = \pi_1\frac{1-\delta_{u_1}}{1-\delta_d\delta_{u_1}}\delta_d$. In this case, given (A.9), it is immediate that $\delta_d\left(\pi_1 - \tilde{t}_1\right) \leq \pi_1 - t^*_{u_1}(S)$ and the proof of the lemma is complete.

Case 2: $\delta_d\left(\pi_1 - \tilde{t}_1\right) = \delta_d\left(\pi_1 - \delta_{u_1}\left(\pi_1 - \pi_2\right)\right)$. Given (A.10), under case 2 we have:

$$
\begin{aligned}
&\delta_d\left(\pi_1 - \delta_{u_1}\left(\pi_1 - \pi_2\right)\right) \geq \pi_1\frac{1-\delta_{u_1}}{1-\delta_d\delta_{u_1}}\delta_d \\
&\Rightarrow \pi_1 - \delta_{u_1}\left(\pi_1 - \pi_2\right) \geq \pi_1\frac{1-\delta_{u_1}}{1-\delta_d\delta_{u_1}} \\
&\Rightarrow \pi_1\left(1-\delta_{u_1}\right) + \pi_2\delta_{u_1} \geq \pi_1\frac{1-\delta_{u_1}}{1-\delta_d\delta_{u_1}} \\
&\Rightarrow \pi_2\delta_{u_1} \geq \pi_1\left(1-\delta_{u_1}\right)\left(\frac{1}{1-\delta_d\delta_{u_1}} - 1\right) \\
&\Rightarrow \pi_2\delta_{u_1} \geq \pi_1\left(1-\delta_{u_1}\right)\left(\frac{\delta_d\delta_{u_1}}{1-\delta_d\delta_{u_1}}\right) \\
&\Rightarrow \pi_2 \geq \pi_1\left(1-\delta_{u_1}\right)\left(\frac{\delta_d}{1-\delta_d\delta_{u_1}}\right) \\
&\Rightarrow \pi_2\left(1-\delta_d\delta_{u_1}\right) \geq \pi_1\left(1-\delta_{u_1}\right)\delta_d \\
&\Rightarrow \pi_2 \geq \delta_d\left(\pi_1\left(1-\delta_{u_1}\right) + \pi_2\delta_{u_1}\right) \\
&\Rightarrow \pi_2 \geq \delta_d\left(\pi_1 - \delta_{u_1}\left(\pi_1 - \pi_2\right)\right) \\
&\Rightarrow \pi_2 \geq \delta_d\left(\pi_1 - \tilde{t}_1\right)
\end{aligned}
\tag{A.11}
$$

Together, (A.9) and (A.11) imply that $\pi_1 - t^*_{u_1}(S) \geq \delta_d \left(\pi_1 - \tilde{t}_1\right)$ which completes the proof of the lemma. $\square$

Now, given lemma 1, we know that under $t^*_{u_1}(S) = \bar{t}_1$ and $t^*_{u_2}(S) = 0$, the present values of the profits for u_1 and u_2 will be $\bar{t}_1$ and 0 respectively. Now I show that u_1 and u_2 cannot strictly profit from deviating from these strategies. First consider u_2. If u_2 deviates to some $t_2 < t^*_{u_2}(S) = 0$, then either its offer will still not get accepted by d, in which case u_2 should still expect a present value profit of 0, or the offer will be accepted in which case u_2 will make a negative profit. Therefore, u_2 cannot strictly benefit from deviating from 0 to a negative number. Also, if u_2 deviates to some $t_2 > t^*_{u_2}(S) = 0$, since according to OSDP we're assuming that u_1 is sticking to its $t^*_{u_1}(S) = \bar{t}_1$ strategy and d is sticking to its strategy in subgames of type VI, it's straightforward to show that d will still only accept u_1's offer, taking the game into a subgame of type I in the next period, implying a profit of 0 to u_2. Thus, u_2 cannot strictly profit by increasing its price either.

Now let's consider potential deviations by u_1. If u_1 deviates to some $t_1 < t^*_{u_1}(S) = \bar{t}_1$, then one can use a similar argument to lemma 1 to show that d will still only accept u_1's offer, taking the game to a type I subgame where nothing will change from then on. This will lead to a present value profit of $t_1 < \bar{t}_1$ for u_1 which is less that what u_1 got under $t^*_{u_1}(S) = \bar{t}_1$. So, u_1 does not have an incentive to reduce its price. To examine potential incentives by u_1 to increase its price, given that $t^*_{u_1}(S) = \bar{t}_1 = \min\left(\pi_1 - \pi_2, \pi_1 \frac{1-\delta_d}{1-\delta_d \delta_{u_1}}\right)$, I consider the follwoing two cases:

Case 1, $\bar{t}_1 = \pi_1 - \pi_2$: if u_1 deviates to $t_1 > t^*_{u_1}(S) = \bar{t}_1$, then given that the OSDP assumption that u_2 is keeping offering a transfer of 0 and d will sticks to the same algorithm when deciding the offers in the next stage of this period, one can show that d will decide to accept u_2's offer since now $\pi_1 - t_1 < \pi_1 - \bar{t}_1 = \pi_1 - (\pi_1 - \pi_2) = \pi_2 = \pi_2 - t^*_{u_2}(S)$.

Case 2, $\bar{t}_1 = \pi_1 \frac{1-\delta_d}{1-\delta_d \delta_{u_1}} < \pi_1 - \pi_2$: if u_1 deviates to $t_1 > t^*_{u_1}(S) = \bar{t}_1$, then we have $t_1 > \pi_1 \frac{1-\delta_d}{1-\delta_d \delta_{u_1}}$. Now I show that under this t_1, we have

$$\delta_d \left(\pi_1 - \tilde{t}_1\right) > \pi_1 - t_1$$

To show this, note that by (A.10), we have:

$$\delta_d\left(\pi_1-\tilde{t}_1\right)=\max\left(\delta_d\left(\pi_1-\delta_{u_1}\left(\pi_1-\pi_2\right)\right),\pi_1\frac{1-\delta_{u_1}}{1-\delta_d\delta_{u_1}}\delta_d\right)$$
$$\geq\pi_1\frac{1-\delta_{u_1}}{1-\delta_d\delta_{u_1}}\delta_d$$
$$=\pi_1-\pi_1\frac{1-\delta_d}{1-\delta_d\delta_{u_1}}=\pi_1-\bar{t}_1>\pi_1-t_1$$

Therefore, given the strategy specified for d in subgames of type VI, d will never accept only u_1's offer, since d at least prefers accepting none of the offers to accepting only u_1's. It's straightforward that d will not accept both. So, the only options are d accepting only u_2's offer of zero or accepting none of the offers in this period. If d, accepts only u_2's offer, then as discussed before, u_1 will make a profit of zero, implying that moving up to $t_1>t^*_{u_1}(S)=\bar{t}_1$ is not a strictly profitable deviation for u_1. If d accepts none of the offers, then given that OSDP assumes that from now on, all firms stick to their original strategies, it's expected that next period d will offer $\left(\tilde{t}_1,0\right)$ and then u_1 will accept and u_2 will reject. This will bring u_1 a profit of $\tilde{t}_1\times\delta_{u_1}=\bar{t}_1$. So, given that u_1 is already making $\bar{t}_1$, this deviation will never be strictly profitable for u_1.

Subgames S of type IV: As a reminder, these are subgames where d is supposed to make a pair of offers (t_1,t_2) to u_1 and u_2. Again, I will first derive d's profit under the strategies specified in section A.2 and then I show there is no deviation available to d that would enable d to make a strictly higher profit than that. The proposed SPNE strategy for d was $t^*_d(S)=\left(\tilde{t}_1,0\right)$. Given the strategies of u_1 and u_2 when considering these offers (see (A.6) and (A.7)), it's expected that u_1 accept d's offer and u_2 will reject. This means that from the next round on, the game will enter a subgame of type I where no further change is supposed to happen. Therefore, the payoff to d is going to be:

$$\pi_1-\tilde{t}_1=\pi_1-\delta_{u_1}\bar{t}_1$$
$$=\pi_1-\delta_{u_1}\min\left(\pi_1-\pi_2,\pi_1\frac{1-\delta_d}{1-\delta_d\delta_{u_1}}\right)$$
$$=\max\left(\pi_1-\delta_{u_1}\left(\pi_1-\pi_2\right),\pi_1\left(1-\delta_{u_1}\frac{1-\delta_d}{1-\delta_d\delta_{u_1}}\right)\right)$$
$$=\max\left(\left(1-\delta_{u_1}\right)\pi_1+\delta_{u_1}\pi_2,\pi_1\frac{1-\delta_{u_1}}{1-\delta_d\delta_{u_1}}\right)\tag{A.12}$$

Now, in order to analyze the possible deviations, let's consider four cases.

Case 1- d deviates from $\left(\tilde{t}_1, 0\right)$ to some (t_1, t_2) where $t_2 > 0$ and $t_1 \leq 0$: In this case, given the strategies of $u_1$and u_2 regarding what offers to accept, only u_2 will accept d's offer. Then the game enters a subgame of type II following the acceptance by u_2 and nothing will not change afterwards. Therefore, the profit to d will be given by $\pi_2 - t_2$ which is strictly less than π_2. But by (A.12), d makes at least $(1 - \delta_{u_1})\,\pi_1 + \delta_{u_1}\pi_2$ under it's current strategy which is higher than π_2. So, d will not have the incentive to make such a deviation.

Case 2- d deviates from $\left(\tilde{t}_1, 0\right)$ to some (t_1, t_2) where $t_2 > 0$ and $t_1 > 0$: In this case, given the strategies of $u_1$and u_2 regarding what offers to accept, both u_1 and u_2 will accept the offers, after which all links are formed and nothing will change in the game. The profit that d will make in this situation will be

$$\pi_3 - t_1 - t_2 < \pi_3 < \pi_2 < \pi_1 - \tilde{t}_1$$

Therefore, d will also not have the incentive to make this type of deviation.

Case 3- d deviates from $\left(\tilde{t}_1, 0\right)$ to some (t_1, t_2) where $t_2 \leq 0$ and $t_1 > \tilde{t}_1$: In this case, given the strategies of $u_1$and u_2 regarding what offers to accept, only u_1 will accept d's offer. Then the game enters a subgame of type I following the acceptance by u_1 and nothing will not change afterwards. Therefore, the profit to d will be given by $\pi_1 - t_1$ which is less than $\pi_1 - \tilde{t}_1$. So, d will not have the incentive to make this type of deviation.

Case 4- d deviates from $\left(\tilde{t}_1, 0\right)$ to some (t_1, t_2) where $t_2 \leq 0$ and $t_1 < \tilde{t}_1$: In this case, given the strategies of $u_1$and u_2 regarding what offers to accept, both u_1 and u_2 will reject the offers. So, the game will enter a subgame of type II where u_1 will make an offer of $\bar{t}_1$ and u_2 will make an offer of 0. Then d will accept only u_1's offer and nothing else will change after that. The profit that d will get out of this process will be equal to

$$\delta_d\,(\pi_1 - \bar{t}_1) < \pi_1 - \bar{t}_1 < \pi_1 - \delta_{u_1}\bar{t}_1 = \pi_1 - \tilde{t}_1$$

which implies that d does not have an incentive to make this type of deviation either. So, I now have shown that in general d does not have an incentive to deviate from $\left(\tilde{t}_1, 0\right)$ given that the rest of the strategies, both by other firms and by d itself on the subsequent subgames stays fixed.

Subgames S of type V: These are subgames where u_1 and u_2 are supposed to simultaneously decide on whether to accept the offers made to them by d, where the offers are t_1 and t_2. Four the following four cases, I first derive the profits to u_1 and u_2 under the proposed strategies in (A.6) and (A.7), and then show that neither u_1 nor u_2 has a strictly profitable deviation.

Case 1) $t_2 > 0$ and $t_1 > 0$: In this case, given the strategies proposed in (A.6) and (A.7), both u_1 and u_2 accept the offers and the profits they make will be equal to t_1 and t_2 respectively. If u_1 deviates and rejects d's offer, then given that u_2 accepts, the game will enter a subgame of type II in the next period and nothing will change afterwards given the strategies of the firms. Therefore, u_1 will make a profit of 0 which is less than t_1. So, u_1 will not have an incentive to not accept. Using a similar argument, u_2 does not have an incentive to deviate.

Case 2) $t_2 > 0$ and $t_1 \leq 0$: In this case, given the strategies proposed in (A.6) and (A.7), only u_2 accepts the offer and the profits u_1 and u_2 make will be equal to 0 and t_2 respectively (given that these strategies will take us to a subgame of type II from which point no further changes are expected to happen). If u_1 deviates to accepting, then the profit to u_1 will be $t_1 \leq 0$. So, u_1 will not have an incentive for such a deviation. If u_2 deviates to rejecting, then in the next round we will enter a subgame of type III, which given the strategies of the firms, is expected to reward u_2 with a profit of 0. But u_2's current profit is $t_2 > 0$ so u_2 will not deviate.

Case 3) $t_2 \leq 0$ and $t_1 < \tilde{t}_1$: In this case, given the strategies proposed in (A.6) and (A.7), neither of u_1 and u_2 accept the offers. So, the game will enter a subgame of type III next round and it will deliver present value payoffs of $\delta_{u_1}\bar{t}_1 = \tilde{t}_1$ and 0 for u_1 and u_2 respectively. If u_1 deviates to accepting, then, given u_2 is rejecting, the game will enter a subgame of type I from which point on no further changes occur. So, the payoff to u_1 would be t_1 which is less than $\tilde{t}_1$, the payoff u_1 is making without deviating. So, u_1 will not deviate. If u_2 deviates to accepting, given that u_1 is rejecting, the game will enter a subgame of type II from which point on no further changes happen. So, the profit this deviation gives to u_2 will be t_2. But t_2 is weakly less than u_2's profit without deviation: 0. So, u_2 also will not deviate.

Case 4) $t_2 \leq 0$ and $t_1 \geq \tilde{t}_1$: In this case, given the strategies proposed in (A.6) and (A.7), only u_1 accepts the offer and the profits u_1 and u_2 make will be equal to t_1 and 0 respectively (given that these strategies will take us to a subgame of type I from which point no further changes are expected to happen). If u_1 deviates to rejecting, given that u_2 is also rejecting, the game will enter a subgame of type III which will give a present value payoff of $\delta_{u_1}\bar{t}_1 = \tilde{t}_1$ to u_1. But u_1 is already making $t_1 \geq \tilde{t}_1$. So, u_1 does not want to deviate. If u_2 deviates to accepting, it will make a profit of t_2 which is weakly less than 0, what u_2 is making without deviation. So, u_2 will not deviate either.

Subgames S of type VI: As a reminder, these subgames are those where d is considering the pair of offers (t_1, t_2) from u_1 and u_2. The proposed strategy for d was specified in (A.8).

Claim 1. If d accepts only u_1's offer, then the present value of the profit to d will be $\pi_1 - t_1$.

To see this, note that once d accepts u_1's offer and rejects u_2's, the game will enter a subgame of type I after which no further change occurs.

Claim 2. If d accepts only u_2's offer, then the present value of the profit to d will be $\pi_2 - t_2$.

A similar argument proves this remark.

Claim 3. If d accepts neither of the offers, then the present value of the profit to d will be $\delta_d \left(\pi_1 - \tilde{t}_1\right)$.

To see this, note that if d does not accept any of the offers, we will enter a subgame of type IV where d will make an offer of $\tilde{t}_1$ to u_1 which will be accepted and offer of 0 to u_2 which will be rejected.

Claim 4. If d accepts both of the offers, then the present value of the profit to d will be $\pi_3 - t_1 - t_2$.

Given the above claims, it's straightforward that the strategy specified in (A.8) is optimal and if d makes a one-shot deviation from this strategy, d will be weakly worse off. So, d will not have an incentive to deviate. This completes the proof of the theorem. □

Theorem 1 shows that the strategies set out in section A.2 are a SPNE of the game in A.1. My next theorem relates this game to the formulation offered in the main text. The idea is that as $\Delta \to^{+} 0$, (that is, as the time difference between two consecutive periods shrinks), the transfers predicted by the SPNE of the game converges to what my bargaining formulation predicts. Before stating the theorem, observe that the network expected to realize on the equilibrium path of the above SPNE is $G^* = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and the corresponding transfer is $t_1^* = \bar{t}_1$, which, from equation (A.1), is equal by definition to $\min\left(\pi_1 - \pi_2, \pi_1 \frac{1-\delta_d}{1-\delta_d\delta_{u_1}}\right)$.

Theorem 2. *Denote $\gamma_i = \frac{r_d}{r_d + r_{u_i}}$. Then, the limit of $\bar{t}_1$ as $\Delta \to^+ 0$ is given by*

$$\lim_{\Delta \to^+ 0} \bar{t}_1 = \min\left(\pi_1 - \pi_2, \pi_1\gamma_1\right) \tag{A.13}$$

Proof of theorem 2:

$$\begin{aligned}
\lim_{\Delta \to^+ 0} \bar{t}_1 &= \lim_{\Delta \to^+ 0} \min\left(\pi_1 - \pi_2, \pi_1 \frac{1-\delta_d}{1-\delta_d\delta_{u_1}}\right) \\
&= \min\left(\pi_1 - \pi_2, \pi_1 \times \lim_{\Delta \to^+ 0}\left(\frac{1-\delta_d}{1-\delta_d\delta_{u_1}}\right)\right) \\
&= \min\left(\pi_1 - \pi_2, \pi_1 \times \lim_{\Delta \to^+ 0}\left(\frac{1-e^{-r_d\Delta}}{1-e^{-\left(r_{u_1}+r_d\right)\Delta}}\right)\right) \\
&= \min\left(\pi_1 - \pi_2, \pi_1 \times \frac{r_d}{r_d + r_{u_1}}\right) \\
&= \min\left(\pi_1 - \pi_2, \pi_1\gamma_1\right)
\end{aligned}$$

□

Note that according to (A.13), this SPNE of the game coincides with the formulation of used in the main text of the paper for the empirical analysis of the health insurance markets. To see this, note that the construction $\gamma_i = \frac{r_d}{r_d + r_{u_i}}$ of γ_i matches the way bargaining parameters are defined in Binmore et al. [1986], Collard-Wexler et al. [2014]. So, γ_1 is interpreted is u_1's bargaining power when negotiating with d. Given that the total additional surplus to u_1 and d from contracting with one another is equal to π_1, it is expected that a Nash Bargaining between u_1 and d divides this total surplus between them in a way that they get γ_1 and $1 - \gamma_1$ fractions of the total surplus respectively. So, the Nash Bargaining transfer between u_1 and d, when d and u_2 are not connected, is expected to be exactly $\pi_1\gamma_1$ which is one of the two terms in (A.13). Now also notice that the other term (i.e. $\pi_1 - \pi_2$), is exactly the highest safe rate that u_1 can charge to d as defined in the definition of best safe reimbursement rates in the main text. To see this, note that if u_1 charges anything more than $\pi_1 - \pi_2$, then d will have strictly profit from replacing u_1 with u_2 at, say, $t_2 = 0$ where u_2 is also willing to participate. But when u_1 is charging $\pi_1 - \pi_2$, the only way for d to strictly profit from replacing u_1 with u_2 is to pay u_2 a negative amount, which violates incentive compatibility for u_2. Therefore, $\pi_1 - \pi_2$ is in fact the highest safe transfer that u_1 can charge. So, to sum up, the strategies that were proven to constitute an SPNE of the game imply that at the equilibrium, u_1 charges the Nash Bargaining price unless under the Nash bargaining price, there is a danger of replacement (that is, unless the Nash Bargaining price is not "safe" from replacement threats), in which case u_1 charges the highest price that is in fact safe from substitution.

The reason why the logic of the formulation in the main text is sustained as an equilibrium outcome of the extensive form game in section A.1 was explained at a technical and detailed level in sections A.2 and A.3 (By "the logic", I mean going with Nash Bargaining if it's safe from replacement threats, and going with the highest safe reimbursement otherwise). To see this at a more intuitive level, suppose that the Nash Bargaining transfer between u_1 and d gives a higher profit to d than d could get from replacing u_1 with u_2 and being charged a price of 0 by u_2. Now consider a subgame of the offers and counter offers game where u_1 and u_2 are simultaneously making offers to d. Suppose that u_2 is asking for a transfer of 0 and u_1 is asking for strictly more than the Nash Bargaining price and also strictly more than $\pi_1 - \pi_2$. In this situation, the optimal response for d isn't one where d accepts u_2's offer and rejecting u_1's. The optimal response for d is, rather, rejecting both offers and waiting till the next period where d will be supposed to make offers, and then offer the Nash Bargaining price to u_1. In other words, as long as the Nash Bargaining price is safe from substitution threats, any threat from d of replacing u_1 with u_2 is an empty threat and it's common knowledge that it is so. Therefore, the replacement threat will not affect the outcome of the bargaining if the Nash Bargaining price is safe from replacement. That's roughly the intuition for why the effects of replacement threats start to kick in the form of a minimum function.

B The Regression Fixed-Point Algorithm

In this appendix, I introduce a computational approach to solve the optimization problem used in the fourth and last step of the estimation procedure (i.e. estimating fixed costs of coverage) in the estimation section of the main text of the paper. Notations in this appendix are separate from the notations in the main text unless otherwise is stated. The reason is that I introduce the algorithm for a general problem that I set up below, not just for the problem in the main text. The rest of this appendix is organized as follows. First, I set up the general estimation problem. Then, I introduce the Regression Fixed-Point approach. Finally, I show that the Regression Fixed-Point algorithm converges to the solution of the estimation problem. I conclude by discussing the more general application of the Regression Fixed-Point approach beyond this paper.

B.1 The Estimation Problem

We have N observations for the following regression:

$$y_n = \beta x_n + \epsilon_n \tag{B.1}$$

where $y \in \mathbb{R}^N$ is the dependent variable, x is a vector of independent variables, β is the vector of true coefficients, and ϵ is an i.i.d normally distributed error term. The key feature of this problem is that y is not fully observed (x is observed). We only know that $y \in A$ where $A \subset \mathbb{R}^N$. One instance is when y is coming from a discrete choice model, like discrete choice single agent decision making, or a model involving network formation, like the model in the main text of this paper. In these cases, for at least some n, instead of observing y_n, we only know a bound on y_n. Or for some n and n', we have a restriction on the relationship between y_n and $y_{n'}$. For instance, in the estimation problem in the main text, for some hospital-insurer pairs, we only had a lower bound for the corresponding fixed costs. Those lower bounds came from the imposed no-gains-from-trade condition for pairs that do not have a link in the data. Similarly, we had a lower bound on the differences between some of the fixed costs. Those bounds came from the restriction that ensured no incentive for replacement.

The objective is to go from the partial identification of y (given by some imposed economic model) to full identification, by assuming that the true y is the one that is best fitted, among all candidates in A, to some observable characteristics x in as specified in the linear regression (B.1) (In the application in the main text, these characteristics were hospital fixed effects and insurer fixed effects). As such, $\hat{y}$ is defined by:

$$\hat{y} = \arg\min_{y \in A} \sigma(y) \tag{B.2}$$

where $\sigma(y)$ is the standard error of the regression when y is plugged into (B.1). Solving this problem using plain constrained optimization can be computationally demanding, especially when N is large or when this problem needs to be solved multiple times within a loop, nested in a larger iterative process. The Regression Fixed-Point algorithm substantially reduces the computational burden.

B.2 The Algorithm

The algorithm has three main steps. The input to the algorithm is an arbitrary $y^0 \in A$. The i-th iteration of the algorithm updates y^{i-1} to y^i. We repeat until the output of the algorithm converges to some $\bar{y}$. In next section, I show that, under some conditions, $\bar{y}$ is a global optimum of the minimization problem in (B.2). Each iteration i of the algorithm has the three following steps:

Step 1: Plug y^{i-1} into (B.1) and run the regression to get estimated coefficients β^{i-1}.

Step 2: Predict a new left-hand side vector $y^{p,i-1}$ using the estimated coefficients:

$$y^{p,i-1} = \beta^{i-1} x \tag{B.3}$$

Step 3: Find y^i as the closest point (by point, I mean a vector in $\mathbb{R}^N$) to $y^{p,i-1}$ among members of A:

$$y^i = \arg\min_{y \in A} \sum_{n \in \{1,\ldots,N\}} \left(y_n - y_n^{p,i-1}\right)^2 \tag{B.4}$$

where, if the argmin is not a singleton, an arbitrary minimizer is chosen, unless y^{i-1} is a member of the argmin, in which case the algorithm chooses $y^i = y^{i-1}$.

Each iteration of this algorithm is computationally straightforward. The only potentially worrisome step is step 3 where we have to solve a constrained optimization problem in $\mathbb{R}^N$. But this step is usually very fast in most natural cases where the restrictions on y given by A are in the form of independent upper- or lower-bounds for different observations or upper- or lower-bounds on the differences. For instance, suppose that A gives an exact value of y_n^A for some n and a lower bound of z_n^A for the rest of the observations. In that case, y^i is given by $y_n^i = y_n^A$ for observations where A specifies a point and by $y_n^i = \max\left(y_n^{p,i-1}, z_n^A\right)$ for observations where A specifies a lower bound.

B.3 Properties of the Regression Fixed-Point Algorithm

In this section, I first show that the algorithm always converges to some $\bar{y} \in A$, and then I characterize the necessary conditions for $\bar{y}$ to be a local optimum to the minimization problem in (B.2).

Theorem 3. *The Regression Fixed-Point algorithm converges to a $\bar{y} \in A$.*

Proof of theorem 3. By the definition of standard error, we know that $\forall y \in \mathbb{R}^N : \sigma(y) \geq 0$. This, combined with the following lemma proves the convergence:

Lemma 2. *For every two consecutive iterations $i-1$ and i, if $y^{i-1} \neq y^i$, then $\sigma\left(y^i\right) < \sigma\left(y^{i-1}\right)$.*

Proof of Lemma 2. If $y^{i-1} \neq y^i$, then, (B.4) and the choice rule for non-singleton argmin imply:

$$\sum_{n\in\{1,\dots,N\}} \left(y_n^i - y_n^{p,i-1}\right)^2 < \sum_{n\in\{1,\dots,N\}} \left(y_n^{i-1} - y_n^{p,i-1}\right)^2 \tag{B.5}$$

But $\sum_{n\in\{1,\dots,N\}} \left(y_n^{i-1} - y_n^{p,i-1}\right)^2 = \sigma^2\left(y^{i-1}\right)$. Therefore:

$$\sum_{n\in\{1,\dots,N\}} \left(y_n^i - y_n^{p,i-1}\right)^2 < \sigma^2\left(y^{i-1}\right)$$

On the other hand, given (B.3):

$$\sum_{n\in\{1,\dots,N\}} \left(y_n^i - y_n^{p,i-1}\right)^2 = \sum_{n\in\{1,\dots,N\}} \left(y_n^i - \beta^{i-1}x\right)^2 \geq \min_\beta \sum_{n\in\{1,\dots,N\}} \left(y_n^i - \beta x\right)^2$$

$$\Rightarrow \min_\beta \sum_{n\in\{1,\dots,N\}} \left(y_n^i - \beta x\right)^2 \leq \sum_{n\in\{1,\dots,N\}} \left(y_n^i - y_n^{p,i-1}\right)^2 \tag{B.6}$$

But by the construction of linear regression, $\min_\beta \sum_{n\in\{1,\ldots,N\}} \left(y_n^i - \beta x\right)^2 = \sigma^2(y)$. Therefore, combining (B.5) and (B.6), we get:

$$\sigma\left(y^i\right) < \sigma\left(y^{i-1}\right)$$

which completes the proof of the lemma as well as the theorem. $\square$

I so far established that the algorithm has a fixed point. Now observe that each global maximum of the optimization problem (B.2) is a fixed point.

Proposition 1. *If $y^i \in \arg\min_{y\in A} \sigma(y)$, then $y^{i+1} = y^i$.*

Proof of proposition 1. The proof follows directly from Lemma 2. $\square$

I now show that if set A is convex, then every fixed point $\bar{y}$ of the algoritm solves the optimization problem in (B.2). Most common applications satisfy convexity of A. For instance, observation-specific lower and/or upper bounds on y_n gives a convex A. So do pair-specific lower and/or upper bounds on $y_n - y_{n'}$.[63]

Theorem 4. *If set A is convex, then every fixed point $\bar{y}$ of the Regression Fixed-Point algorithm solves (B.2).*

Proof of theorem 4. Define $B = \left\{y \in \mathbb{R}^N : \exists \beta \in \mathbb{R}^M : y = \beta x\right\}$. That is, B is the set of all y that satisfy all N equations specified by (B.1) with zero error terms.

Lemma 3. *B is a closed and convex set.*

Proof of Lemma 3. For all β and β' in $\mathbb{R}^M$ and all $\alpha \in [0,1]$, the linear combination $\beta'' = \alpha\beta + (1-\alpha)\beta'$ is also in $\mathbb{R}^M$. So, $\beta'' x \in B$. Thus B is convex. The proof of closedness is more also straightforward and is left to the reader. $\square$

Now suppose $\bar{y}$ is a fixed point of the Regression Fixed-Point algorithm. Also take

$$\tilde{y} \in \arg\min_{y\in B} \sum_{n\in\{1,\ldots,N\}} \left(y_n - \bar{y}_n\right)^2 \tag{B.7}$$

In other words, $\tilde{y} = \bar{\beta}x$ where $\bar{\beta}$ comes from regressing $\bar{y}$ on x. Now observe that by $\bar{y}$

[63] Also for non-convex A it can likely be shown that for any fixed-point $\bar{y}$ of the Regression Fixed-Point algorithm that does NOT solve (B.2), the algorithm reaches $\bar{y}$ at its fixed point only if it starts out at $y^0 = \bar{y}$ (In other words, $\bar{y}$ is an "unstable" fixed point). But proving this result is beyond the scope of this appendix.

being a fixed point, we also have:

$$\bar{y} = \arg\min_{y \in A} \sum_{n \in \{1,\ldots,N\}} (y_n - \tilde{y}_n)^2 \tag{B.8}$$

That is, $\bar{y}$ is the closest point in A to $\tilde{y}$ and also $\tilde{y}$ is the closest point in B to $\bar{y}$. Note that the distance between $\bar{y}$ and $\tilde{y}$ is exactly $\sigma(\bar{y})$. In order to show that $\bar{y}$ solves the optimization problem (B.2), one would need to show there is no $\bar{y}^{'} \in A$ with $\sigma\left(\bar{y}^{'}\right) < \sigma(\bar{y})$. A sufficient condition for there being no such $\bar{y}^{'}$ is proven in the following lemma.

Lemma 4. *There is no pair* $\left(\bar{y}^{'}, \tilde{y}^{'}\right)$ *with* $\bar{y}^{'} \in A$ *and* $\tilde{y}^{'} \in B$ *such that*

$$\sum_{n \in \{1,\ldots,N\}} \left(\tilde{y}_n^{'} - \bar{y}_n^{'}\right)^2 < \sum_{n \in \{1,\ldots,N\}} (\tilde{y}_n - \bar{y}_n)^2 \tag{B.9}$$

Proof of Lemma 4. I use convexity of A (by assumption) and convexity of B (by lemma 3) and apply the separating hyperplane theorem twice.

First observe that $\tilde{y}$ is the closest point in B to $\bar{y}$ and B is convex. Hence, by the separating hyperplane theorem, for all $\tilde{y}^{'} \in B$ we have:

$$\left(\tilde{y}^{'} - \tilde{y}\right) \cdot (\tilde{y} - \bar{y}) \geq 0 \tag{B.10}$$

Similarly, by convexity of A, for all $\bar{y}^{'} \in A$, the separating hyperplane theorem gives:

$$\left(\bar{y} - \bar{y}^{'}\right) \cdot (\tilde{y} - \bar{y}) \geq 0 \tag{B.11}$$

Now observe that

$$\begin{aligned} \left(\tilde{y}^{'} - \bar{y}^{'}\right) \cdot (\tilde{y} - \bar{y}) &= \left[\left(\tilde{y}^{'} - \tilde{y}\right) + (\tilde{y} - \bar{y}) + \left(\bar{y} - \bar{y}^{'}\right)\right] \cdot (\tilde{y} - \bar{y}) \\ &= \left(\tilde{y}^{'} - \tilde{y}\right) \cdot (\tilde{y} - \bar{y}) + (\tilde{y} - \bar{y}) \cdot (\tilde{y} - \bar{y}) + \left(\bar{y} - \bar{y}^{'}\right) \cdot (\tilde{y} - \bar{y}) \end{aligned} \tag{B.12}$$

The first and the third term in (B.12) are non-negative by (B.10) and (B.11) respectively. Therefore:

$$\left(\tilde{y}^{'} - \bar{y}^{'}\right) \cdot (\tilde{y} - \bar{y}) \geq (\tilde{y} - \bar{y}) \cdot (\tilde{y} - \bar{y})$$

But given that $(\tilde{y} - \bar{y}) \cdot (\tilde{y} - \bar{y}) = |\tilde{y} - \bar{y}|^2$ and given that $\left(\tilde{y}^{'} - \bar{y}^{'}\right) \cdot (\tilde{y} - \bar{y}) \leq |\tilde{y} - \bar{y}| \times \left|\tilde{y}^{'} - \bar{y}^{'}\right|$, we can write:

$$|\tilde{y} - \bar{y}| \times \left|\tilde{y}^{'} - \bar{y}^{'}\right| \geq |\tilde{y} - \bar{y}|^2$$

$$\Rightarrow \left|\tilde{y}^{'} - \bar{y}^{'}\right| \geq |\tilde{y} - \bar{y}|$$

$$\Rightarrow \left|\tilde{y}^{'} - \bar{y}^{'}\right|^2 \geq |\tilde{y} - \bar{y}|^2$$

$$\sum_{n\in\{1,\ldots,N\}} \left(\tilde{y}_n^{'} - \bar{y}_n^{'}\right)^2 \geq \sum_{n\in\{1,\ldots,N\}} (\tilde{y}_n - \bar{y}_n)^2$$

which completes the proof of the lemma. □

Now I use lemma 4 to complete the proof of the theorem. Take some $\bar{y}^{'} \in A$. I now show $\sigma\left(\bar{y}^{'}\right) \geq \sigma(\bar{y})$ which will complete the proof of the theorem. Observe that:

$$\sigma^2\left(\bar{y}^{'}\right) = \min_{y\in B} \sum_{n\in\{1,\ldots,N\}} \left(y_n - \bar{y}_n^{'}\right)^2 \tag{B.13}$$

Denote one of the minimizers of (B.13) by $\tilde{y}^{'} \in B$. This gives:

$$\sigma^2\left(\bar{y}^{'}\right) = \sum_{n\in\{1,\ldots,N\}} \left(\tilde{y}_n^{'} - \bar{y}_n^{'}\right)^2$$

But by lemma 4 we know

$$\sum_{n\in\{1,\ldots,N\}} \left(\tilde{y}_n^{'} - \bar{y}_n^{'}\right)^2 \geq \sum_{n\in\{1,\ldots,N\}} (\tilde{y}_n - \bar{y}_n)^2 = \sigma^2(\bar{y})$$

So, $\sigma^2\left(\bar{y}^{'}\right) \geq \sigma^2(\bar{y})$, and, hence, $\sigma\left(\bar{y}^{'}\right) \geq \sigma(\bar{y})$. □

I used this algorithm in estimation section of the main text to estimate the hospital system specific fixed costs of coverage f_{sj} to insurers. In step 3 of each iteration of the algorithm, I use an approximation to which the results are not sensitive.

C The Constrained Optimization Problem for the Estimation of Fixed Costs of Inclusion and Bargaining Parameters

In the main text, I mentioned that the estimation procedure backs out the fixed costs of inclusion $\hat{f}_{sj}$ and bargaining parameters $\hat{\gamma}_{sj}$ by finding the values that minimize the standard deviation of the residuals in the specification $f_{sj} = \Omega \times \chi_s + FE_j + \upsilon_{sj}$ of the fixed costs, subject to some constraints. The formal statement of those constraints is left to this appendix. As a reminder, there were two types of constraints. First, for each system-insurer pair sj, the corresponding bargaining parameter γ_{sj} is equal to γ_{star} if s is either Partners or Tufts Medical Center, and equal to $\gamma_{nonStar}$ otherwise. Of course, γ_{star} and $\gamma_{nonStar}$ are both within the interval $[0, 1]$. The second restrictions came from imposing the equilibrium condisions of the model (i.e., pairwise stability for network formation and nash in nash bargaining with threats of replacement for bargaining) on the observed market outcome (G^*, T^*, P^*).

This above arrangement turns out the optimization problem into one in terms of the tuple $(f_{sj}, \gamma_{star}, \gamma_{nonStar}, \Omega, FE_j, \nu_{sj})$. The object of minimization in this problem is

$$\sum_{s,j} \nu_{sj}^2 \tag{C.1}$$

This is subject to the following constraints:

1. The characterization of the fixed costs of inclusion:

$$f_{sj} = \Omega \times \chi_s + FE_j + \upsilon_{sj} \tag{C.2}$$

2. Bargaining parameters:

$$(\gamma_{star}, \gamma_{nonStar}) \in [0, 1]^2 \tag{C.3}$$

3. Equilibrium condisions imposed on (G^*, T^*, P^*):

   (a) Lower-bound on f_{sj} from no gains-from-trade for every sj with $g_{sj}^* = 0$

$$f_{sj} \geq \underline{f}_{sj} \tag{C.4}$$

   (b) Upper-bound on f_{sj} from gains-from-trade for every sj with $g_{sj}^* = 1$

$$f_{sj} \leq f_{sj}^{NB(\gamma_{star}, \gamma_{nonStar})} \tag{C.5}$$

   (c) Differential-bound from no-replacement. For every $g_{sj}^* = 1$, for all s' such that $g_{s'j}^* = 0$, it must be that

$$f_{s'j} - f_{sj} \geq f_{sj}^{diff} \tag{C.6}$$

Some terms in (C.4), (C.5), and (C.6) need further explanation. In (C.4), the lower-bound $\underline{f}_{sj}$ is the minimum value for f_{sj} that would rationalize there being no gains from trade for hospital-system s and insurer j. Given that at this point in the estimation procedure, we have already estimated the demand for hospitals and insurance plans as well as the non-inpatient costs ξ_j to insurers, the value of $\underline{f}_{sj}$ can be computed. To see how the computation is done, denote

$$\pi_{d_j}^{NFC}(G,T,P) = \pi_{d_j}(G,T,P) + \sum_{s \; s.t. \; g_{sj}=1} f_{sj}$$

That is, $\pi_{d_j}^{NFC_s}(G,T,P)$ is the profit to insurer d_j under market outcome (G,T,P), just not accounting for the fixed costs of inclusion. Note that in order to compute $\pi_{u_s}(G,T,P)$ and $\pi_{d_j}^{NFC}(G,T,P)$, knowing the outcomes of the first three stages of the estimation process (i.e., hospital choice function, plan demand function, and non-inpatient costs) sufficient. We would not need to know either the fixed costs f_{sj} or bargaining parameters γ_{sj}. Now, $\underline{f}_{sj}$ is given by:[64]

$$\underline{f}_{sj} = \left(\max_{t_{sj}}\left(\pi_{u_s}\left(G^*_{+sj}, T^*_{+t_{sj}}\right) + \pi_{d_j}^{NFC}\left(G^*_{+sj}, T^*_{+t_{sj}}\right)\right)\right) - \left(\pi_{u_s}(G^*, T^*) + \pi_{d_j}^{NFC}(G^*, T^*)\right) \tag{C.7}$$

This way, for any $f_{sj} < \underline{f}_{sj}$, there will be some contract t_{sj} that system u_s and insurer d_j can sign and get a strict pareto-improvement in their profits compared to the status quo of (G^*, T^*, P^*). But there is no strict pareto-improving t_{sj} in case $f_{sj} \geq \underline{f}_{sj}$, which is exactly what we expect based on the pairwise stability condition.

In (C.5), $f_{sj}^{NB(\gamma_{star},\gamma_{nonStar})}$ is the level of fixed cost f_{sj} that would rationalize the observed reimbursement rate t^*_{sj} as the outcome of Nash Bargaining between u_s and d_j given the rest of the network structure and reimbursement rates, and assuming that $\gamma_{sj} = \gamma_{star}$ for $s \in \{Partners, Tufts\}$ and $\gamma_{sj} = \gamma_{nonStar}$ otherwise. Note that to compute $f_{sj}^{NB(\gamma_{star},\gamma_{nonStar})}$, we do not need to have identified $\pi_{d_j}(G,T,P)$. As with (C.4), here $\pi_{u_s}(G,T,P)$ and $\pi_{d_j}^{NFC}(G,T,P)$ –which are both identified at this point in the estimation procedure– are sufficient. The intuition behind the inequality in (C.5) would, therefore, be that either hospital-system u_s is charging the insurer the Nash Bargaining rate (i.e.,

[64]Note that in (C.7), like section 3 of the main text, I am suppressing the notation on P given that its response to formation of a link is anticipated by firms.

$f_{sj} = f_{sj}^{NB(\gamma_{star},\gamma_{nonStar})}$), or u_s is charging less than the Nash Bargining rate in order to deter a threat of replacement, in which case it must be that $f_{sj} < f_{sj}^{NB(\gamma_{star},\gamma_{nonStar})}$.

In (C.6), $f_{ss'j}^{diff}$ is the minimum amount by which $f_{s'j}$ needs to be larger than f_{sj} in order for the model to rationalize that insurer d_j would not strictly profit from replacing system u_s with system $u_{s'}$. Based on the bargaining conditions with threats of replacement that were specified in the main text of the paper, $f_{ss'j}^{diff}$ is computed in the following way:

$$f_{ss'j}^{diff} = \pi_{d_j}^{NFC}\left(G^*_{-sj+s'j}, T^*_{-sj+t^R_{s'j}(G^*,T^*,s)}\right) - \pi_{d_j}^{NFC}\left(G^*, T^*\right) \tag{C.8}$$

where $t_{s'j}^R\left(G^*, T^*, s\right)$ is d_j's reservation rate for replacing u_s with $u_{s'}$. That is, d_j's threat to u_s of replacing it with $u_{s'}$ would take place at the reimbursement rate of $t_{s'j}^R\left(G^*, T^*, s\right)$. This, in line with the definition of best safe reimbursement rate in the model section of the main text, would be the lowest rate that would satisfy incentive-compatibility and no-commitment conditions. Note that $t_{s'j}^R\left(G^*, T^*, s\right)$ is not a function the fixed costs of inclusion and can be computed at the end of the first three stages of the estimation procedure (it can be described in terms of the $\pi_{u_s}\left(\cdot,\cdot\right)$ functions only). Note that at least one of the inequalities in (C.6) has to hold with equality in case (C.5) holds strictly. The reason is that, if hospital system u_s is charging below the Nash Bargaining rate in order to deter a threat of replacement, according to the formulation of such threats in my model, it has to be charging its highest safe rate. This implies that at there should be at least one $u_{s'}$ such that u_s is just deterring a threat of being replaced with $u_{s'}$ by d_j.

The above derivations formally complete the optimization problem (C.1) over $(f_{sj}, \gamma_{star}, \gamma_{nonStar}, \Omega, FE_j, \nu_{sj})$, which I solve using the Regression-Fixed-Point algorithm. A computationally powerful feature of this procedure is that the constraints on each f_{sj} are computed completely independently of the values of other f_{sj}. This feature may not be present in other potential applications. For instance, if I were estimating a set of marginal costs instead of fixed costs, then the value of an in-network hospital's marginal cost would factor into the bounds on the marginal cost of an out-of-network hospital calculated from the equilibrium conditions of the model. In those cases, the procedure developed in this paper can still be applied but with the further complication that the bounds become interdependent.

D Further Discussion of Modeling Choices

D.1 Discussion of Model Choice

I believe that the model developed in this section is better suited than other potential candidates in the economic theory literature for structural empirical analysis of interactions among network formation, bargaining, and downstream price (i.e., premium) setting in vertical markets. Cooperative-based models of network formation and bargaining (e.g., Jackson [2005]) do not restrict the transfers to be only between firms who are contracting with one another. Such models, therefore, always deliver efficiency, which for the health insurance market, implies that only one insurer stays in the market and pays off all other insurers to stay out, in order to extract the most out of consumers. Non-cooperative models like that of Abreu and Manea [2012] offer the advantage of having a well-defined extensive form; but heavily restrict externalities by assuming for each link there is exactly one unit of surplus to divide between the two sides no matter what the rest of the network looks like. Models of many-to-many matching with transfers (e.g., Hatfield et al. [2013]) also heavily restrict externalities. Additionally, none of the models above captures the idea that transfers among firms can affect the total surplus (e.g., reimbursement rates can affect premiums and, hence, total market demand). Papers on vertical contracting (such as Segal [1999], Rey and Whinston [2013], Prat and Rustichini [2003]) give the full bargaining power to one side (either upstream or downstream) and do not capture the full range. Also, these papers examine settings with either a single upstream or a single downstream.[65] I do not use a model with insurer commitment to a hospital network (such as Liebman [2017] and Ho and Lee [2017a]) for multiple reasons as I discussed in the related-literature section in the main text.

I also do not use some other alternative bargaining models to NiN that have been developed in the literature. For instance, models such as Stole and Zwiebel [1996a,b], Dranove et al. [2016], Lee and Fong [2013], unlike NiN, capture the idea that if an insurer drops a hospital from its network, it can re-negotiate with other hospitals (e.g. Stole

[65] The only paper with multiple firms on both sides, to my knowledge, is Prat and Rustichini [2003], which restricts the externalities in a way that cannot accept standard demand functions and premium competition, as modeled in my paper and other empirical papers on vertical markets.

and Zwiebel [1996a,b], Dranove et al. [2016], Lee and Fong [2013]). Unlike my model, however, these approaches are based on the idea that the renegotiation takes place *after* the dropping and not *simultaneously* to it. This takes away the insurer's ability to play hospitals off against each other for network inclusion and can still predict unrealistically high reimbursement rates in response to network adequacy regulations. Also these approaches, when applied to settings with endogenous network formation and bargaining, are very computationally complex. Also the dynamic-based model of Lee and Fong [2013] is worth further discussion. Lee and Fong [2013] analyze both bargaining and network formation in a dynamic game context using the concept of Markov Perfect Equilibrium (MPE) a la Maskin and Tirole [1988]. To my knowledge, this approach has not yet been applied to an empirical setting most likely because MPE combined with network formation and bargaining is very computationally demanding, even if we abstract away from strategic premium setting as Lee and Fong [2013] do. In addition, that approach would need long panel data on many variables including some that may be hard to observe, like costs that firms incur from forming, retaining, and removing links with other firms.

Finally, my approach in this paper to incorporating threats of replacement into the bargaining formulation is one among multiple options. My approach says the negotiated reimbursement rate is the minimum between the Nash Bargaining rate and the highest safe rate. Another approach, for instance, would be to keep using the Nash Bargaining formulation (without taking any minimums) but have d_j's "outside option profit" (which is equal to $\pi_{d_j}\left(G^*_{-ij}, T^*_{-ij}\right)$ in standard NiN) equal to the maximum profit d_j could attain by either severing the link with u_i or replacing u_i with some $u_{i'}$ in a way that would satisfy incentive-compatibility and no-commitment constraints as specified in the model section of the main text. The reason why I choose my specific formulation over other possible options is that I can show that for a class of 2×1 games, it can be sustained as the outcome of a subgame perfect Nash equilibrium of an extensive form game which is constructed by extending the alternating-offers game in Binmore et al. [1986] to from a bilateral bargaining to two upstream firms and one downstream. See section A for more details.

D.2 Discussion of the Choice of Estimation Procedure

My estimation procedure is based on the one developed by Gowrisankaran et al. [2013]. I take a linear characterization of the insurers' fixed costs of inclusion f_{sj} based some some observables as the moment condition (Gowrisankaran et al. [2013] do this for insurer-specific margainal costs of hospitals). The characterization was:

$$f_{sj} = \mathit{\Omega} \times \chi_s + FE_j + \upsilon_{sj} \tag{D.1}$$

I then minimize the deviation from that moment condition (i.e., minimize the standard deviation of $\hat{\nu}$) by searching not only over the values of the right-hand-side variables, but also those of f_{sj} values themselves as well as bargaining parameters γ_{star} and $\gamma_{nonStar}$ subject to the equilibrium conditions of the model.

Alternatively, I could take a maximum likelihood approach. I instead of minimizing the standard devaiation of the error terms in the linear characterization of the fixed costs based on observables, I search for parameters that would maximize the likelihood of the observed data being an equilibrium of the model. I decided against that approach as it would be very computationally intensive. For each parameterization, it would be necessary to draw numerous instances of all ν_{sj} and simulate the equilibrium for each of them to compute the likelihood of the observed market outcome arising at equilibiurm; and that would be only one instance of the likelihood function. Finding the MLE estimates would require carrying out this expensive procedure iteratively and looking for the optimum.

Another alternative approach would be moment inequalities. I could exclude the term ν_{sj} from (D.1) and find the values of f_{sj}, γ_{star}, $\gamma_{nonStar}$, $\mathit{\Omega}$, and FE_j that would minimally violate the inequalities coming from the equilibium conditions of the model. I decided against that approach because using a moment inequalities estimation procedure can run into difficulties in exercises like this paper, where counterfactual analysis is an integral part to the research. In moment inequalities, the estimated parameters do not fully rationalize the model. Therefore, they would predict a different equilibrium than the observed market outcome in the data. So, running counterfactual simulation to evaluate the effects of a policy, it would not be clear how much of the difference between the simulation results and the data is because of the policy and how much of it is simply because the model did not rationalize the data in the first place.[66] This problem has been noticed in the literature. In their model of bargaining, Crawford et al. [2015] add shocks to firm profits *after* they estimate their model. In order to make sure that the shocks do rationalize the data, they draw them from truncated normal distributions, which by construction makes the shocks non-mean-zero. I prefer the approach taken in this paper which allows more control over the profit shocks ν_{sj} so that the mean is zero and the variance is minimized.

Finally, another possibility would be to make the approach more similar to Gowrisankaran et al. [2013]. That is, I could include no fixed costs of inclusion f_{sj} into the model, and instead consider hospital marginal costs c_i to be insurer specific, implying that certain hospital u_i could have different marginal costs of providing inpatient care c_{ij} to the same patient, depending on which plan d_j the patient is enrolled with. The next step would be to characterize c_{ij} based on some observables and minimize the deviation from that linear characterization subject to the equilibrium conditions of the model and subject to bargaining powers being no less than 0 and no more than 1. I decided against that approach for three reasons. First, such an approach would not use the available data on hospital marginal costs. Second, using such an approach could imply unrealistic estimates for c_{ij} values. It would, for example, imply that the c_{ij} for u_i =Boston Medical Center (the hospital, not the plan) and d_j =Celticare would be about twice as high as that for u_i =Boston Medical Center and $d_{j'}$ =NHP, as, absent fixed costs, that would be the only way to rationalize the fact that Celticare is excluding BMC hospital and includes MGH and BWH at very expensive rates, and NHP is including the BMC hospital at much a much lower rate. I find the fixed-cost explanation for Celticare including MGH and BWH but excluding BMC to be more realistic than the particular c_{ij} for u_i =Boston Medical Center and d_j =Celticare to be overly high even though BMC is a fairly cheap hospital compared to MGH and BWH. Finally, the third reason why I dediced against the c_{ij} approach –not as important as the two aforementioned reasons– is computational complexity. As appendix C points out, with

[66] Of course an alternative would be to compare the counterfactual simulation results to the results of simulating the equilibrium of the estimated model with *no policy*, instead of comparing it to the data. However, such an approach is implicitly assuming that the direction and magnitude of the difference between the model's prediction under no policy and the data is exactly the same as that between the model's prediction under the policy and the prediction that a model that would rationalize the data would make. I find this assumption strong.

the c_{ij} approach, the bounds on each c_{ij} implied by the equilibrium conditions of the model would be a function of the values of the other c_{ij}. This further complicates the process of finding the optimal c_{ij} values. I implemented the process for an earlier version of the analysis, so in principle, it can be implemented. But I do not discuss the further complications of that optimization process here, since we do not run into those problems with the fixed-costs approach.

E The Algorithm for Computing Equilibrium Market-Outcome (G^*, T^*, P^*), and Discussion of Existence and Uniqueness

In this appendix, first I will outline the algorithm that I use to compute the equilibrium market outcome (G^*, T^*, P^*). I will then qualitatively discuss existence and uniqueness of the equilibria for the algorithm.

E.1 The Computational Algorithm for (G^*, T^*, P^*)

The computational algorithm works by starting from and initial market outcome (G, T, P) and iteratively updating it using an updating procedure which modifies (G, T, P) if and only if it does not satisfy all of the equilibrium conditions outlined in the model section of the main text. I keep updating (G, T, P) until it satisfies all the equilibrium conditions and, hence, is not modified by further updates. At this point, the algorithm stops and its most recently updated market outcome (G, T, P) is output as the equilibrium market outcome (G^*, T^*, P^*). In the rest of this section, I first explain a procedure called LINK_UPDATE which is the cornerstone of my updating algorithm. Roughly, the LINK_UPDATE procedure takes an input market outcome $\left(G^{Input}, T^{Input}, P^{Input}\right)$ plus a pair (s, j) representing hospital system s and insurer j, and outputs $\left(G^{Output}, T^{Output}, P^{Output}\right)$ which is an updating of $\left(G^{Input}, T^{Input}, P^{Input}\right)$ *with respect to the sj link* in a sense that I will detail in this section. Once I lay out how LINK_UPDATE works, I explain how it's used as the building block of a procedure called MARKET-OUTCOME-UPDATE which updates the market outcome completely, not just with respect to a particular link.

E.1.1 The LINK_UPDATE Procedure

Every time the procedure LINK_UPDATE is applied to input information $\left(G^{Input}, T^{Input}, P^{Input}\right)$ and (s, j), the following algorithm is run to produce output $\left(G^{Output}, T^{Output}, P^{Output}\right)$:

Step 1: Check if there are gains from trade (GFT) for the pair sj under $\left(G^{Input}, T^{Input}, P^{Input}\right)$. That is, fix $G^{Input}_{-sj}, T^{Input}_{-sj}, P^{Input}_{-j}$ and check if there is any reimbursement t_{sj} that hospital system s and insurer j can contract on and both get higher profits than if they do not have a contract. In making these profit comparisons, assume that insurer j optimally responds to any reimbursement rate t_{sj} (or lack thereof). If there is no such t_{sj} then there is no GFT. In this case, set $\left(G^{Output}_{-sj}, T^{Output}_{-sj}, P^{Output}_{-j}\right) = \left(G^{Input}_{-sj}, T^{Input}_{-sj}, P^{Input}_{-j}\right)$. Also set g^{Output}_{sj} to 0, and t^{Output}_{sj} to $\emptyset$. Then set p^{Output}_j to insurer j's optimal premium response to $\left(G^{Output}, T^{Output}, P^{Output}_{-j}\right)$. Then exit the LINK_UPDATE procedure.

In sum, what step 1 does is it checks if link sj can be sustained. If not, it drops the link, adjusts insurer j's premium accordingly, and exits.

Step 2: (Note: If the algorithm enters this step, it means there were GFT in step 1.) Set $\left(G^{Interim}_{-sj}, T^{Interim}_{-sj}, P^{Interim}_{-j}\right) = \left(G^{Input}_{-sj}, T^{Input}_{-sj}, P^{Input}_{-j}\right)$. Set $g^{Interim}_{sj}$ to 1, and . Then, I follow the bargaining formulation developed in the main text. As a reminder, the formulation was:

$$t^*_{ij} = \min\left(t^{NB}_{ij}\left(G, T, \gamma_{ij}\right), \hat{t}_{u_i}\left(G, T, d_j\right)\right) \tag{E.1}$$

Following (E.1), I set

$$t^{Interim}_{sj} = \min\left(t^{NB}_{sj}(G^{Interim}, T^{Interim}, \gamma_{sj}), \hat{t}_{u_s}(G^{Interim}, T^{Interim}, d_j)\right) \tag{E.2}$$

where in computing both $\hat{t}_{u_s}$ and t^{NB}_{sj}, the optimal reaction of p_j is anticipated. Finally, set $p^{Interim}_j$ to insurer j's optimal response to $\left(G^{Interim}, T^{Interim}, P^{Interim}_{-j}\right)$.

In sum, step 2 sets the reimbursement rate for link sj to the Nash Bargaining price or the Best Safe Price, whichever is smaller, according to the model outlined in the model section of the main text. It then adjusts insurer j's premium accordingly. Next steps make sure that, in order to deter replacement threats, hospital system s does not offer a reimbursement so low that it actually makes losses on the sj link.

Step 3: Construct another market outcome $\left(G^{Interim2}, T^{Interim2}, P^{Interim2}\right)$ by setting $\left(G^{Interim2}_{-sj}, T^{Interim2}_{-sj}, P^{Interim2}_{-j}\right) = \left(G^{Interim}_{-sj}, T^{Interim}_{-sj}, P^{Interim}_{-j}\right)$, setting $g^{Interim2}_{sj}$ to 0, and $t^{Interim2}_{sj}$ to $\emptyset$, and finally setting $p^{Interim2}_{j}$ to insurer j's optimal response to $\left(G^{Interim2}, T^{Interim2}, P^{Interim2}_{-j}\right)$. If

$$\pi_{u_s}\left(G^{Interim}, T^{Interim}, P^{Interim}\right) \geq \pi_{u_s}\left(G^{Interim2}, T^{Interim2}, P^{Interim2}\right) \tag{E.3}$$

then set $\left(G^{Output}, T^{Output}, P^{Output}\right) = \left(G^{Interim}, T^{Interim}, P^{Interim}\right)$ and exit.

In sum, step 3 checks whether the reimburse rate $t^{Interim}_{sj}$ as specified in (E.2) is too small a rate to be sustainable. Observe that according to definition of best safe reimbursement rates, and equation (E.2), the rate $t^{Interim}_{sj}$ has to be small enough to disincentivize insurer j from replacing system s with any currently out of network system s'. But a reimbursement rate this small might also be small enough so that hospital system s would rather drop the sj link than keep such a low rate (e.g. when $t^{Interim}_{sj}$ falls well below the marginal cost of hospital system s). This would violate network stability conditions as defined in the main text. That's why step 3 checks for this condition. If $t^{Interim}_{sj}$ is still high enough not to violate network stability conditions, step 3 approves of it and exits the algorithm. Otherwise, hospital system s is not capable of deterring replacement threats. Therefore, the algorithm moves on to step 4 to actually allow a replacement to happen.

Note that if the algorithm enters step 4, it means that condition E.3 in step 3 did not hold. This can only be the case when:

$$t^{Interim}_{sj} = \hat{t}_{u_s}(G^{Interim}, T^{Interim}, d_j)$$

That is, $t^{Interim}_{sj}$ is the highest safe price for system s when dealing with insurer j. Therefore, under $t^{Interim}_{sj}$, insurer j is just indifferent between keeping the contract with system s and replacing system s with some other system s'. Formally, there is a system s' and reimbursement rate $t_{s'j}$ such that

$$\pi_{d_j}\left(G^{Interim}_{-sj+s'j}, T^{Interim}_{-sj+t_{s'j}}\right) = \pi_{d_j}\left(G^{Interim}, T^{Interim}\right) \tag{E.4}$$

$$\pi_{u_{s'}}\left(G^{Interim}_{-sj+s'j}, T^{Interim}_{-sj+t_{s'j}}\right) \geq \pi_{u_{s'}}\left(G^{Interim}, T^{Interim}\right) \tag{E.5}$$

$$\pi_{u_{s'}}\left(G^{Interim}_{-sj+s'j}, T^{Interim}_{-sj+t_{s'j}}\right) \geq \pi_{u_{s'}}\left(G^{Interim}_{-sj}, T^{Interim}_{-sj}\right) \tag{E.6}$$

Step 4 simply implements the above replacement.

Step 4: Set $\left(G_{-s'j}^{Output}, T_{-s'j}^{Output}, P_{-j}^{Output}\right) = \left(G_{-s'j}^{Interim2}, T_{-s'j}^{Interim2}, P_{-j}^{Interim2}\right)$. Then set $g_{s'j}^{Output}$ to 1, and $t_{s'j}^{Output}$ to $t_{s'j}$, and finally set p_j^{Output} to insurer j's optimal response to $\left(G^{Output}, T^{Output}, P_{-j}^{Output}\right)$.

Step 4 was the last step of the LINK_UPDATE procedure. My next procedure, MARKET-OUTCOME-UPDATE, simply applies LINK_UPDATE multiple times.

E.1.2 The MARKET_OUTCOME_UPDATE procedure

This procedure starts with $\left(G^{Input}, T^{Input}, P^{Input}\right)$ and outputs $\left(G^{Output}, T^{Output}, P^{Output}\right)$. It does this job in the following steps:

Step 1: Produce a random ordering of all possible links sj.

Step 2: Set $\left(G^{Interim_0}, T^{Interim_0}, P^{Interim_0}\right) = \left(G^{Input}, T^{Input}, P^{Input}\right)$. Then for each $index \in \{1, ..., S \times n\}$, set

$$\left(G^{Interim_{index}}, T^{Interim_{index}}, P^{Interim_{index}}\right) =$$

$$LINKUPDATE\left(G^{Interim_{index-1}}, T^{Interim_{index-1}}, P^{Interim_{index-1}}, (s,j)_{index}\right)$$

where $(s,j)_{index}$ means the $index$-th link according to the random ordering generated in step 1 of the algorithm.

Step 3: Finally, set $\left(G^{Output}, T^{Output}, P^{Output}\right) = \left(G^{Interim_{S\times n}}, T^{Interim_{S\times n}}, P^{Interim_{S\times n}}\right)$.

The reason why I have a randomization over the ordering of the links to update in step 1 is that this, along with using multiple starting points, is a good way to check for multiple equilibria.

It is straightforward to verify that a market outcome that does not change if MARKET_OUTCOME_UPDATE is applied to it satisfies all the equilibrium conditions (i.e., the conditions in the definitions of pairwise stability as well as bargaining with threats of replacement in the main text).

E.2 Existence

As mentioned before, In this section I will not give formal results on existence. Instead, I briefly discuss some of the most important issues regarding it.

1. Capturing replacement threats helps existence: My proposed bargaining formulation does much better than NiN bargaining when it comes to equilibrium existence. Using NiN for bargaining as part of a network-formation and bargaining model sometimes leads to non-converging cycles in the algorithm for finding the equilibrium market outcome. To illustrate, the following cycle in the algorithm is possible when using NiN: hospital system u_s is in the network of insurer d_j with some t_{sj} as the Nash Bargaining outcome. Reimbursement rate t_{sj} is high enough so that there is gain for insurer d_j to also sign a contract with hospital $u_{s'}$ and steer some of its patients from u_s to $u_{s'}$. But once the $s'j$ link is added to the network, insurer d_j's outside option profit in bargaining with u_s increases because now d_j has $u_{s'}$ in its network and losing u_s would not have as big of an effect on d_j as it did before. This enables d_j to negotiate a lower reimbursement rate with u_s. This lower reimbursement rate is low enough so that d_j loses the incentive to keep its contract with $u_{s'}$. So, the $s'j$ link severs. But this compromises d_j's outside option profit in bargaining with u_s, bringing the sj negotiated reimbursement back up. But then, the sj reimbursement rate is now again high enough so that d_j will have the incentive to sign a contract with $u_{s'}$, and so forth.

The type of cycle described above happens with NiN bargaining often in general and *always* when I simulate network adequacy regulations. With such regulations, when insurer d_j is just abiding by the regulation (e.g. when it covers exactly $X\%$ of the hospitals where X is the mandated minimum percentage), NiN would imply very high negotiated reimbursement rates due to the fact that the insurer is not allowed to drop any of the hospitals. These extremely high rates incentivize the insurer to bring in at least another hospital. But then the regulation is not binding anymore and none of the hospitals that used to charge high sums due to the bind of the regulation are able to keep doing that. Therefore all those reimbursements fall and take back away the insurer's incentive to have the new additional hospitals in its network. But once those hospitals are out of network, the regulation becomes binding again, leading to very high reimbursement rates, and so forth.

2. The importance of no-commitment conditions for existence: The no-commitment condition in the definition of best safe reimbursement rates is important for existence when there are *multiple insurers*. To illustrate, that condition makes sure the following cycle does not happen in the algorithm that finds the equilibrium (G^*, T^*, P^*): hospital system $u_{s'}$ is willing to replace system u_s in insurer d_j's hospital network and charge a low price. The reason for $u_{s'}$'s willingness is that by taking u_s out of d_j's network, d_j will become less popular among customers and some of d_j's enrollees will now switch to $d_{j'}$ which also covers $u_{s'}$ and pays a high reimbursement to it. But once the replacement takes place, $u_{s'}$ would profit from exiting d_j's network because of the low rare it's charging. Once that happens, d_j brings u_s back. Then $u_{s'}$ is again willing to substitute u_s for a low price, and so on and so forth. The no-commitment condition prevents this by imposing the extra restriction that the d_j will not replace $u_{s'}$ for u_s if u_s is willing to participate only at rates that it will find unacceptable after the replacement takes place.

3. Existence and horizontal differentiation: When there is horizontal differentiation among hospitals (e.g. heterogeneity in locations), there could be a potential non-existence issue, which happened (though not frequently) in my simulations. To illustrate, suppose hospital systems u_s and $u_{s'}$ are located in two different locations. Also suppose there are two other hospital systems u_{s_1} and $u_{s'_1}$ which are located near u_s and $u_{s'}$ respectively. The following cycle can happen through which u_{s_1} and $u_{s'_1}$ are always covered by an insurer d_j but u_s and $u_{s'}$ keep replacing each other: u_s is in network and $u_{s'}$ isnt. Here, u_s's presence in the market reduces u_{s_1}'s rate due to the fact that d_j's outside option profit in the Nash Bargaining with u_{s_1} increases (because if u_{s_1} drops out of d_j's network, some of its patients go to u_s). But the low rate charged by u_{s_1}, in turn, makes the presence of u_s less necessary. Therefore, insurer d_j replaces u_s by $u_{s'}$. After that, the price by u_{s_1} goes back up due to the fact that now, if d_j drops $u_{s_1}1$, patients in that geographical area may drop the insurer rather than spillover to u_s. Also, due to presence of $u_{s'}$ in the market, the rate charged by geographically nearby $u_{s'_1}$ goes down, making the presence $u_{s'}$ a little less necessary. These two, together, can lead the insurer to replace u_s back for $u_{s'}$ in the network, and the cycle repeats from here on.

Two natural questions arise about this situation. First, the main purpose of adding replacement threats to the NiN bargaining solution was that it captured the competition that, say, u_{s_1} faces from the neighboring u_s even when u_s is out of d_j's network. So, why is it that when u_s is left out, u_{s_1} gets to charge a higher rate to d_j than when u_s is in-network, which leads, in some cases, to the cycle described in the previous paragraph? The answer is in the particular formulation of replacement threats in the model as constructed in equation (E.1). This formulation assumes that when u_s is out of network, u_{s_1} will charge the *highest* rate at which d_j would not want to swap u_s and u_{s_1}, or even a higher rate. Therefore, even though equation (E.1) captures the competition that u_{s_1} faces from u_s when u_s is out-of-network, this competition exists to a lesser extent than it does when u_s is in-network. This, then leads to the cyclical mechanism mentioned above. Note that there would be possible formulations of replacement threats than (E.1). However, I decided for (E.1) and against those other formulations for two reasons. First, the non-existence cases described above do not happen very often. Second, and more importantly, equation (E.1) has the non-cooperative support detailed in section A but for no other formulation did I succeed in finding any non-cooperative support even for 2×1 games.

The second natural question is whether there is a workaround in these cases so that the algorithm converges? What I do in these cases, where d_j oscillates by replacing u_s and $u_{s'}$ for one another, is that I make the algorithm choose the network that gives a higher profit to the insurer. I assume that the hospital system that is in that profitable network (say it is $u_{s'}$) charges its "lowest acceptable rate" to the insurer, which is the lowest rate under which $u_{s'}$ would not want to leave the network. Under this rate, of course, the insurer would like to replace $u_{s'}$ by u_s but I assume this does not happen to prevent the cycle from happening. In all of my simulations that led to such cycles, this approach assured convergence. Also, in all cases, the two networks between which the algorithm oscillates are very similar to one another in terms of the profits to the firms and the consumer welfare implications. So, a simple algorithm for choosing between the two, like the one I implement here, would be a reasonable approach to deal with the issue of oscillations in the algorithm that searches for the equilibrium market outcome.

E.3 Uniqueness

Like with existence, I do not have formal proofs regarding uniqueness. Nevertheless, in my simulations the equilibrium has always been unique. I conjecture that in situations where non-unique equilibria are natural, my model can generate all of the equilibria (e.g. a product differentiation situation where one insurer covers only the cheap, lower-quality, hospitals to be able to charge a low premium and appeal to more price-sensitive consumers, and another insurer covers all hospitals and charges a higher premium and appeal to more quality-sensitive consumers.) Nonetheless, I have not come across this type of situation in the analysis of CommCare. In this section, I qualitatively discuss features of the model that, based on simulations, I have concluded prevent the existence of multiple equilibria in ways that are not as economically natural as, say, product differentiation.

1. Capturing replacement threats helps uniqueness: NiN bargaining could lead to non-uniqueness of equilibria. To illustrate, suppose there is one insurer d and two hospitals u_1 and u_2. Suppose that if the insurer does not have any hospital in its network, then signing a contract with hospital u_1 will give gains from trade to both the hospital and the insurer. Also suppose that if the insurer is contracting only with u_1 and paying the corresponding Nash Bargaining price, it does not have an incentive to bring in also u_2, due to the fact that there are fixed costs of bringing in u_2 and that u_1 already being in the network gives less marginal value to u_2. In addition, suppose similar conditions hold between insurer d and hospital u_2. That is, first they do have an incentive to sign a contract if the insurer is not already covering u_1, and second, once they are signing their Nash Bargaining contract, the insurer does not have an incentive to sign with u_1. Under these circumstances, a model of network formation with NiN bargaining would imply that networks $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ could both be sustained as part of equilibria. Importantly, even if the insurer prefers the Nash Bargaining contract with hospital u_1 to that with hospital u_2, a contract with hospital u_2 could still be an equilibrium since the only ways the insurer can deviate from this equilibrium is either by dropping the contract or by adding u_1 to the network, none of which the insurer would like to do. This issue is not there in my proposed bargaining formulation. In my model, if network $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ can be sustained as part of the equilibrium market outcome, it means that hospital u_1 is able to reduce its price by enough so that even the lowest price from u_2 will not incentivize insurer d to replace u_2 for u_1. This directly implies that if the current network is $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$, hospital u_2 cannot offer a low enough price that would beat hospital u_1's lowest price, and hence, insurer d will inevitably substitute u_1 for u_2, alleviating the issue of multiple equilibria.

2. Capturing sequential premium setting helps uniqueness: A second thing that could lead to multiple equilibria is assuming that premium setting (i.e., step 2 in the four-stage game set up in the main text) happens simultaneously to network formation and bargaining (i.e., step 2 in the four-stage game set up in the main text). That is, when bargaining, firms take the premiums as fixed and do not anticipate them to change in response to the outcome of the bargaining. This assumption has been used in at least two papers on upstream-downstream bargaining in vertical markets (Ho and Lee [2017b], Crawford et al. [2015]) due to the substantial computational simplicity that it provides. Those papers do not fully endogenize network formation. But when network formation is also endogenous, this assumption on premium setting could undermine uniqueness. To illustrate, suppose there are two hospitals u_1 and u_2 and one insurer d. Hospital u_1 is a prestigious hospital with very high marginal cost and hospital u_2 is the opposite: low quality but also low cost. Now consider two market outcomes. In the first one, the insurer covers hospital u_1, pays a high reimbursement to u_1 due to u_1's high marginal cost and quality, and charges a high premium accordingly. In this situation, hospital u_2 could, in principle, provide a strong replacement threat for insurer d to use against hospital u_1 when bargaining. Since hospital u_2 is cheap, replacing it for hospital u_1 could allow the insurer to lower its premium and appeal to more consumers. Nevertheless, a model that assumes premiums are set simultaneously with the bargaining would only allow for the insurer to substitute hospital u_2 for u_1 without changing the premium, making the replacement threat much weaker. A similar issue exists with the reverse situation: Consider a market outcome where the insurer covers hospital u_2, pays a low reimbursement to u_2 due to u_2's low marginal cost and quality, and charges a low premium accordingly. In this situation, hospital u_1 could provide a strong replacement threat for insurer d to use against hospital u_2. Since hospital u_1 is prestigious, replacing it for hospital u_2 could make the insurance plan more attractive. But the insurer might only want to do this if it can raise its premium to cover the higher reimbursement that it has to pay the expensive u_1. Nevertheless, a model that assumes premiums are set simultaneously with the bargaining would only allow for the insurer to substitute hospital u_1 for u_2 without changing the premium, making the replacement threat much weaker. So both market outcomes could possibly be sustained as equilibria. We do not face this issue if the premium setting happens *after* network formation and bargaining. In this latter case, only one of the two above mentioned market outcomes is sustainable under my network stability and bargaining conditions. If the insurer prefers covering u_1 and charging a high premium over covering u_2 and charging a low premium, then u_1 can deter a threat of replacement by u_2 but u_2 cannot do the same. The reverse is true if the insurer prefers covering u_2 and charging a low premium over covering u_1 and charging a high premium. Thus, we get only one equilibrium.

F Exogeneity of Premium Variation in Plan Demand Estimation

In this appendix, I argue that the variation of monthly plan premiums in my plan demand estimation procedure is exogenous to the variation in demand from those income groups for which I seek to estimate premium coefficients (i.e. income groups of between 100% and 200% of poverty and between 200% and 300% of poverty). I do this by arguing that plans' strategies regarding the below poverty income group -who always paid zero premium and for whom I do not estimate premium a coefficient- drove all the premium variation in the data that I use. I first detail the premium variation in my data and then discuss the institutional details behind the variation.

F.1 Premium Variation in the Data

Before I use my enrollments dataset (described in the data section of the main text) to estimate the plan demand model (i.e., step 2 in the four-stage estimation procedure set up

Fiscal year	BMC	Celticare	Fallon	NHP	Network Health
2012	447	360	429	415	360
2013	346	377	415	415	355

Table 6: Premiums of all CommCare plans, in \$/month/person, in fiscal years of 2012 and 2013

in the main text), I restrict it in two ways. First, I restrict the dataset to fiscal years of 2012 and 2013 as, unfortunately, my data on FY2011 and before has missing records. Second, I restrict to consumers only from the two above poverty income groups, and drop the below poverty ones. The reason for this restriction is the particular policy adopted by CommCare towards this specific income group. I will detail that policy in section F.2. Given that in my model I assume all income groups have the same brand preference and the care to the same extent about hospital networks when choosing insurance plans, dropping this group will not undermine the identification of those coefficients.

CommCare premiums from 2011 on were "community rated." That is, each of the insurers charged the same pre-subsidy premium to all CommCare consumers, regardless of demographics. Table 6 shows these pre-subsidy premiums set by all CommCare plans in fiscal years of 2012 and 2013.

The premium paid by each consumer depends only on the premiums reported in table 6 and subsidy parameters described in the followingi formulation copied from the main text:

$$p_{jk}^{sub} = a_{y(k)} \times p_j - b_{y(k)} \tag{F.1}$$

The pass-through rate subsidy parameter a_y for each income groups 100%-200% of poverty and 200%-300% of poverty were approximately 50% and 90% for both 2012 and 2013 fiscal years.[67] Table 7 shows the after-subsidy premium changes from FY2012 to

[67]Note that as mentioned before, the income-grouop specific subsidies are based on a finer categorization than the one I use in this paper. 100%-150% of poverty and 150%-200% of poverty groups have different subsidy rates. Same with 200%-250% of poverty and 250%-300% of poverty groups. But since in my datasets, I only see the coarser categorization, I simply assume that the pass through rate for the bigger group of 100%-200% of poverty is the average of the rates for 100%-150% and 150%-200% and make a similar assumption for the 200%-300% of poverty group.

Income Group	BMC	Celticare	Fallon	NHP	Network Health
100%-200% of poverty	-50	8.41	-6.93	0	-2.48
200%-300% of poverty	-91.41	15.38	-12.67	0	-4.53

Table 7: Changes in CommCare plans' premiums (after accounting for subsidies) from FY2012 to FY2013.

FY2013 for all CommCare plans.

F.2 Role of the Below Poverty Income-Group in the Premium Variation

Since FY2012, CommCare mandated that, every year, those below poverty CommCare consumers that did not hold a CommCare plan in the past year, had to choose a plan only between the two cheapest plans offered in the market, as long as those two plans each had a pre-subsidy premium of 380$/month/person or less.[68] This rule, Celticare and Network Health lowered their premiums from 404$/month/person and 425$/month/person respectively in FY2011 to 360$/month/person in FY2012. The two became the cheapest two plans on CommCare for 2012 and enjoyed duopoly for serving the new below poverty enrollees.[69] Below, I argue that almost all of the premium variation between FY2012 and FY20133, as shown in table 7, comes from the insurers' reactions to this particular regulation.

BMC's premium change of 101$ accounts for a large portion of the total premium variation. The institutional setting detailed in the previous paragraph strongly suggests that BMC reduced its premium by this large amount in order to get access to the new below poverty enrollees, which it successfully did. Also Celticare's premium increased from 360$/month/person to 377$/month/person, just below the 380$ cap described above. This strongly suggests that in FY2012 Celticare fiercely competed for the two spots for serving the below poverty group; and when it succeeded, it increased its price to 377$/month/person,

[68] If only one plan had a premium no higher than 380$/month/person, that plan would be the only choice for new below-poverty enrollees. If no plan satisfied this premium requirement, there would be another open enrollment preiod for the new below poverty enrollees. None of these two situations took place in CommCare however.

[69] In order to be able to go cheap, Network Health dropped some prestigeous hospitals out of its network.

the highest level under which it could still serve the new below poverty enrollees, only to get surprised by the unanticipated large reduction in BMC's premium. These changes in BMC's and Celticare's premiums, hence, seem to have arisen from their strategies regarding the below poverty group, and not affected by unobservable changes in the demand for insurance plans from the other two groups.

Of course beside the premium changes for BMC and Celticare, there are other premium changes as well between FY2012 and FY2013. Nevertheless, they do not account for as large variations in premiums as the chances for Celticare, and in particular BMC do. Fallon reduced its premium by 14 dollars and this reduction cannot be attributed to the below poverty group. But Fallon's market share in CommCare was around 3% and this change is not likely to largely affect the estimated premium coefficients. Also Network Health reduced its premium going into FY2013. Nevertheless, this premium changes was only by 5$/month/person, which is not comparable to the much more substantial premium change from slightly more popular BMC. Therefore, I argue that a very large portion of the variation in premiums was exogenous to the demand functions of the two income groups y for which I estimate price coefficients β_y.

G Constructing measures of hospital marginal costs c_i and reimbursement rates t^*_{ij}

In the main text, I treated both hospitals' marginal costs c_i of inpatient care and the reimbursements matrix $T^* = \left[t^*_{ij}\right]_{m\times n}$ as data. These quantities, however, are not directly observed in my data. In this appendix, I detail how I construct measures of these quantities using hospitals' cost reports data and medical claims data.

G.1 Measuring hospital marginal costs c_i

In my model, c_i is the *severity adjusted inpatient marginal cost* to hospital u_i. In order to construct a measure of c_i from HCRIS data, I follow these four steps:

Step 1: I construct a measure of *total variable cost* to each hospital. To do that, I sum over all of the costs items in the hospitals cost report for 2011, except for those that I believe are fixed costs. More specifically, I include all cost items except for all capital related costs, employee benefits, operation of plant, nursing school costs, all costs regarding interns and residents, and all non-reimbursable costs.

Step 2: I construct a measure of *total variable inpatient cost* to hospital u_i. This cost measure is a fraction of *total variable cost* determined by how large of a fraction of the total revenue to the hospital is from inpatient care:

$$\textit{total variable inpatient cost} = \textit{total variable cost} \times \frac{\textit{total inpatient revenue}}{\textit{total revenue}} \tag{G.1}$$

To illustrate, if 40% of a hospital's revenue comes from inpatient care, then also 40% of the cost must have come from costs to provide inpatient care. Of course, this approach is non-ideal since it implicitly assumes that the hospital's profit margins from inpatient and outpatient care are equal. But it is a good approximation and is used in other work (e.g. see Schmitt [2015]).

Step 3: I construct a measure of *inpatient marginal cost* to hospital u_i simply by dividing the *total variable inpatient cost* by total discharges.

Step 4: I construct the *severity adjusted inpatient marginal cost* c_i by dividing the *inpatient marginal cost* of hospital u_i by hospital u_i's Case Mix Index (which is a measure of average severity of the inpatient admissions by the hospital constructed by CMS).

G.2 Measuring reimbursement rates t^*_{ij}

Reimbursement rate t^*_{ij} is the average payment by insurer d_j to hospital u_i for an inpatient admission an enrollee of d_j at hospital u_i. I estimate each t^*_{ij} off of MA-APCD's medical claims file by simply dividing the total payments from insurer d_j to hospital u_i for inpatient care in the fiscal year of 2011 by the total number of discharges of d_j enrollees and u_i in that year. The following notes about these rate estimates are worth mentioning.

First, I am not taking MS-DRG severity indices into account in constructing the t^*_{ij} measures. Other papers, (e.g. Gowrisankaran et al. [2013], Ho and Lee [2017b]) assume payments are proportional to DRG severity scores. I examined this assumption for CommCare by taking the following regression:

$$\ln(t_a) = FE_{ij,year} + \rho \times \ln(\psi_a) + \varrho_a \tag{G.2}$$

In this regression, a represents an inpatient admission event. t_a represents the payment from the corresponding insurer to the corresponding hospital for that admission. $FE_{ij,year}$ is a hospital-insurer-year fixed effect. DRG severity score corresponding the principal diagnosis for admission is denoted ψ_a. Finally, ϱ_a is an idiosyncratic error term. In this regression, coefficient ρ represents the effect of severity score on the payment. If the payment is, on average, proportional to the severity score, we should expect an estimated $\hat{\rho}$ close to 1 and significant. However, I estimate $\hat{\rho}$ at 0.028 and insignificant at the 5% level. It remains insignificant when I also include the age of the patient as another control variable, and its magnitude remains almost the same. I also anecdotally noticed that there was at least one hospital-insurer pair for which t_a had the same value for all admissions. Thus, I decided against incorporating ψ_a into the modeling of the reimbursement rates.

The second note about the t^*_{ij} measures is that, unfortunately, I am missing data on Network Health claims from 2012 and before. Therefore, I am not able to estimate t^*_{ij} for $j =$ "Network Health". I instead, calibrate those rates from the outcomes of the estimation process in Shepard [2015] which examines the same market. Shepard [2015] reports average reimbursement rates for hospitals (across all insurers). Given that I estimate the rates for all insurers but Network Health, I use his numbers (in Table 1, Panel A) to back out what the average rate for Network Health must have been. Also, given that Shepard [2015] reports the rates for only 10 of the most expensive hospitals, I cannot use this approach for the rest of the hospitals. For them, I instead use my estimated Network Health rates for 2013. I ran the estimation procedure for a range of "reasonable" alternative calibrations and the results are not sensitive to that.

Finally, before using the t^*_{ij} measures in the model for estimating cost functions and running counterfactual simulations, I carry out another further processing. Some of my t^*_{ij} measures happen to fall below my measures of hospital marginal costs c_i. This will be inconsistent with the equilibrium conditions of my model, given that $t^*_{ij} < c_i$ would imply an incentive for the hospital u_i to break off the ij link, unless this hospital is part of a bigger hospital system s which includes other hospitals that charge high enough to d_j to cover the loss from $t^*_{ij} < c_i$. To avoid this and help my model rationalize the date as an equilibrium, I bump up the t^*_{ij} measures in cases where it leads to the aforementioned problems to $c_i + 300$, measures in \$/admission. This procedure modifies 15 of the t^*_{ij} measures out of a total of 112. It increases the average reimbursement rate paid to hospitals by BMC, Celticare, NHP, and Network Health, respectively by 1%, 0%, 2%, and 1%. The $t^*_{ij} < c_i$ could be results of measurement errors in estimating t^*_{ij} and/or c_i potentially both by the econometrician and the hospitals.

H Consumer risk aversion and bargaining leverage for star hospitals

In this appendix, I describe a hypothesis about the effects of consumers' risk aversion in how they evaluate health plans against each other, and I argue that, not capturing this risk aversion, standard models of insurance plan demand (developed by Capps et al. [2003], further completed by Ho [2006], and then widely used in the literature) might miss an important source of bargaining leverage for star hospitals in negotiating rates with insurers. Therefore, if such demand models are used in conjunction with a supply model (such as the model I develop in this paper) to estimate bargaining powers, they may over estimate the bargaining parameters of star hospitals $\hat{\gamma}_{star}$, as the bargaining leverage from risk aversion is being picked up by $\hat{\gamma}_{star}$. Formalizing and empirically testing this hypothesis would be a separate paper. In this appendix, I only describe this hypothesis and discuss its implication for estimation of bargaining parameters.

Hypothesis: *Due to risk aversion, consumers place additional value on insurance plans that cover star hospitals in their network, to a much greater extent than their likelihood of using those star hospitals would suggest. That is, even those consumers who are not likely to use the star hospitals (e.g. due to geographical distance) would still place a large value on insurance plans that cover them, just in order to have a peace of mind in case of complex but very rare conditions that start hospitals have a clear advantage at treating.*

If the above hypothesis is true, then consumer risk aversion provides a channel through which including a star hospital in network can give an insurer additional enrollees without proportionally giving the star hospital additional patients (due to the fact that those enrollees need the hospital mainly for peace of mind). Therefore, the only way for the star hospital to charge for this "peace of mind effect" would be to load the charge on the reimbursement rates for the patients that the hospital does get, thereby increasing the average reimbursement rate paid by the insurer to the hospital. Of course if this effect is there in the market and we use the demand model that does not capture it, the high reimbursement rates resulting from the peace of mind effect would have to be picked up by another parameter in the model, which in the case of my model is $\hat{\gamma}_{Star}$.