Bargaining on price on behalf of price-insensitive downstream consumers

Guy Arie* Paul L. E. Grieco† Shiran Rachmilevitch‡§

October 10, 2022

Abstract

There are settings in which linear prices are negotiated by procurement agents and final consumption decision made by end users who are indifferent to negotiated prices. For example, a patient seeking medical treatment is indifferent to the treatment's cost, if it is covered by his insurance program. We study bargaining for per-unit prices between suppliers and an intermediary who represents price-insensitive consumers. Under simultaneous bargaining with all suppliers, the resulting prices exceed the value of the good (or service) being delivered, provided that the suppliers have sufficiently large bargaining power. This overpricing is solved if simultaneous negotiations are replaced by sequential ones. The problem with sequential negotiations is that they necessitate treating the suppliers asymmetrically, even if they are symmetric. We utilize the result about sequential negotiations as a building block in a multi-period model that resolves the issue: in this model, overpricing is prevented and all suppliers are treated the same.

Keywords: Bargaining; Nash-in-Nash; Overpricing; Price insensitivity.

*Unaffiliated. Arie was at University of Rochester while working on this project and since moved to Amazon.com. Amazon.com did not review or approve this research or paper.

†Pennsylvania State University: paul.grieco@psu.edu

‡University of Haifa: shiranrach@econ.haifa.ac.il

§We thank colleagues in our institutions, participants in the UECE Lisbon Meetings (2014), IIOC (2015), the FTC 2015 Microeconomics Conference, Drexel University, and in particular Ali Yurukoglu, Henry Mak, Bob Town, Mike Ostrovsky and several anonymous referees.

1 Introduction

In most economic settings, agents care about prices and costs. Consumers care about the prices of the goods they buy, firms about their production costs, and so on. However, there are many settings in which linear prices are negotiated by procurement agents and final consumption decision made by end users who are indifferent to negotiated prices. For example:

- A fully-insured patient seeking medical treatment that is covered by medical insurance—he ignores the treatment's price when taking his decision since it is paid by the insurer;
- A military commander choosing which (or how much) ammunition to use in battle—he is unlikely to consider its costs.[1]
- The end user of a music subscirption service is indifferent to the price the service provider pays for access to any particular music song.

A common feature of these examples is the vertically structured industry in which an intermediary (e.g., an insurance company, Department of Defense, a streaming service) bargains prices with multiple potential suppliers (e.g., hospitals, military contractors), under the assumption that the negotiated prices will have minimal to no effect on the end users' consumption decisions. In turn, final payments are determined by the negotiated price and the quantity consumed.

We study the implications of this scenario on popular models of negotiations. For concreteness, we consider an insurer who bargains with hospitals for per-treatment prices, and whose goal is to maximize the consumers' expected surplus net of prices

[1] An account (in Hebrew) of ammunition-overuse in the IDF can be found in Shelah (2015), pages 44-45.

payed. However, our results apply to any environment which is characterized by the aforementioned vertical structure and price insensitivity.

The dominant approach in the literature models this bargaining enviornment with multilateral simultaneous negotiations that follows an approach that originated in Horn and Wolinsky (1988), where the insurer bargains with each hospital separately, in accordance to the Nash bargaining solution (Nash 1950), taking the prices with all other hospitals as given. Collard-Wexler et al. (2019), in a work to which we will refer in more detail shortly, call it *Nash-in-Nash* (NiN); we follow their terminology. An important distinction between the bargaining environment studied in Collard-Wexler et al. (2019) and our setting is that in thier enviornment bargaining is over a lump sum payment, whereas in our setting bargaining is over linear prices, with quantities determined by downstream consumers. We feel that our setting is commonly observed in the real world, and indeed has been dominant in applied work. For example, in their analysis of bargaing in the multichannel TV market, Crawford and Yurukoglu (2012) assume distributors and content providers bargain over per-subscriber licensing fees. Similarly, analyses of bargainging between hospitals and insurers assume insurers and hospital groups bargain over capitation rates rather than lump sum payments (e.g., Gowrisankaran et al. 2015, Ho and Lee 2017). We will refer to our approach as the *linear NiN* model to distinguish it from the lump sum approach of Collard-Wexler et al. (2019).

In the linear NiN model, the insurer bargains with each hospital, holding the contracts with all other hospitals fixed. The Nash product is formed by comparing the hospital-network's surplus with and without the bargained-with hospital. Consider the case of two hospitals, A and B. Consider bargaining with A, given the negotiated price with B. The network's surplus without A depends not only on B's price but also on A's patients that would substitute to B if A left the network, and their value for

B. Since patients are price-insensitive, they will go to hospital B even if their value for B is lower than B's price. Thus, in the event that A drops out of the network, these patients may generate negative surplus: the utility they get from B—their second choice hospital—is less than the price of B. As a result, A's marginal value per-patient when added to the network (given B's price) is higher than its actual per-patient value, since its addition prevents the aforementioned negative surplus creation. We show that in any network with two or more hospitals, as long as some patients substitute to an in-network hospital when their most preferred hospital is out of the network, prices exceed patient valuations, given that hospitals' bargaining power is sufficiently large. We call this phenomenon *Nash overpricing*. In Theorem 1 we show that if the hospitals' bargaining power is large enough, Nash overpricing will occur for every hospital in the network, which we refer to as *complete Nash overpricing*.

Mathematically, the overpricing in our model is a consequence of super-additivity of mean hosptial valuations, which is itself a consequence of patients' demand being insensitive to negotiated prices. To see this, consider again the two-hospital example, and denote by $v(AB)$, $v(A)$, $v(B)$, and $v(\emptyset)$ the insurer's value (i.e., the total surplus generated) from the four possible hospital networks. It is easy to check that maximizing Nash products implies that the price payed to hospital $j = A, B$ is $t_j^N = \beta[v(AB) - v(j)]$, where $\beta \in (0, 1)$ is the hospitals' bargaining power parameter. Therefore, if β is close to one, the insurer's surplus from the full network, $v(AB) - t_A^N - t_B^N$, is approximately:

$$v(AB) - t_A^N - t_B^N = v(AB) - v(AB) + v(A) - v(AB) + v(B) = \\ v(A) + v(B) - v(AB),$$

and the last term is negative if v is super-additive, which it is in our model. Our first main contribution is not highlighting the significance of super-additivity in applied bargaining settings—this is already known.[2] Instead, it is showing how the combination of price-insensitivity and bargaining over per-unit prices implies super-additivity, and hence implies the overpricing that goes along with it.

NiN is a popular framework for studying bargaining problems, especially in applied settings. Collard-Wexler et al. (2019) provide an analysis of this framework, in a model that generalizes Rubinstein's (1982) alternating offers game. Though the game they study is non-cooperative, the Nash bargaining solution can be applied to their model, and the prices it provides approximate the ones obtained in the non-cooperative equilibrium of their game as the players become infinitely patient; that it, "Rubinstein prices" converge to "Nash prices."[3]

Whereas Collard-Wexler et al. (2019) provide a rather general account of NiN, our model is not a special case of theirs. There are two substantial differences: First, Collard-Wexler et al. (2019) take their primitive payoff functions to be price-independent; second, they take bargaining to be over lump-sum transfers. We, by contrast, consider an objective that depends on per-unit prices crucially. Bargaining over linear (per-unit) prices is the typical assumption in applied work.[4] That per-unit prices play a central role in the emergence of overpricing can be seen by a careful comparison to Collard-Wexler et al. (2019): their Lemma 2.2 establishes that in their environment (where bargaining is over lump-sum transfers), Nash prices are always below the value of the good being delivered, and this is so independent of any super- or sub-additivity properties.[5]

[2]See Bloch and Jackson (2007).
[3]This is a general version of the known "Rubinstein-to-Nash convergence" (Binmore et al. 1986).
[4]See, e.g., Ho and Lee (2017).
[5]Super-additivity ensures equilibrium existence in their non-cooperative game.

Our next contribution is to establish that overpricing is a consequence of the simultaneous negotiations framework. If negotiations are carried out sequentially rather than simultaneously, overpricing is prevented. With J hospitals ordered in a sequence, given that prices with hospitals $\{1, \cdots, J-1\}$ have been determined, the insurer faces a standard 2-player bargaining problem with hospital J, in which his payoff is positive. In the negotiation with hospital $J-1$ this is taken into account, and so negotiations with hospital $J-1$ are also a standard 2-player bargaining problem in which the insurer's payoff is positive, and so on.

Under sequential negotiations the insurer's payoff is independent of the order of negotiations. The reason is that the prices that are obtained by maximizing Nash products internalize the effects of earlier negotiations on later ones, and this internalization turns out to be perfect: switching from any negotiations order to any other results in price-adjustments that exactly off-set the changes induced by price-insensitivity. By contrast, hospitals do care about the order of negotiations, at least under some conditions. For example, this is the case if there are only two hospitals. In this case, the first hospital's price is lower than that of the second, because once there is disagreement with the first hospital and it "drops out" of the game, the second hospital becomes a monopolist; by contrast, there is no such effect for the first hospital when the second hospital drops out. Hospitals' payoffs increase in their bargaining-sequence position also in the case of arbitrarily many symmetric hospitals. Interestingly, the first hospital in the sequence earns a low payoff, given any number of hospitals, and regardless of whether they are symmetric or not. As we explain shortly, this finding turns out to be useful.

The result that the order of negotiations matters to hospitals is a challenge for applied work since in general the precise order of negotiations will be unknown. Therefore, our third contribution is to briefly sketch a model in which overpricing is resolved, but hospitals are treated symmetrically. Specifically, we propose the following "semi-cooperative" multi-period model, in which, in addition to the bargaining-power parameter β, there is one additional parameter—the hospitals' discount factor, δ. The model is such that every period the insurer makes simultaneous price offers to all hospitals, to which they simultaneously respond by acceptance or rejection. If all accept, the offered prices are contracted for the period, and the game moves one period ahead. In equilibrium everybody accepts the offers, so a rejection means a deviation; once there is a rejection by some hospital, say j^D, then the model enters an absorbing phase in which only payoffs are described, but the underlying actions are not (hence the "cooperative" in our "semi-cooperative" terminology). These payoffs are the ones derived from the sequential negotiations model mentioned above, in which the sequence is selected at random, out of the set of sequences in which the deviating hospital (j^D) is placed first.[6] As mentioned above, the "punishment price" that this first-placed hospital obtains is low, hence it deters the hospitals from rejecting their price offers.

We focus on the equilibrium which is best for the insurer. As $(\beta, \delta) \to (1, 1)$, the equilibrium price payed to each hospital converges to the abovementioned "punishment price," and the insurer's payoff converges to a positive number. When $(\beta, \delta) \sim (1, 1)$, the "punishment price" payed to a hospital is approximately that hospital's *standalone value*—what the hospital contributes on average per treated patient, when it is the only hospital in the network. This number is smaller than the hospital's value, because when the hospital is the only available option, it also serves patients for whom it is the second-best choice.

The rest of the paper is organized as follows. Section 1.1 reviews the literature.

[6] As far as we can tell, we are the first to consider a model in which behavior is described explicitly "on the path" but "off the path" only payoffs are specified, and these are derived by means of a cooperative solution concept.

Section 2 lays down the environment, Second 3 considers simultaneous negotiations, Section 4 considers sequential negotiations, Section 5 considers our multi-period semi-cooperative model, and Section 6 concludes with a discussion. The overpricing problem that we identify in the NiN model is derived under the assumption that the pool of hospitals is exogenous. In Appendix A we show that being able to exclude some hospital from the pool (*ex ante exclusion*), as well as excluding some hospitals after contracts with them have already been signed (*ex post exclusion*), do not provide a satisfactory solution to overpricing. Thus, the overpricing problem is robust. In Appendix B we discuss the insurer's outside option, Appendix C provides estimating equations for the multi-period model and Appendix D collects proofs.

1.1 Literature

Our paper belongs to a strand of literature that concerns bilateral bargaining in vertically-structured markets.[7] Horn and Wolinsky (1988) and Collard-Wexler et al. (2019) are central references in this regard.

One of our non-trivial findings is that under sequential negotiations, the insurer's payoff is independent of the hospital order. There are bargaining settings in which the order of negotiations matters (Manea 2018, Xiao 2018), and there are settings in which it does not, at least under some assumptions (Marx and Shaffer 2007, Krasteva and Yildirim 2012). A deeper investigation on order-dependence (or independence) in an environment with price-insensitive consumers is beyond the scope of the present paper.

The bargaining externalities in our paper (the price payed to B influences bargaining with A) make it relate to the literature on contracting with externalities, though

[7] A more general framework is that of bargaining in networks. E.g., Abreu and Manea (2012), De Fontenay and Gans (2013), Stole and Zwiebel (1996).

much of this literature concerns externalities among agents, whereas we focus on the principal who is contracting with them (the insurer).[8]

Finally, though our study is theoretical, our models are inspired by the applied theoretical work from the health economic literature; an important reference in this regard is the handbook chapter by Gaynor and Town (2011). The linear NiN model is also used to estimate bargaining environments by Crawford and Yurukoglu (2012), Gowrisankaran et al. (2015) and Ho and Lee (2017), among others.

2 The environment

An insurer bargains with $J \geq 2$ hospitals on behalf of a mass of heterogeneous patients. Under the full network, which comprises all J hospitals, the quantity of patients treated by hospital j is $q_j > 0$. The expected value for a patient who goes to hospital j (given the full network) is $v_j > 0$.[9] Regardless of what hospitals are in the network, patients always have the (outside) option of not seeking medical treatment, which is associated with the value zero.

If a hospital, say j, is not part of the network, then it is not available for patients to seek treatment. This event only affects those patients who prefer to be treated at j, who then go to their next preferred hospital. The hospital choice of patients that chose hospital $l \neq j$ when j is in the network does not change. The mass of additional consumers for hospital k when hospital j is dropped from the network and those patients' expected value are denoted by $q_{k,-j}$ and $v_{k,-j}$, respectively.

Hospitals can treat patients with a marginal cost $c_j \geq 0$. Therefore, under the

[8] See Galasso (2008) and the references therein.

[9] The expectation is over patients: the patients are heterogeneous, and distinct patients for whom j is the most preferred hospital may value it differently.

full network hospital j's profit is:

$$\pi_j = (p_j - c_j)q_j.$$

If hospital j is out of the network, it receives zero profit. However, if another hospital, say k, is out of the network, j receives more patients and its profits become:

$$\pi_{j,-k} = (p_j - c_j)(q_j + q_{j,-k}) \ .$$

The hospitals are *symmetric* if $\{q_j, v_j, q_{k,-j}, v_{k,-j}, c_j\}$ are independent of k and j. We make the following assumptions:

- (I) For all j: $\sum_{k \neq j} q_{k,-j} > 0$;
- (II) For all distinct j and k: $q_{j,-k} > 0 \Rightarrow v_j > v_{j,-k}$;
- (III) For all j: $c_j < \min\{v_{j,-k} : q_{j,-k} > 0\}$;
- (IV) For all distinct j and k with $q_{k,-j} > 0$: $v_j - c_j > v_{k,-j} - c_k$.

(I) says that for every hospital j, at least some patients have a second-choice-hospital that they prefer over the outside option. (II) requires that patients whose first choice of a hospital is j, on average value hospital j more than patients for whom j is the second choice. (III) says that the surplus from providing service to patients for whom the service provider is the second choice is still a positive surplus. This has two important implications. First, it implies—because of (I) and (II)—that $c_j < v_j$, and so the surplus from the full network is positive. Second, it means that negative surplus for the insurer—the thing around which our paper pretty much revolves—is *only* because of overpricing, not because of providing service by "technologically expensive/inefficient" hospitals. Finally, (IV) is a bound that means that the surplus generated by the first-choice hospital is large enough; specifically, it is greater than the surplus that would have been generated had that first choice been removed from the network and then we looked at what surplus the patients who remain in the network generate in any other hospital. This assumption follow from (II) if all hospitals have the same cost.

In all of our examples the cost, for simplicity, is taken to be zero. Under zero cost, (III) and (IV) follow from (I) and (II).

For expositional clarity, we assume the insurer maximizes patient surplus, net of prices payed. Therefore, the insurer's value from the full network, given a price vector $p = (p_1, \ldots, p_J)$, is:

$$F(p) = \sum_{j=1}^{J} (v_j - p_j) q_j. \tag{1}$$

This assumption abstracts away from the insurer adjusting downstream contacts between it and patients as a result of disagreements (such as making changes to premiums) as these are not the focus of our study. It corresponds to two plausible scenarios. First, the insurer will maximize patient valuation if it is acting as an agent for consumers seeking, as one might suppose of a self-insuring employer offering medical insurance as a benefit to employees. Second, an insurer will do so if it has complete market power over consumers and is able to fully extract the surplus.

Similarly, the insurer's surplus from the network without hospital j, given the remaining hospital prices, is:

$$F_{-j}(p) = \sum_{k \neq j} \left[(v_k - p_k) q_k + (v_{k,-j} - p_k) q_{k,-j} \right]. \tag{2}$$

We study several bargaining protocols between the insurer and the hospitals, to be specified in the next sections; the (asymmetric) Nash bargaining solution is common to them, which justifies the following terminology: we say that *Nash overpricing* occurs if there is some hospital j whose price exceeds its value, i.e., $p_j > v_j$. If this is true for every j, then we say that there is *complete Nash overpricing.*

3 Simultaneous negotiations

We start by considering the case where prices between the insurer and each hospital are set following the Nash bargaining solution, holding all other prices fixed. The hospitals' bargaining power parameter is $\beta \in (0,1)$. We refer to this model as the *linear NiN model*, and to its prices as *NiN prices.*

The NiN prices $(p_1^N, \cdots, p_J^N)$ satisfy:

$$p_j^N = \max_{p_j} [F(p_j, p_{-j}^N) - F_{-j}(p_{-j}^N)]^{(1-\beta)} \cdot [q_j(p_j - c_j)]^{\beta}.$$

Maximization of the Nash product gives:

$$p_j^N = \beta[v_j - \frac{\sum_{k \neq j}(v_{k,-j} - p_k^N)q_{k,-j}}{q_j}] + (1-\beta)c_j. \tag{3}$$

A solution to this system of equations is called an *equilibrium.*

Theorem 1. *In the NiN model, an equilibrium exists, and it is unique. There exists a $\bar{\beta} < 1$, such that if the hospitals' bargaining power parameter satisfies $\beta \in (\bar{\beta}, 1)$, then each of the NiN prices exceeds the value of service in the corresponding hospital. That is,*

$$p_j^N > v_j \quad \forall j = 1, \cdots, J.$$

Namely, there is complete Nash overpricing.

Proof. We start by establishing existence and uniqueness. Note that we may assume that prices never exceed some (possibly large) bound M: clearly, no hospital can obtain a price that exceeds its value plus all the "adverse selection prevention" that its addition to the network can bring about. Then the RHS of (3) describes a map from $[0, M]^J$ into itself. Though this is a map of vector-to-vector, it can be viewed as an operator on functions because a vector is a constant function. It is easy to check that this operator—i.e., the RHS of (3)—satisfies Blackwell's sufficient conditions for contraction (monotonicity and discounting). Therefore, (3) has a unique solution; that is, an equilibrium exists, and is unique.

We now turn to complete Nash overpricing. Consider the formulas for NiN prices:

$$p_j^N = \beta v_j + \frac{\beta}{q_j} \sum_{k \neq j} (p_k^N - v_{k,-j}) q_{k,-j} + (1 - \beta) c_j. \tag{4}$$

Let p^0 be the vector of prices that solves the above system (uniquely), when $q_{k,-j} = 0$ for all distinct k and j. That is, $p_k^0 = \beta v_k + (1 - \beta) c_k$. Suppose that β is large enough, so that each p_k^0 is sufficiently close to v_k, so that the following holds: $p_k^0 > v_{k,-j}$ for all $j \neq k$.

Now increase $\{q_{k,-1}\}_{k>1}$ from zero to their true values. These $(J - 1)$ coefficients only appear in the formula for the first price, and because $p_k^0 > v_{k,-1}$ for all $k \neq 1$ this price increases: it changes from p_1^0 to some $\tilde{p}_1^0 > p_1^0$. The change $p_1^0 \mapsto \tilde{p}_1^0$ increases any other price p_k^0, and since prices depend positively on one another, the new price vector that results, call it p^1, satisfies $p_k^1 > p_k^0$ for all k.

Now increase $\{q_{k,-2}\}_{k \neq 2}$ from zero to their true values. By the same logic, the resulting price vector, call it p^2, satisfies $p_k^2 > p_k^1$ for all k. Repeating this process iteratively we end up with the vector of NiN prices, p^N. The analysis above implies that the following holds for all k and $j \neq k$:

$$p_k^N > \cdots > p_k^2 > p_k^1 > p_k^0 > v_{k,-j}.$$

Thus, when $\beta \sim 1$ it follows from (4) that $p_j^N \sim v_j + \frac{1}{q_j}\sum_{k \neq j}(p_k^N - v_{k,-j})q_{k,-j} > v_j$. □

In the next subsection we dive further into the linear NiN model by considering a 2-hospital example in detail. A reader who is less interested in details and is more interested in the big picture of our work can skip to Section 4.

Example: Two symmetric hospitals

Consider two hospitals, A and B. The cost of serving a patient is zero. The market has four types of patients, $\{ab, a0, ba, b0\}$. Patients of type ab (resp. ba) have a value of $u^h = 10$ from being served by hospital A (resp. B) and $u^l = 5$ from being served by the other hospital. Patients of type $a0$ (resp. $b0$) have a value of $u^h = 10$ from hospital A (resp. B) but would leave the network if hospital A (resp. B) leaves the network. In other words, the ab and $a0$ types prefer a over the alternatives but disagree on their second choice hospital (B or out-of-network). The ab and ba types (equivalently $a0$ and $b0$) disagree on whether their first option is A or B.

With both hospitals in the network, the insurer expects a unit mass of patients for each hospital. There are α patients of type ab and the same for ba, and $(1-\alpha)$ patients of each of types $a0$ and $b0$.

Thus:

$$F(p) = 20 - p_A - p_B \text{ ; and } F_{-A}(p) = \alpha \cdot (5 - p_B) + (10 - p_B) \text{ .}$$

Under Nash bargaining, if we hold p_B fixed then p_A solves:

$$\max_p (F - F_{-A})^{(1-\beta)} \cdot p^\beta$$

The price response is given by:

$$p_A^N = \beta \cdot \left[10\,(1-\alpha) + \alpha\left(5 + p_B^N\right)\right]. \tag{5}$$

Equation (5) shows that A obtains a fraction β of the surplus it generates. The $(1-\alpha)$ new patients each account for 10 utils. The α patients that instead would have went to B only gain 5 directly from going to their preferred hospital, but also save the payment of p_B.

Of these two consumer segments ($1-\alpha$ and α), prices may surpass value only because of the second (α) group. In particular, whenever $p_B > 5$, the insurer actually generates negative surplus serving these patients without A in the network. The surplus that A generates to these patients is then higher than it's ex-post per-patient value of 10. If the hospitals' bargaining power is sufficiently high so that hospitals obtain most of the surplus they generate, price will be higher than the ex-post value.

Formally, prices are obtained by solving (5) and the symmetric equation for p_B. This gives:

$$p_j^N = 5\beta\frac{2-\alpha}{1-\beta\alpha},$$

for both $j \in \{A, B\}$. The following is easy to verify:

$$p_j^N \leq 10 \iff \beta \leq \frac{2}{2+\alpha}. \tag{6}$$

Since the value from being served by the top choice is 10, it follows that for this example, $\bar{\beta}$ from Theorem 1 equals $\frac{2}{2+\alpha}$.

The example is easily generalized for any positive u^h and u^l. In particular, letting $u^l = \lambda u^h$ for $\lambda \in (0,1)$ the insurer generates negative surplus (i.e., $p_j^N > u^h$) iff:

$$\beta \geq \frac{1}{1+\alpha(1-\lambda)}.$$ [10]

In particular, for $\alpha, \lambda \in (0,1)$, there is some $\bar{\beta} < 1$ such that for any $\beta > \bar{\beta}$ the price of each service is higher than its value.

4 Sequential negotiations

The difficulty in the linear NiN model lies in the hypothetical "disagreement event" associated with each negotiation. For example, in the 2-hospital case, when A is out of the network (i.e., there is disagreement with A) its entry-contribution exceeds its true value because its absence from the network generates an adverse shift in patients going to B which would be welfare reducing at B's price. Avoiding this outcome crucially depends on constructing a bargaining mechanism where the impact of disagreement in a certain problem on other bargaining problems is taken into account. We now take this approach, in our *sequential Nash model*, which is as follows: the hospitals are ordered in a (commonly known) sequence, and if negotiation breaks down with some hospital j, all subsequent negotiations assume that j is not in the insurer's network. That is, the economic environment is the same as the one considered in linear NiN, the only difference is that the Nash products reflect the order of negotiations.

It is easy to see that the insurer's surplus is positive under sequential negotiations.

[10] For $\lambda = 0.5$ one obtains (6).

Consider again the 2-hospital case with A going first and B going second. Given any outcome of negotiations with A, the bargaining problem with B must result in a non-negative addition to the insurer's overall surplus (or else the insurer will not sign a deal with B). Additionally, one can map any possible outcome in the A-negotiations, say o, to the subsequent bargaining problem with B, say $P(o)$. Since the insurer's surplus in $P(o)$ is non-negative given any possible o, the bargaining problem with A boils down to a standard bargaining problem in which both parties make positive profits. This idea generalizes to any length of hospital-sequence.

The sequential order of negotiations does not affect the insurer's surplus, but it does affect negotiated prices and hospital payoffs. For example, in the 2-hospital case where A is first, disagreement with A automatically makes B a monopolist. In contrast, disagreement with B cannot have such a favorable effect on A's bargaining position, since it can only happen after the interaction with A has concluded. Specifically, either (i) a deal with A has already been signed and so A's price is fixed, or (ii) A has dropped out, and is no longer in the network. We show that if there are only two hospitals, or if there are J symmetric hospitals, then prices are increasing in the hospital's position in the negotiation-sequence.

We start by presenting the 2-hospital case and then turn to the J-hospital case.

4.1 Two hospitals

The insurer negotiates first with hospital A. Consider the insurer's bargaining with hospital B given that bargaining with hospital A has concluded. Since p_A has already been determined, this bilateral negotiation is effectively equivalent to the linear NiN bargaining setup. Without hospital B, the insurer's surplus is determined by his agreement with hospital A. The insurer's surplus with only A is:

$$F_{-B} = q_A(v_A - p_A) + q_{A,-B}(v_{A,-B} - p_A).$$

For any price p_B the insurer's value with both hospitals in the network is:

$$F = q_A(v_A - p_A) + q_B(v_B - p_B).$$

Since the additional surplus accruing to the insurer from adding hospital B to a network with only hospital A is bounded below by zero (else B would not be added), it is given by:

$$F - F_{-B} = \max\{0, q_B(v_B - p_B) - q_{A,-B}(v_{A,-B} - p_A)\}.$$

For the moment, we assume that the addition of B is worthwhile, hence:

$$F - F_{-B} = q_B(v_B - p_B) - q_{A,-B}(v_{A,-B} - p_A).$$

After making some calculations, we will verify, ex post, that this assumption is indeed correct.

The Nash product is $[F - F_{-B}]^{1-\beta} \cdot [q_B(p_B - c_B)]^{\beta}$. Maximizing it gives the price:

$$p_B = \beta \frac{q_B v_B - q_{A,-B}(v_{A,-B} - p_A)}{q_B} + (1 - \beta)c_B. \tag{7}$$

Substituting this p_B into the expression for F we obtain:

$$\begin{aligned} F &= q_A(v_A - p_A) + q_B v_B - q_B \left[\beta \frac{q_B v_B - q_{A,-B}(v_{A,-B} - p_A)}{q_B} + (1-\beta)c_B\right] \\ &= q_A(v_A - p_A) + q_B(v_B - c_B)(1-\beta) + \beta q_{A,-B}(v_{A,-B} - p_A) \ . \end{aligned} \tag{8}$$

Now consider bargaining with A. Without A, the insurer will bargain with B, when the insurer's outside option is zero. Maximizing the Nash product for this problem gives:

$$F_{-A} = (1-\beta)(q_B(v_B - c_B) + q_{B,-A}(v_{B,-A} - c_B)).^{11}$$

Therefore,

$$F - F_{-A} = q_A(v_A - p_A) + \beta q_{A,-B}(v_{A,-B} - p_A) - (1-\beta)q_{B,-A}(v_{B,-A} - c_B).$$

Maximizing the Nash product $[F - F_{-A}]^{1-\beta} \cdot [q_A(p_A - c_A)]^{\beta}$ gives:

$$p_A = \beta \frac{q_A v_A + \beta q_{A,-B} v_{A,-B} - (1-\beta) q_{B,-A}(v_{B,-A} - c_B)}{q_A + \beta q_{A,-B}} + (1-\beta)c_A \ . \tag{9}$$

Equipped with these formulas, we can turn to the results.

Proposition 1. *In the sequential Nash model with two hospitals, the insurer's surplus is independent of the order of negotiations.*

Proof. Consider the case where hospital A is first and B is second. Plugging p_A from

[11] The insurer obtains a $(1-\beta)$-fraction out of the surplus.

(9) into the expression for the surplus, (8), gives the surplus:

$$(1-\beta)q_A(v_A - c_A) + (1-\beta)q_B(v_B - c_B) + \beta(1-\beta)q_{A,-B}(v_{A,-B} - c_A) + \\ + \beta(1-\beta)q_{B,-A}(v_{B,-A} - c_B).$$

Clearly, the same expression obtains if B goes first. □

We now turn to verify that the above derivation is valid; namely, that it is worthwhile to have the second hospital join the network at the second stage, and the optimum network is the full one. W.l.o.g, it suffices to show that the full network brings greater surplus than the one consisting only of B. Utilizing the expression derived in Proposition 1's proof, what needs to be verified is:

$$(1-\beta)q_A(v_A - c_A) + (1-\beta)q_B(v_B - c_B) + \beta(1-\beta)q_{A,-B}(v_{A,-B} - c_A) + \\ + \beta(1-\beta)q_{B,-A}(v_{B,-A} - c_B) > (1-\beta)q_B(v_B - c_B) + (1-\beta)q_{B,-A}(v_{B,-A} - c_B),$$

which simplifies to:

$$q_A(v_A - c_A) + \beta q_{A,-B}(v_{A,-B} - c_A) > (1-\beta)q_{B,-A}(v_{B,-A} - c_B).$$

By definition, $q_A \geq q_{B,-A}$, and by assumption (IV) we have that $v_A - c_A > v_{B,-A} - c_B$. Thus, the first term on the LHS is larger than the RHS. The remaining element on the LHS is positive, by assumption (III).

The intuition behind Proposition 1 is that when bargaining with the first hospital in the sequence, A, the insurer internalizes the effect that this bargaining will have on the next problem in the sequence, namely negotiations with B. This is exemplified by the various terms on the RHS of (9). Most importantly, the denominator $q_A + \beta q_{A,-B}$ expresses this internalization. If, for example, $q_{A,-B}$ is large, then one implication of signing a deal with A is that if there will be disagreement with B, many of B's would-be patients would now go to A, and in order to prevent the possibility of overpricing a la linear NiN, the price needs to be adjusted sufficiently downwards, and this is exactly what happens according to the price formula.

We now turn to the hospitals' profits. Denote by π_j^1 the profit of hospital j if it is the first hospital in the sequence, and denote by π_j^2 its profit if it is second.

Proposition 2. *Consider the sequential Nash model with two hospitals. There exists a $\beta^* < 1$ such that if $\beta \in (\beta^*, 1)$ then $\pi_j^1 < \pi_j^2$ for both $j = A, B$.*

The intuition behind Proposition 2 is that being last in the negotiations sequence confers a monopolistic position on the hospital, and this has no counterpart for the first position in the sequence. The proof is relegated to Appendix C.

4.2 An arbitrary number of hospitals

In the case of $J \geq 3$ hospitals we add the following assumption: we assume that patients leave the insurer if their top two options leave the network. To see the importance of this assumption, suppose that the hospitals are ordered from 1 to J, and consider negotiations with hospital $j - 1$. If there is disagreement in these negotiations, then the insurer moves on to bargain with hospital j. Now, in order to formulate the Nash product for *these* negotiations, it is important to know what happens in case there is disagreement with j; in particular, we need to know what would happen to the patients who had $j - 1$ as their top choice, and would choose j if $j - 1$ is out of the network. Our assumption allows us to ignore these patients; that is, to assume that they leave the network. The following is a generalization of

Proposition 1.

Proposition 3. *In the sequential Nash model with J hospitals, the insurer's surplus is independent of the order of negotiations.*

The intuition behind Proposition 3 is the same as the one behind Proposition 1. The following is a generalization of Proposition 2, under the restriction of symmetry. In its statement, π^l is the profit of a hospital if it is in the l-th position in the sequence.

Proposition 4. *Consider the sequential Nash model with J symmetric hospitals. There exists a $\tilde{\beta} < 1$ such that if $\beta \in (\tilde{\beta}, 1)$ then π^l is strictly increasing in l.*

The proofs of Propositions 3 and 4 involve some tedious algebra, and are therefore relegated to Appendix C. A key element in the analysis is the derivation of prices when negotiations happen in a sequence. Since it will be useful in the next section, we now provide (for the time being, without a proof) the price formula.

Recall that the hospitals have names: hospital 1, hospital 2, etc. Suppose that the negotiation sequence is given by these labels, namely hospital 1 is the first in the sequence, hospital 2 is second, and so on. Then, the price obtained by hospital i is:

$$p_i = \beta \frac{q_i v_i + \beta \sum_{k=i+1}^{J} q_{i,-k} v_{i,-k} - \sum_{j=1}^{i-1} q_{j,-i}(v_{j,-i} - p_j) - (1-\beta)\sum_{j=i+1}^{J} q_{j,-i}(v_{j,-i} - c_j)}{q_i + \beta \sum_{k=i+1}^{J} q_{i,-k}} + \\ + (1-\beta)c_i \quad (10)$$

This formula generalizes (9). Setting in the formula $i = 1$ makes the middle term in the numerator disappear.[12] Therefore, when $\beta \sim 1$ this price is approximately:

$$p_1^D = \frac{q_1 v_1 + \sum_{k=2}^{J} q_{1,-k} v_{1,-k}}{q_1 + \sum_{k=2}^{J} q_{1,-k}}. \quad (11)$$

[12]This terms refers to i's predecessors, which, for $i = 1$, do not exist.

This hospital 1's *standalone value*—what it contributes (per patient) when it is the only hospital in the network. It is easy to see that $p_1^D < v_1$.

5 A multi-period model

We now turn to utilizing our results from the previous section in the following infinite-horizon model. In each period $t = 1, 2, \cdots$, the insurer makes simultaneous price offers to the hospitals. These are given by a publicly-observed vector $(p_1(t), \cdots, p_J(t))$, where $p_j(t)$ is the offer made to hospital j. The hospitals react simultaneously by accept/reject responses. If all accept, the prices are implemented and play moves on to the next period. Once there is a rejection by a single hospital, say j^D, the following applies:

- With every $j \neq j^D$, the price $p_j(t)$ is contracted with for period t, where t is the period being considered (i.e., the one when the deviation occurred).

- The price contracted with j^D is $\tilde{p}_{j^D}$; it is determined by bargaining between this hospital and the insurer, but as we will soon see there is no importance to what exactly this price is, or to the bargaining mechanism that generates it.

- From $t+1$ onwards, prices are set at $(p_1(j^D), \cdots, p_J(j^D))$ in every period, where these are the prices obtained from the sequential negotiations model, where the negotiations order is one of the orders in which j^D is placed first in the sequence. Specifically, the order is selected by a uniform randomization over all orders in which j^D is placed first.

This model is sparse on information. In particular, it is not a fully-specified extensive-form game. First, only "on-path" behavior is described, and once play goes "off path" prices are set to the aforementioned levels $(p_1(j^D), \cdots, p_J(j^D))$. Second, our description of the shift from the on-path to the off-path phase is also partial, because we do not specify what happens once several hospitals reject their offers simultaneously.

We assume that all hospitals share the discount factor $\delta \in (0,1)$ and we look for a time-independent price vector $(p_1^*, \cdots, p_J^*)$ that will be accepted by all hospitals. Consider hospital 1. Its associated incentive constraint is:

$$(1-\delta)\tilde{p}_{1^D} + \delta p_1^D \leq p_1^*. \tag{12}$$

The analogous equation holds for any $j = 2, \cdots, J$. A solution to this model is a vector of prices, $(p_1^*, \cdots, p_J^*)$, that satisfy the J constraints and maximize the insurers' payoff. Clearly, these are the prices under which the J inequalities hold as equalities, and as $\delta \to 1$ each p_i^* converges to p_i^D. By equation (11), as $\beta \to 1$ the price p_i^D converges to i's standalone value, $\frac{q_i v_i + \sum_{k \neq i} q_{i,-k} v_{i,-k}}{q_i + \beta \sum_{k \neq i} q_{i,-k}}$. Since this standalone value is strictly smaller than v_i, the insurer makes positive profits under the model's solution when $(\beta, \delta) \sim (1,1)$.

Empirical Applicability

The linear NiN framework has served as a workhorse model for the empirical investigation of multilateral bargaining, in part due to it's empirical tractability (Crawford and Yurukoglu 2012, Gowrisankaran et al. 2015, Ho and Lee 2017). This tractability stems from the fact that, given observed prices and estimated demand elasticities, it is possible to recover implied costs from the linear NiN model by solving a linear system of equations. This permits straightforward estimation of linear NiN bargaining parameters and costs via generalized method of moments. In Appendix C, we show that our multi-period model is also tractable for estimation, and represents a statistical generalization of linear NiN with one additional parameter—the discount factor δ, and the system of equations that define costs under the multi-period model remains linear (see equation 17 in Appendix C). Thus, there is no increase in computational complexity when estimating our proposed model instead of linear NiN.

As hospitals' discount factor δ approaches zero, the multi-period model, informally speaking, converges to a static model, and one would like to know what is the relation between the resulting "static" model and NiN. Our analysis regarding estimation sheds light on this question:

Proposition 5. *Hospital costs and bargaining parameter estimates implied by linear NiN model estimation are equivalent to those implied by multi-period model estimation constrained to* $\delta = 0$.

Consequently the linear NiN model is testable against our more general proposal using a straightforward hypothesis test. We establish this result in Remark 1 in Appendix C.

6 Closing comments

We have studied bargaining between an intermediary and suppliers in which the total surplus from a suppliers' network is non-linear in the quantity sold by each supplier. This feature makes sense when the intermediary acts on behalf of price-insensitive users. We showed that under the common NiN approach, suppliers may charge unit prices that surpass the unit value of their service, because of the negative surplus which is created from directing users to second-best choices. If negotiations happen in a sequence, this overpricing problem cannot arise.

We have also constructed a multi-period model in which all suppliers are treated identically in terms of their place in the bargaining mechanism (as opposed to one-shot sequential negotiations), and the intermediary makes a positive payoff (as opposed to one-shot simultaneous negotiations). This model has a non-cooperative aspect—offers and responses on the path—and a cooperative aspect—Nash-bargaining payoffs off the path. To the best of our knowledge, we are the first to consider such a hybrid approach. A final question is whether using the more complicated multi-period model is likely to result in different conclusions than linear NiN. This is of course an empirical question and outside the scope of this paper, so we leave it to future work.

7 Appendix A: Exclusion in the linear NiN framework

7.1 Ex ante exclusion

In the *linear NiN model with ex-ante exclusion* a group of J hospitals is selected out of a finite pool of potential hospitals, and only then, in a second stage, the NiN interaction occurs with the selected hospitals. The solution concept for this model is subgame perfect equilibrium: the insurer selects a group of hospitals such that its surplus will be maximized, given the second-stage NiN prices.

Proposition 6. *In the linear NiN model with ex-ante exclusion there exists a* $\beta^* < 1$ *such that if* $\beta \in (\beta^*, 1)$ *then in any subgame perfect equilibrium the selected network consists of a single hospital.*

Proof. Let $\mathcal{X}$ denote the hospital pool. By Theorem 1, for each subset of hospitals $X \subset \mathcal{X}$ with at least two hospitals, there is a discount factor $\beta_X < 1$ such that if $\beta > \beta_X$ the insurer makes a negative surplus in the linear NiN model in which the hospital network is X. Since the insurer makes a positive surplus when it Nash bargains with any single hospital, the result follows by taking β^* to be the maximum of the β_X's. □

Proposition 5 implies that this two-stage structure cannot resolve the overpricing problem satisfactorily: within the two-stage framework, the threat of overpricing leads to a monopoly.

7.2 Ex post exclusion

An alternative to ex ante exclusion is *ex-post exclusion linear NiN*, suggested in Crawford and Yurukoglu (2012). Here, the insurer may remove hospitals from the network after the negotiation is complete. Ex-post exclusion guarantees the insurer at least zero surplus—it is also possible to "undo" contracts—and thus avoids complete Nash overpricing. However, *some* Nash overpricing may persist. That is, it is still possible that for a sufficiently large β *some* hospital's price exceeds the value of service. This is illustrated in the following example.

Example: Partial Nash overpricing

Consider two hospitals and two types of patients ab, ba, with the same $u^h = 10$ and $u^l = 5$ as in Example 1. However, in contrast to Example 1, assume a unit measure of patients, with s of the patients type ab and $1 - s$ of the patients of type ba.

The insurer's value from a full network given prices p_A, p_B is:

$$F(p) = \max\{0, 10 - p_A s - p_B(1 - s), 10s + 5(1 - s) - p_A, 10(1 - s) + 5s - p_B\}$$

Without A, the insurer's value given p_B is:

$$F_{-A}(p) = \max\{0, 10(1-s) + 5s - p_B\}$$

The Nash bargaining price responses are:

$$p_A(p_B) = \begin{cases} \beta(p_B + 5) & p_B \leq 10 - 5s \\ \beta\frac{10 - p_B(1-s)}{s} & p_B \geq 10 - 5s \end{cases}; \quad p_B(p_A) = \begin{cases} \beta(p_A + 5) & p_A \leq 5(1+s) \\ \beta\frac{10 - p_A s}{1-s} & p_A \geq 5(1+s) \end{cases}$$

The pair $\hat{p}_A = 10\frac{\beta}{s(1+\beta)}$ and $\hat{p}_B = 10\frac{\beta}{(1-s)(1+\beta)}$ is a solution if β and s are such that $\hat{p}_A > 5(1+s)$ and $\hat{p}_B > 10 - 5s$. Both $\hat{p}_j$ increase with β and are continuous for $\beta \geq 0$. Assume $\beta = 1$. Then for $s \in (0.5 - \hat{\xi}, 0.5 + \hat{\xi})$, with $\hat{\xi} = 1 - \frac{\sqrt{5}}{2}$ we have $\hat{p}_A = \frac{5}{s} > 5(1+s)$ and $\hat{p}_B = \frac{5}{1-s} > 10 - 5s$. Therefore, for any such s which is different from 0.5 one of the prices will exceed 10. By continuity, this is true also for all sufficiently large β's below 1.

8 Appendix B: The outside option

Under complete Nash overpricing, the insurer's objective assumes a negative value. In particular, not signing contracts with some hospitals is not a feasible alternative and the insurer's outside option is not zero, despite the fact that its payoff would have been zero if it did not sign any contract. It should be noted, however, that there is a limit to how low the insurer's payoff can be. We illustrate this for the 2-hospital case, though the idea is more general. Let $\mathcal{P}$ be the set of prices (p_A, p_B) that are consistent with an equilibrium of the model. It is enough to show that there is some bound $\bar{p}$ such that $p_A \leq \bar{p}$ for every $(p_A, p_B) \in \mathcal{P}$ that satisfy $p_B \leq p_A$. Consider then such prices. It holds that $p_A = \beta[v(AB) - v(B)] \leq \beta[\bar{v} + \alpha p_B]$, for some numbers $\bar{v} > 0$ and $\alpha \in (0,1)$. Therefore, $p_A \leq \beta[\bar{v} + \alpha p_A]$, hence $\bar{p} = \frac{\beta \bar{v}}{1 - \alpha\beta}$. In particular, prices are bounded above by $\frac{\beta \bar{v}}{1-\beta}$.

9 Appendix C: Estimating equations for the multi-period model

Estimation of patient hospital demand is independent of the bargaining problem. Therefore, we assume that the demand system is estimated prior to estimating costs and bargaining parameter.[13] This means that we can treat consumer expected valuations and choices conditional on the hospital network as known. The remaining

[13] For example, demand could be estimated following Capps, Dranove and Satterthwaite (2003). In principle, it would be more efficient to estimate demand and supply jointly in a simultaneous equations framework, which is conceptually straightforward although computationally more demanding.

parameters to estimate are the bargaining parameters (β, δ) and hospital costs. We assume that hospitals face an constant marginal cost per patient which is a function of a set of observable cost shifters, z_j and a hospital specific error term,

$$c_j = \lambda z_j + \omega_j$$

where λ is a vector of cost parameters to estimate. To estimate (β, δ, λ) we rely on the pricing equations derived in the previous section. Specifically, similar to estimation of the linear NiN model, marginal costs can be derived as a function of demand and bargaining parameters, allowing us to recover ω_j as a function of (β, δ, λ). We then jointly estimate demand and cost parameters via non-linear generalized method of moments (GMM).

Recall that the NiN price p^N negotiated if the hospital rejects the initial take-it-or-leave it offer is determined using treating all other prices as given. Letting p^* denote the observed (and multi-period equilibrium) prices, equation (3), the linear NiN prices satisfy:

$$p^D = \beta\theta^D + \beta\Gamma^D p^* + (1-\beta)c \tag{13}$$

Where, θ^D is a vector and Γ^D is a matrix defined by:

$$\theta^D_j = v_j - \frac{\sum_{\ell \neq j} v_{\ell,-j} q_{\ell,-j}}{q_j} \ \ ; \ \Gamma^D_{j,l} = \frac{q_{l,-j}}{q_j} \ \ ; \ \Gamma^D_{j,-J} = 0$$

Equation 13 proves Proposition 5: The price vector on the right hand side is the observed prices. In particular, $p^N = p^*$ if and only if the linear NiN model is correct.

Prices under the multi-period model are a convex combination of NiN prices and each hospital's "punishment" price from reverting to the sequential model with that hospital negotiating in the first (least-favorable) position determined by the incentive constraint (12). Rewriting this formula in terms of markups and using matrix notation obtains,

$$p^F = \theta^F(\beta) + (I + \Psi^F(\beta)) \cdot c \tag{14}$$

The matrix $\Psi^F(\beta)$ accounts for the impact of disagreement with hospital j on the cost of treating patients,

$$\Psi^F_{j,-J} = \beta, \qquad \Psi^F_{j,k\neq j} = \beta(1-\beta)\frac{q_{k,-j}}{q_j + \beta\sum_{k\neq j} q_{j,-k}}.$$

From this formula we see that a unit increase in hospitals cost causes price to go up by $1+\beta$. And is increasing in the costs of rival hospitals in proportion to substitution to those hospitals when j is dropped from the system. That is, in the sequential game, j's bargaining position is enhanced when it's patients are likely to substitute to high cost hospitals. This feature of the sequential model will be inherited by the multi-period solution but is absent from the linear NiN solution, which considers only the importance of hospital j's provision of surplus holding other hospital prices fixed.

The vector $\theta^F(\beta)$ is,

$$\theta_j^F(\beta) = \beta \frac{q_j v_j + \sum_{k \neq j} [\beta q_{j,-k} v_{j,-k} - (1-\beta) q_{k,-j} v_{k,-j}]}{q_j + \beta \sum_{k \neq j} q_{j,-k}}$$

The first term of the numerator represents the hospitals contribution to surplus, the remaining terms are adjustments to hospital j's bargaining position based on cross-hospital substitution. If hospital j is rewarded to the extent that it can serve as a substitute for hospital k in the event that bargaining with k fails. On the other hand, j's bargaining position is reduced if other hospitals are strong substitutes in the event of its own disagreement.

Next, construct the multi-period model estimator by merging the linear NiN and punishment price equations. Following the analysis in Section 5, the observed price in the multi-period model is given by

$$p^* = \delta p^F + (1-\delta) p^D$$

Using equations (13) and (14) we have an expression that is linear in prices and costs,

$$p^* = \delta \theta^F(\beta) + \delta(I + \Psi^F(\beta))c + (1-\delta)\beta(\theta^D + \Gamma^D p^*) + (1-\delta)(1-\beta)c$$

To solve for either p^* or c, rewrite as:

$$0 = \Psi(\beta,\delta)c + \theta(\beta,\delta) + \Gamma(\beta,\delta)p^* \tag{15}$$

Here, the terms multiplying the cost vector c and price vector p are aggregated into the matrices Ψ and Γ, and the constant terms are aggregated into the vector θ:

$$\begin{aligned} \Psi(\beta,\delta) &= I(1-\beta(1-\delta)) + \delta \Psi^F(\beta,\delta) \\ \Gamma(\beta,\delta) &= (1-\delta)\beta\Gamma^D - I \\ \theta(\beta,\delta) &= \delta\theta^F(\beta) + (1-\delta)\beta\theta^D \end{aligned} \tag{16}$$

To back out costs from the demand system and bargaining parameters, solve (15) for costs,

$$c(\beta,\delta) = -\Psi(\beta,\delta)^{-1}(\theta(\beta,\delta) + \Gamma(\beta,\delta)p^*). \tag{17}$$

Proposition 5 is now immediate by observation: The linear NiN model is a special case of the multi-period model with the value for δ set to zero.

Estimation of the supply side parameters using the multi-period model therefore is similar to the existing method for linear NiN with the additional parameter (δ) and a slightly more complicated non-linear function for costs. Specifically, given a set of bargaining parameters the structural error in costs is,

$$\omega_j(\alpha_0, \alpha_1, \beta, \delta) = c_j(\beta,\delta) - \gamma z_j.$$

To construct the moments, define $h_{j,n}$ as a vector of instruments for hospital j in market n. The data moments are:

$$g_{NJ}(\alpha_0, \alpha_1, \beta, \delta) = \frac{1}{NJ} \sum_{n,j} h_{n,j} \omega_{n,j}(\alpha_0, \alpha_1, \beta, \delta) \ .$$

The GMM estimator is:

$$\underset{\alpha_0, \alpha_1, \beta, \delta}{\operatorname{argmin}} \, g_{NJ}(\alpha_0, \alpha_1, \beta, \delta)' W g_{NJ}(\alpha_0, \alpha_1, \beta, \delta).$$

Where W is a symmetric positive definite weight matrix. We use the standard 2-step GMM to derive the optimal weight matrix for each dataset.

The observed cost shifters z_j, are available as instruments, but we clearly need two additional instruments to identify the two bargaining parameters. Candidates are most likely to come from the exogenous variation in the demand system, which generates variation in the substitutability of hospitals (e.g., $v_{j,-k}$ and $q_{j,-k}$) that represent the key primitives in the matrices defined in (16). An example of such an instrument could be the distances between hospitals, or the relative weights of different types of observable consumers which vary across markets. In the Monte Carlo analysis below, we will assume the existence of an observable demand shifter which will serve as instruments.

To conclude this section, note that, due to (5) the simultaneous linear NiN model nests the multi-period model by fixing $\delta = 0$, allowing us to test linear NiN against a more general alternative. To facilitate comparison with the earlier literature, the following remark establishes the connection between our notation and that of GNT.

Remark 1. *To facilitate comparison with existing work, in particular Gowrisankaran et al. (2015), rewrite (3) as*

$$p_j^N - c_j = \frac{\beta}{1-\beta} \left(v_j - p_j^N - \frac{\sum_{l \neq j} q_{l,-j}(v_{l,-j} - p_l)}{q_j} \right) \tag{18}$$

Next, let ξ_j equal to the right hand side of equation (18) *for hospital j and define Λ as a diagonal matrix with elements $\Lambda_{j,-J} = -\frac{1}{\xi_j q_j}$. Then,*

$$-\Lambda(p_j^N - c_j) = q_j$$

Λ is provided from the first stage estimates and observed prices. Note that this is exactly equation 13 in Gowrisankaran et al. (2015), with the only differences that Λ in that paper has off-diagonal elements to account for cross hospital effects for hospitals within the same hospital system and that we omit their Ω which captures patient price sensitivity of demand.

10 Appendix D: Proofs

Proof of Proposition 2: Wlog, consider hospital A. Since the quantities are unaffected by price and by the order of negotiations, it is enough to show that for all sufficiently large β's we have $p_A^1 < p_A^2$, where p_A^l is the price corresponding to π_A^l. It is enough to verify that that's the case when $\beta = 1$.

Setting $\beta = 1$ in (9) gives:

$$p_A^1 = \frac{q_A v_A + q_{A,-B} v_{A,-B}}{q_A + q_{A,-B}}. \tag{19}$$

Setting $\beta = 1$ in the analog of (7) gives:

$$p_A^2 = \frac{q_A v_A - q_{B,-A}(v_{B,-A} - p_B^1)}{q_A}.$$

We argue that $\frac{q_A v_A + q_{A,-B} v_{A,-B}}{q_A + q_{A,-B}} < \frac{q_A v_A - q_{B,-A}(v_{B,-A} - p_B^1)}{q_A}$. Simplifying this expression we get:

$$q_A q_{A,-B}(v_{A,-B} - v_A) < -q_{B,-A}(v_{B,-A} - p_B^1)(q_A + q_{A,-B}).$$

The LHS is negative since $v_{A,-B} < v_A$. Therefore, it is enough to prove that the RHS is positive, or that $p_B^1 > v_{B,-A}$. Clearly, it is enough to show that $p_A^1 > v_{A,-B}$. This follows immediately from (19), since $v_A > v_{A,-B}$. □

Lemma 1. *Consider Nash bargaining between the insurer and a hospital, under the following assumptions:*

1. *The insurer's profit without the hospital is V_0.*

2. *The hospital's unit cost is c and its bargaining power parameter is β.*

3. *If the hospital joins the network it serves a population of mass q.*

4. *Adding the hospital to the network at unit price p increases the insurer's profit by $K - p \cdot y$.*

Then the price is:

$$p = \beta \frac{K}{y} + (1 - \beta)c. \tag{20}$$

The lemma's proof boils down to a simple maximization of a Nash product, and is therefore omitted.

In what follows we consider sequential negotiations between the insurer and the hospitals, where the hospitals are order in a particular, commonly known order. Given the order, each hospital has a position in it—the first in line, the second in line, etc. In addition to that, the hospitals have names—hospital 1, hospital 2, etc.—and each name is associated with particular model-parameter-values, such as q_1, q_2, and so on. We use the term **canonical order** to denote the order where the two label-systems coincide; that is, hospital 1 is placed first in the canonical order, hospital 2 is second, and so on.

Lemma 2: Consider sequential negotiations according to the canonical order, $\{1, \cdots, J\}$. Given this order, denote the insurer's surplus after bargaining and signing contracts with hospitals $\{1, \cdots, i\}$ for prices $(p_1, \cdots, p_i)$, and facing "future hospitals" $\{i+1, \cdots, J\}$, by $V(p_1, \cdots, p_i; \{i+1, \cdots, J\})$. This surplus is given by:

$$V(p_1, \cdots, p_i; \{i+1, \cdots, J\}) =$$
$$= \sum_{j=1}^{i}[q_j(v_j - p_j) + \beta \sum_{k=i+1}^{J} q_{j,-k}(v_{j,-k} - p_j)] +$$
$$+ (1-\beta) \sum_{j=i+1}^{J} [q_j(v_j - c_j) + \beta \sum_{k=i+1, k\neq j}^{J} q_{j,-k}(v_{j,-k} - c_j)].$$

The price obtained by hospital i in these negotiations is:

$$p_i = \beta \frac{q_i v_i + \beta \sum_{k=i+1}^{J} q_{i,-k} v_{i,-k} - \sum_{j=1}^{i-1} q_{j,-i}(v_{j,-i} - p_j) - (1-\beta) \sum_{j=i+1}^{J} q_{j,-i}(v_{j,-i} - c_j)}{q_i + \beta \sum_{k=i+1}^{J} q_{i,-k}} +$$
$$+ (1-\beta) c_i$$

Before we turn to the proof, it is worthwhile to consider the equation for the above-mentioned value function V. The RHS is composed of two terms, one corresponding to the already-contracted-with hospitals and one corresponding to the "future hospitals," and the first term is such that for each element in the summation (each $j = 1, \cdots, i$) we have the "direct value" from the hospital plus β times the value that this hospital generates assuming that all "future hospitals" drop out. It will be useful to bear this meaning in mind later on in the proof.

Proof of Lemma 2: Clearly, $V(p_1, \cdots, p_J; \emptyset) = \sum_{j=1}^{J} q_j(v_j - p_j)$. When negotiating with hospital J, given that contracts with all previous hospitals have been signed, the insurer's outside option is the value $V(p_1, \cdots p_{J-1}; \emptyset) = \sum_{j=1}^{J-1}[q_j(v_j - p_j) + q_{j,-J}(v_{j,-J} - p_j)]$. Thus, the gain from adding J is:

$$W_J(p_1,\cdots,p_J;\emptyset) \equiv V(p_1,\cdots,p_J;\emptyset) - V(p_1,\cdots,p_{J-1};\emptyset) = q_J(v_J - p_J) - \sum_{j=1}^{J-1} q_{j,-J}(v_{j,-J} - p_j) =$$

$$= \underbrace{q_J v_J - \sum_{j=1}^{J-1} q_{j,-J}(v_{j,-J} - p_j)}_{K} - p_J \underbrace{q_J}_{y},$$

where the "K" and "y" are the notations of Lemma 1. Applying this lemma we obtain the price p_J:

$$p_J = \beta \frac{q_J v_J - \sum_{j=1}^{J-1} q_{j,-J}(v_{j,-J} - p_j)}{q_J} + (1-\beta)c_J. \tag{21}$$

Having obtained the price and surplus for the last bargaining problem in the sequence, we turn to the next-to-last bargaining. The value of these negotiations is $V(p_1,\cdots,p_{J-1};\{J\}) = \sum_{j=1}^{J-1} q_j(v_j - p_j) + q_J(v_J - p_J)$.[14] Note that we slightly abuse notation to have p_J on the RHS is legitimate even though it does not appear as an argument of the V function on the LHS, because of (21)—namely, p_J is pinned down by the previous prices (and the other model parameters). Substituting p_J into the expression gives:

$$V(p_1,\cdots,p_{J-1};\{J\}) = \sum_{j=1}^{J-1}[q_j(v_j - p_j) + \beta q_{j,-J}(v_{j,-J} - p_j)] + (1-\beta)q_J(v_J - c_J). \tag{22}$$

The outside option in the next-to-last negotiations has the value $V(p_1,\cdots,p_{J-2};\{J\})$. It follows from (22) that this value is:

$$V(p_1,\cdots,p_{J-2};\{J\}) = \sum_{j=1}^{J-2}[\underbrace{q_j(v_j - p_j) + q_{j,-(J-1)}(v_{j,-(J-1)} - p_j)}_{L} + \beta q_{j,-J}(v_{j,-J} - p_j)] +$$

$$+ (1-\beta)\underbrace{[q_J(v_J - c_J) + q_{J,-(J-1)}(v_{J,-(J-1)} - c_J)]}_{M}.$$

Here, L is the counterpart of $q_j(v_j - p_j)$ from (22), and M is the counterpart of $q_J(v_J - c_J)$ (expected value minus expected cost at the final hospital).[15]

[14] The value when the prices up to and including p_{J-1} have been contracted, and the insurer *expects* p_J to be contracted next is $\sum_{j=1}^{J} q_j(v_j - p_j)$.

[15] This is true because (22) holds also for a sequence of length $J' = J - 1$.

The gain from the $(J-1)$-th bargaining is $W_{J-1}(p_1, \cdots, p_{J-1}; \{J\}) = V(p_1, \cdots, p_{J-1}; \{J\}) - V(p_1, \cdots, p_{J-2}; \{J\})$, or:

$$
\begin{aligned}
&W_{J-1}(p_1, \cdots, p_{J-1}; \{J\}) = \\
&= \underbrace{q_{J-1}v_{J-1} + \beta q_{J-1,-J}v_{J-1,-J} - \sum_{j=1}^{J-2} q_{j,-(J-1)}(v_{j,-(J-1)} - p_j) - (1-\beta)q_{J,-(J-1)}(v_{J,-(J-1)} - c_J) -}_{K} \\
&\qquad - p_{J-1}\underbrace{(q_{J-1} + \beta q_{J-1,-J})}_{y}.
\end{aligned}
$$

By Lemma 1,

$$
\begin{aligned}
&p_{J-1} = \\
&= \beta \frac{q_{J-1}v_{J-1} + \beta q_{J-1,-J}v_{J-1,-J} - \sum_{j=1}^{J-2} q_{j,-(J-1)}(v_{j,-(J-1)} - p_j) - (1-\beta)q_{J,-(J-1)}(v_{J,-(J-1)} - c_J)}{q_{J-1} + \beta q_{J-1,-J}} + \\
&\qquad + (1-\beta)c_{J-1}.
\end{aligned}
$$

It follows from equation (22) that:[16]

$$
\begin{aligned}
V(p_1, \cdots, p_{J-2}; \{J-1, J\}) &= \sum_{j=1}^{J-2}[q_j(v_j - p_j) + \beta q_{j,-J}(v_{j,-J} - p_j)] + (1-\beta)q_J(v_J - c_J) + \\
&+ q_{J-1}v_{J-1} + \beta q_{J-1,-J}v_{J-1,-J} - p_{J-1}(q_{J-1} + \beta q_{J-1,-J}).
\end{aligned}
$$

Combining this with the formula for p_{J-1} gives:

$$
\begin{aligned}
&V(p_1, \cdots, p_{J-2}; \{J-1, J\}) = \\
&= \sum_{j=1}^{J-2}[q_j(v_j - p_j) + \beta \sum_{k=J-1}^{J} q_{j,-k}(v_{j,-k} - p_j)] + (1-\beta) \sum_{j=J-1}^{J} [q_j(v_j - c_j) + \beta \sum_{k=J-1, k\neq j}^{J} q_{j,-k}(v_{j,-k} - c_j)].
\end{aligned}
$$

Now assume that given the contracted prices $\{p_1, \cdots, p_i\}$, and "future hospitals" $\{i+1, \cdots, J\}$, the insurer's payoff is:

[16] This is simply writing separately the $(J-1)$-th term from the first summation, leaving the first $J-2$ elements in the sum.

$$V(p_1, \cdots, p_i; \{i+1, \cdots, J\}) =$$
$$= \sum_{j=1}^{i} [q_j(v_j - p_j) + \beta \sum_{k=i+1}^{J} q_{j,-k}(v_{j,-k} - p_j)] + (1-\beta) \sum_{j=i+1}^{J} [q_j(v_j - c_j) + \beta \sum_{k=i+1, k\neq j}^{J} q_{j,-k}(v_{j,-k} - c_j)]. \tag{23}$$

As we have shown above, this assumption is indeed correct given a fixed J and $i \in \{J-2, J-1\}$. Basically, the same arguments can be applied given that the hospital sequence is of length $J' = J - 1$: the formula still holds with J' replacing J, $i = J' - 1$, and also for $i = J' - 2$ provided that this is a positive integer. But, one has to be careful in the application and note the role of our assumption that when i drops out, all of its consumers that choose $k \neq i$ as their second choice leave the network if the second choice drops out as well. This is what makes the application work, and hence (23) implies:

$$V(p_1, \cdots, p_{i-1}; \{i+1, \cdots, J\}) =$$
$$= \sum_{j=1}^{i-1} [q_j(v_i - p_j) + q_{j,-i}(v_{j,-i} - p_j) + \beta \sum_{k=i+1}^{J} q_{j,-k}(v_{j,-k} - p_j)] +$$
$$+ (1-\beta) \sum_{j=i+1}^{J} [q_j(v_j - c_j) + \beta \sum_{k=i+1, k\neq j}^{J} q_{j,-k}(v_{j,-k} - c_j)].$$

Note that, like in the explanation that preceded the proof, each element in the first summation has a direct benefit component and an additional component, when in writing down these components we have invoked the abovementioned assumption regarding what happens when i drops out.

Thus, the gain from bargaining with i is:

$$W_i(p_1, \cdots, p_i; \{i+1, \cdots, J\}) = V(p_1, \cdots, p_i; \{i+1, \cdots, J\}) - V(p_1, \cdots, p_{i-1}; \{i+1, \cdots, J\}),$$

or:

$$q_i(v_i - p_i) + \beta \sum_{k=i+1}^{J} q_{i,-k}(v_{i,-k} - p_i) - \sum_{j=1}^{i-1} q_{j,-i}(v_{j,-i} - p_j) - (1-\beta) \sum_{j=i+1}^{J} q_{j,-i}(v_{j,-i} - c_j) =$$

$$= \underbrace{q_i v_i + \beta \sum_{k=i+1}^{J} q_{i,-k} v_{i,-k} - \sum_{j=1}^{i-1} q_{j,-i}(v_{j,-i} - p_j) - (1-\beta) \sum_{j=i+1}^{J} q_{j,-i}(v_{j,-i} - c_j)}_{K} -$$

$$- p_i \underbrace{\left(q_i + \beta \sum_{k=i+1}^{J} q_{i,-k}\right)}_{y}.$$

Applying Lemma 1 we obtain:

$$p_i = \beta \frac{q_i v_i + \beta \sum_{k=i+1}^{J} q_{i,-k} v_{i,-k} - \sum_{j=1}^{i-1} q_{j,-i}(v_{j,-i} - p_j) - (1-\beta) \sum_{j=i+1}^{J} q_{j,-i}(v_{j,-i} - c_j)}{q_i + \beta \sum_{k=i+1}^{J} q_{i,-k}} + $$

$$+ (1-\beta) c_i \quad (24)$$

With (23) and (24) established, the proof is completed. □

Proof of Proposition 3: It follows from (23) that the insurer's value, before he approaches the first hospital in the canonical order, is:

$$V(\emptyset; \{1, \cdots, J\}) = (1-\beta) \sum_{j=1}^{J} [q_j(v_j - c_j) + \beta \sum_{k \neq j} q_{j,-k}(v_{j,-k} - c_j)],$$

and the RHS is independent of the order. □

Proof of Proposition 4: Consider J symmetric hospitals and let $q \equiv q_j$, $\hat{q} \equiv q_{j,-k}$, $v \equiv v_j$ and $\hat{v} \equiv v_{j,-k}$ (recall that symmetry means independence of these quantities of j and k). Let $p_i^* \equiv \lim_{\beta \to 1} p_i$, where p_i is given by (24). Setting $\beta = 1$ at (24) gives:

$$p_i^* = \underbrace{\frac{qv + \hat{q}\hat{v}(J-i)}{q + \hat{q}(J-i)}}_{A} + \underbrace{\frac{\hat{q} \sum_{j=1}^{i-1} (p_j^* - \hat{v})}{q + \hat{q}(J-i)}}_{B}.$$

Claim 1: A is increasing in i.

Proof of Claim 1: The sign of $\frac{\partial A}{\partial i}$ is the same as the sign of $-\hat{q}\hat{v}[q + \hat{q}(J-i)] + \hat{q}[qv + \hat{q}\hat{v}(J-i)]$, and the latter is positive if and only if $v > \hat{v}$, which is true.

Claim 2: B is increasing in i.

Proof of Claim 2: It is enough to prove that $p_j^* > \hat{v}$. This is true for $j = 1$ in virtue of $v > \hat{v}$, and the fact that it is true for all $j' < j$ implies that it is true also for j. $\square$

References

[1] Abreu, D. and Manea, M., 2012. Bargaining and efficiency in networks. Journal of Economic Theory, 147(1), 43-70.

[2]Binmore, K., Rubinstein, A. and Wolinsky, A., 1986. The Nash bargaining solution in economic modelling. The RAND Journal of Economics, 176-188.

[3] Bloch, F. and Jackson, M.O., 2007. The formation of networks with transfers among players. Journal of Economic Theory, 133(1), 83-110.

Capps, C., Dranove, D. and Satterthwaite, M., 2003. Competition and market power in option demand markets. RAND Journal of Economics, 43(4), 737-763.

[5] Collard-Wexler, A., Gowrisankaran, G. and Lee, R.S., 2019. "Nash-in-Nash" bargaining: A microfoundation for applied work. Journal of Political Economy, 127(1), 163-195.

[6] Crawford, G.S. and Yurukoglu, A., 2012. The welfare effects of bundling in multichannel television markets. American Economic Review, 102(2), 643-85.

[7] De Fontenay, C.C. and Gans, J.S., 2014. Bilateral bargaining with externalities. The Journal of Industrial Economics, 62(4), 756-788.

[8] Galasso, A., 2008. Coordination and bargaining power in contracting with externalities. Journal of Economic Theory, 143(1), 558-570.

[9] Gaynor, M. and Town, R.J., 2011. Competition in health care markets. Handbook of health economics, 2, 499-637.

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

[11] Ho, K. and Lee, R.S., 2017. Insurer competition in health care markets. Econometrica, 85(2), pp.379-417.

[12] Horn, H. and Wolinsky, A., 1988. Bilateral monopolies and incentives for merger. The RAND Journal of Economics, 408-419.

[13] Krasteva, S. and Yildirim, H., 2012. Payoff uncertainty, bargaining power, and the strategic sequencing of bilateral negotiations. The RAND Journal of Economics, 43(3), 514-536.

[14] Manea, M., 2018. Intermediation and resale in networks. Journal of Political Economy, 126(3), 1250-1301.

[15] Marx, L.M. and Shaffer, G., 2007. Rent shifting and the order of negotiations. International Journal of Industrial Organization, 25(5), 1109-1125.

[16] Nash, J.F., 1950. The Bargaining Problem. Econometrica, 18(2), 155-162.

[17] Rubinstein, A., 1982. Perfect equilibrium in a bargaining model. Econometrica:

Journal of the Econometric Society, 97-109.
[18] Shelah, O. 2015. Dare to Win, Yediot books.
[19] Stole, L.A. and Zwiebel, J., 1996. Intra-firm bargaining under non-binding contracts. The Review of Economic Studies, 63(3), 375-410.
[20] Xiao, J., 2018. Bargaining orders in a multi-person bargaining game. Games and Economic Behavior, 107, 364-379.