WORKING PAPERS

Simulating Hospital Merger Simulations

David Balan
Keith Brand

WORKING PAPER NO. 334

March 2018

BUREAU OF ECONOMICS
FEDERAL TRADE COMMISSION
WASHINGTON, DC 20580

Simulating Hospital Merger Simulations

David J. Balan * Keith Brand †

March 20, 2018

Abstract

We assess the performance of three hospital merger simulation methods by means of a Monte Carlo experiment. We first specify a rich theoretical model of hospital markets and use it to generate "true" price effects of a large number of hospital mergers. We then use the theoretical model to generate the data that would be available in a real-world prospective merger analysis and apply the merger simulation methods to those data. Finally, we compare the predictions of the merger simulation methods to the true price effects. While there is some heterogeneity in performance, all three simulation methods perform reasonably well.[1]

Keywords: Hospitals, Mergers, Hospital Mergers, Nash Bargaining, Merger Simulation, Oligopoly
JEL Classification: L11, L13, L31, L38, I11, I18

*Bureau of Economics, Federal Trade Commission, dbalan@ftc.gov
†Bureau of Economics, Federal Trade Commission, kbrand@ftc.gov
[1]We are grateful to Leemore Dafny, Gary Fournier, Christopher Garmon, Martin Gaynor, Gautam Gowrisankaran, Daniel Hosken, Thomas Koch, Robert McMillan, Ted Rosenbaum, David Schmidt, and Robert Town; to seminar participants at George Washington University, the University of Illinois, and the University of Oklahoma; to brown bag participants at the Department of Justice and at the Federal Trade Commission; and at the 2015 Kellogg School of Management Fourth Annual Conference on Healthcare Markets for their helpful comments. The views expressed in this article are those of the authors and do not necessarily reflect those of the Federal Trade Commission.

1 Introduction

In recent years, the economics literature has produced a number of methods for simulating the price effects of hospital mergers. These merger simulation methods have been used in internal analyses at the Federal Trade Commission (Farrell et al. (2011)). They have also been used by testifying economic experts in recent litigated hospital merger cases.[2]

The main purpose of this paper is to make a contribution to evaluating the accuracy of three of these hospital merger simulation methods. Specifically, we evaluate a variant of the Willingness-to-Pay (*WTP*) method originally exposited in Capps et al. (2003) (CDS), as well as an extension to the CDS method described in Brand (2013). We also evaluate what is known as the "Upward Pricing Pressure" (*UPP*) approach to predicting the price effects of hospital mergers.

All of these simulation methods have the important advantage of being tractable (with the *UPP* method being extremely tractable), but this tractability is the result of important simplifying assumptions, the validity of which are uncertain. The simulation methods can be thought of as approximations to a richer and more realistic theoretical model, and the accuracy of the methods in predicting the price effects of mergers will depend, in part, on the closeness of those approximations.

We present such a rich theoretical model that captures the key features of hospital markets in the United States. Among these are: (i) health insurers typically act as intermediaries between hospitals and consumers; and (ii) hospital prices are typically determined via bilateral bargaining between hospitals and insurers rather than being posted by hospitals. The primitives of the model are defined on hospital attributes (location, quality, cost, and system affiliation), consumer attributes (location and probability of using inpatient care), and consumer preferences over hospitals and insurers. We assume profit-maximizing behavior for both hospitals and insurers. The solution concept is standard "Nash-In-Nash," meaning that the equilibrium vectors of hospital prices and insurance premiums simultaneously comprise: (i) a Nash Equilibrium of solutions to a set of Nash Bargaining equations that model the bargaining between hospitals and insurers; and (ii) a Nash Equilibrium in a Bertrand game played by insurers. Our theoretical model is broadly similar to other models developed in recent papers (see Gaynor and Town (2012), Gowrisankaran et al. (2015), and Gaynor et al. (2015)).

We perform a Monte Carlo experiment in which we generate simulated data to evaluate how closely the hospital merger price effects predicted by the simulation methods approximate the "true" price effects from the theoretical model. We emphasize that it is by no means obvious *a priori* that the simulation methods must be a close approximation to the theoretical model. While the simulation methods (like the theoretical model) are based upon the realistic assumption that hospital

[2] For example, merger simulation based on *WTP* was used in the ProMedica Health System matter (https://www.ftc.gov/sites/default/files/documents/cases/2012/06/120625promedicaopinion.pdf and https://www.ftc.gov/sites/default/files/documents/cases/2012/06/120328promedicaroschopinion.pdf). A version of the *UPP* method was used in the Federal Trade Commission and State of Illinois vs. Advocate Health Care Network, Advocate Health and Hospitals Corporation, and North Shore University Health System matter (public trial transcript of Dr. Steven Tenn, April 11, 2016).

prices are set via bilateral bargaining between hospitals and insurers, they omit potentially important features that are included in the theoretical model. These include competition in the health insurance market, uncertainty at the consumer level over anticipated healthcare utilization, and the group purchase of health insurance. In addition, two of three simulation methods do not account for strategic responses in prices among non-merging hospitals. The simulation methods may also mis-specify the objective functions of insurers. The reason is that, as discussed in the next section, while insurers are best thought of as profit maximizers, the simulation methods do not model them as such. For these reasons, the simulation methods are not guaranteed to closely approximate the theoretical model. If our Monte Carlo experiment shows that they do in fact closely approximate the theoretical model, that would constitute meaningful evidence of their real-world efficacy insofar as the theoretical model is a reasonably accurate representation of the real world.

Our experiment proceeds in three stages. First, we solve the theoretical model for a large number of simulated markets under a wide variety of model parameterizations, and, for each simulated market, we calculate the price effect of every possible merger between two hospital systems. That is, we calculate the equilibrium set of hospital prices before and after every possible pairwise merger. Comparing the pre- and post-merger prices generates what we refer to as the true price effect of each merger. Second, for each simulated market, we generate the types of data that would be available in a real-world prospective merger analysis: pre-merger prices and individual-level hospital discharge data. We then apply the merger simulation methods to those data to generate a prediction of the price effect of the merger. Third, we compare the true price effects to the predicted effects of each of the three simulation methods, and evaluate their overall performance, as well as how that performance varies across model parameterizations.[3]

We determine the set of possible values of the model parameters by calibrating our results against real-world metrics, including hospital prices and costs. We include a wider range of parameter values than this calibration would suggest, both because of uncertainty on which combinations of model parameters correspond most closely to the real world and to cover real-world heterogeneity in these metrics across markets.

We find that the merger simulation methods generally perform quite well. The method based on CDS exhibits a tendency to modestly under-predict the true merger price effects, with a mean prediction error of around -15% of the true price effect. For example, if the mean true price effect is 5%, then the mean predicted price effect using that simulation method is about 4.25%. The method based on Brand (2013) exhibits a tendency to modestly over-predict, with a mean prediction error

[3]Our Monte Carlo experiment is similar to those performed by Miller et al. (2016) and Miller et al. (2017). In each of those papers, as in ours, the accuracy of a merger simulation method is evaluated by using simulated data to compare its predictions to the true results of a richer, more realistic model. The key difference is that those papers simulate mergers of differentiated products with posted prices, and ours simulates hospital mergers with negotiated prices. Another difference is that our theoretical model, while broadly similar to existing models, was developed specifically for this paper and represents a modest contribution to the theoretical literature.

of around 14% of the true price effect. Overall, *UPP* performs less well, and its performance varies much more by the magnitude of the true price effect. Its mean prediction error is 34.9% of the true price effect for mergers whose true price effect is between 4.5% and 5.5%, falling to only 3.8% of the true price effect for mergers whose true price effect is between 9.5% and 10.5%, and falling all the way to -13.3% of the true price effect for mergers whose true price effect is between 14.5% and 15.5%.

We also apply a performance measure based on the median absolute prediction error (MAPE), which captures the dispersion of the predicted effects about the true effects. The simulation methods perform quite well by this measure as well. For the method based on CDS, the MAPE is typically about 20%-25% of the true price effect. The method based on Brand (2013) performs significantly better, with the MAPE typically about 12%-14% of the true price effect. Once again, *UPP* performs somewhat less well, and its performance varies significantly with the magnitude of the true price effect.

Based on these results, we conclude that the simulation methods perform at least reasonably well in predicting the true price effects from our theoretical model. And while there is some variation in the methods' performance across different parameterizations of the theoretical model, they generally perform reasonably well throughout the parameter space. This suggests that the methods are likely to be useful even if we do not know which parts of the parameter space in our simulations correspond most closely to the real world.

A roadmap for the rest of the paper is as follows. Section 2 presents some background and discusses previous literature. Section 3 lays out our theoretical model. Section 4 specifies the parameterization of the theoretical model. Section 5 lays out the merger simulation methods. Section 6 presents the results of our Monte Carlo simulations comparing the price effects from the simulation methods with the true price effects from the theoretical model. Section 7 contains a discussion of these results. Section 8 concludes.

2 Background and Previous Literature

We begin by discussing the theoretical basis of the merger simulation methods evaluated in this paper. The methods based on CDS and Brand (2013) involve constructing a measure of hospital market power using individual-level inpatient discharge data. A key component of this market power measure was initially developed in Town and Vistnes (2001) and CDS, the latter of which first applied the now commonly-used term "Willingness to Pay" (*WTP*). As the name suggests, *WTP* is intended to capture the incremental valuation that consumers place on having a particular hospital or hospital system in their insurer's provider network. For closely related reasons, *WTP* can also be thought of as proportional to the amount by which an insurer's gross profits (gross of payments to hospitals) would decline if that hospital or hospital system was excluded from its network. In the context of the bilateral bargaining framework in which hospital prices are determined, WTP can be thought of as a measure of the difference between the insurer's gross payoff if an agreement is reached versus if it is not. For reasons discussed in Farrell et al. (2011) and in Section 5 below, the ability to negotiate a price above marginal cost is more appropriately measured as WTP divided by expected hospital volume (WTP/Q) rather than by WTP itself.

As detailed in Section 5 below, the merger simulation methods based on CDS and Brand (2013) proceed by using a least squares regression to estimate the relationship between price and the measure of market power (WTP/Q or some variation thereof) and other controls. This estimated relationship is then used to predict the price effect of a merger. The UPP method is based on a simple theory-based calculation of diversion ratios and hospital gross margins, as described in Haas-Wilson and Garmon (2009) and Garmon (2017).

A major virtue of the simulation methods is that they reflect the fact that prices are set through bargaining between hospitals and insurers. The simulation methods based on CDS and Brand (2013) also have the important advantage that they are reasonably inexpensive to evaluate, and that the individual-level inpatient discharge or claims data that they require are often available in the context of antitrust investigations. The UPP method is simpler to evaluate, and the data requirements are even lower.

While the simulation methods are consistent with the basic intuition derived from a standard bargaining model, they typically do not account for some potentially important components of bargaining between hospitals and insurers. For example, they abstract away from the oligopoly game played by competing insurers, so they do not directly account for the effect of competition in the insurance market on the bargaining between hospitals and insurers. In addition, of the three simulation methods, only the method based on Brand (2013) accounts for the dependence of a hospital's price on the prices of its competitors, or for price changes at non-merging hospitals in response to a merger.

The purpose of this paper is to evaluate the extent to which these limitations affect the accuracy of the simulation methods in predicting the price effects of hospital mergers. As noted above, we do this by performing a Monte Carlo experiment to compare the price effects predicted by the simulation methods with the true price effects of a richer theoretical model. Our paper can be characterized as treating the price effects from the WTP-based simulation methods as approximations of the true effects that come from our theoretical model, and then testing the closeness of those approximations.

To our knowledge, three previous papers have attempted to assess the accuracy of the predictions of simulation methods based on WTP. Fournier and Gai (2007) find the price increase predicted by a WTP-based merger simulation somewhat under-predicted the price effect estimated by a retrospective analysis. May and Noether (2014) compared the predictions of WTP to the price effects of two hospital mergers (estimated retrospectively) and find that the merger predicted to have the larger price effect in fact had the smaller retrospective price effect.

The prior study that is most relevant to our paper is Garmon (2017). While the methodologies are different, the central objective of the two papers is the same. Both papers attempt to evaluate the performance of relatively modern methods for predicting hospital merger price effects. The present paper does this using a Monte Carlo simulation in which the predicted price effects from the simulation methods are compared to the true price effects generated by a theoretical model. In contrast, Garmon (2017) uses real-world data on twenty-eight consummated hospital mergers over the period 1997-2012 and compares the price effects predicted by the methods to retrospectively measured effects.[4]

In Section 7 below, we discuss the advantages and disadvantages of each approach. Here we simply compare the results of the two papers. Broadly speaking, the results are similar. Both papers find that the modern methods perform reasonably well, and perform much better than traditional methods based on market structure and concentration metrics.

While the broad conclusions are similar, the specific analyses are somewhat different. Like us, Garmon (2017) examines the performance of *WTP/Q* and *UPP*. (Garmon (2017) does not study the merger simulation method based on Brand (2013).) Garmon (2017) analyzes *WTP/Q* in two ways. The first is to simply calculate percent changes in the value of the *WTP* metric resulting from a merger, without estimating the relationship between price and *WTP/Q*. He adopts a 6% change in *WTP* as a threshold for flagging whether the merger caused a price increase of any magnitude. The second is to estimate the relationship between price and *WTP/Q* to predict the price effect of a merger similar to the method based on CDS discussed above. The analysis of *UPP* in Garmon (2017) is broadly similar to ours.[5]

Garmon (2017) finds that the first of his two *WTP/Q*-based analyses and *UPP* perform much better than traditional methods based on market structure and concentration metrics. However, he finds that the second of his *WTP/Q*-based analysis (merger simulation) performs poorly, which is contrary to our results using that same method. But, as discussed in Garmon (2017), this result may be due to an important data limitation. Specifically, he does not use data on actual hospital prices, as these data are not publicly available. Instead, he generates hospital prices using financial information contained in the CMS Healthcare Cost Report Information system (HCRIS). This may introduce substantial measurement error and also reduces the amount of cross-sectional variation in the regression model. However, this limitation may not apply in real-world merger analyses

[4] Another distinction between these papers is that Garmon (2017) analyzes methods that can be implemented using data sources that generally are publicly available, whereas the methods we consider require data that would likely be available only in the context of an antitrust investigation.

[5] As discussed in Section 6, the main distinction between the analyses of *UPP* is that we adopt a 5% threshold as a flag for whether a merger caused a price increase of at least 5%, and Garmon (2017) adopts a 4% threshold as a flag for whether a merger caused a price increase of any magnitude.

as data on actual hospital prices are often available in the context of an antitrust investigation. Another limitation is that Garmon (2017) assumes that the relationship between price and WTP/Q is constant across all MSAs within a state. In contrast, in our simulations, the relationship between price and WTP/Q varies significantly across simulated markets. Hence, we do not view our WTP/Q results as necessarily contradictory to those of Garmon (2017), because these limitations may explain the differences.[6]

3 Theoretical Model

In this section, we present our theoretical model, which explicitly incorporates the aforementioned important features of hospital markets. Consumers do not directly purchase inpatient hospital care but rather access such care by purchasing health insurance either as an individual or through a group. Consumers who utilize inpatient hospital care choose their most preferred hospital from among the hospitals in their insurer's hospital network. Hospital prices are determined via simultaneous bilateral Nash bargaining between hospitals and insurers, and premiums are set via a Bertrand game among the insurers. The model can be solved for any given hospital market structure and then solved again for any alternative market structure in which two or more hospitals or hospital systems have merged. This generates the true price effects to which the price effects predicted by the simulation methods will be compared.

Each simulated market consists of a set of insurers M and a set of hospitals J. The hospitals in J are randomly assigned into a set of systems S. Some hospital systems consist of a single hospital. We use m as a general index for insurers, j as a general index for hospitals, and s as a general index for hospital systems. When referring to a specific insurer, we use m; when referring to a specific hospital, we use k; and when referring to a specific hospital system, we use t.

Each system bargains with insurers on an all-or-nothing basis. Each hospital j produces care at a constant cost c_j per admission. An agreement between insurer m and hospital j consists of a linear per-admission price p_{jm}. Each insurer m sells a single insurance product consisting of access to hospital network J_m and other attributes not related to inpatient care Z_m at a premium π_{J_m}. We assume that each insurer posts a single premium for the entire market. Insurers also incur a per event administrative cost τ. Each simulated market also includes a population of consumers indexed by i.

Given each insurer's network, its set of negotiated hospital prices, and the premium set by competing insurers, insurers choose their profit-maximizing premiums via a Bertrand pricing game. Hospital prices affect the profits of the insurers both directly as costs, and indirectly through the

[6]The sample of mergers available for a retrospective study is non-random; it includes only of mergers that were not blocked, and that were considered interesting enough to be studied. As discussed in Garmon (2017), this may introduce systematic bias into the evaluation of the merger simulation methods. No such bias exists in our analysis, as the simulated mergers are chosen exogenously.

equilibrium insurance premium. Hence, hospital price increases (decreases) are, in part, passed on to consumers in the form of higher (lower) premiums.

Consumers choose from insurers or go without insurance. At the time of the purchase decision, consumers face uncertainty over whether they will need inpatient care and in their preferences over hospitals. There is only one type of health condition that requires inpatient care, and consumer i utilizes inpatient care with probability $\rho_i \sim F(\rho)$. Conditional on seeking care, consumers are treated at their most preferred hospital in the insurer's network. We assume that consumers face no difference in their out-of-pocket expenses for inpatient care across hospitals.

3.1 Consumer Preferences and the Insurance Market

Consumer preferences over hospitals are defined as

$$U_{ij} = V_{ij} + \epsilon_{ij}, \forall j \in J, \tag{1}$$

where V_{ij} and ϵ_{ij} denote systematic and idiosyncratic components, respectively. Consumers who receive a draw of ρ_i that causes them to utilize inpatient care choose the hospital that provides the greatest utility given the realization of $\{\epsilon_{ij}\}_{j \in J}$. However, the uncertainty about the draws of ρ_i and $\{\epsilon_{ij}\}_{j \in J}$ are unresolved when individual consumers or groups choose their insurer.

We randomly assign each consumer to one of a set of buying groups G. This captures that fact that in the United States most consumers obtain health insurance through a buying group, often their employer. We assume that each insurance buying group has a single decision maker. We define the systematic component of decision maker's preferences as the arithmetic mean of the systematic component of preferences of the group's individual members. Hence, we define the utility of the decision maker for buying group g, consisting of consumers denoted by the set I_g, for insurer n with a network consisting of some set of hospitals J_n as

$$U_{gn} = Z_n - \theta \pi_{J_n} + \frac{\lambda}{\# I_g} \sum_{i \in I_g} \rho_i E_\epsilon \left[\max_{j \in J_n} \{V_{ij} + \epsilon_{ij}\} \right] + \zeta_{gn}. \tag{2}$$

Recall that π_{J_n} and Z_n denote the premium and the non-inpatient care attributes of insurer n, respectively. $\# I_g$ denotes the cardinality of the set I_g. The term inside the summation represents the expected utility for consumer i from having access to insurer n's hospital network. This is defined as the expected value of the utility from the ex-post most preferred hospital, times the probability of requiring inpatient care ρ_i. The parameter λ scales the expected utility that the decision maker gets from the insurer's hospital network into the utility that they receive from choosing that insurer. Similarly, the parameter θ translates the insurer's premium into the utility that they receive from choosing that insurer. ζ_{gn} denotes a single idiosyncratic draw for the decision maker that is assumed to be unknown to all agents when the bargaining between hospitals and insurers takes place.[7]

Applying the closed form to the consumer's expected utility, and assuming that ζ_{gn} is a Type I Extreme Value draw, the probability that buying group g will choose to buy the insurance product of insurer n is

$$\Lambda_{gn}(\pi_{J_n}) \equiv \frac{\exp\left\{Z_n - \theta\pi_{J_n} + \frac{\lambda}{\#I_g}\sum_{i\in I_g}\rho_i Emax_{iJ_n}\right\}}{1+\sum_{m\in M}\exp\left\{Z_m - \theta\pi_{J_m} + \frac{\lambda}{\#I_g}\sum_{i\in I_g}\rho_i Emax_{iJ_m}\right\}}, \tag{3}$$

where, under the assumption that $\{\epsilon_{ij}\}_{j\in J}$ are IID Type I Extreme Value draws,

$$Emax_{iJ_m} \equiv \ln \sum_{j\in J_m} \exp\{V_{ij}\}. \tag{4}$$

We use $\Lambda_{gn}(\pi_{J_n})$ to denote (3) when all hospital system-insurer combinations reach an agreement and $\Lambda_{gn}(\pi_{J_n=J\setminus k})$ to denote (3) when all hospital system-insurer combinations other than (k, n) reach an agreement.

Conditional on a vector of hospital prices, insurers play a Bertrand pricing game taking expectations over the distribution of both idiosyncratic components ζ_{gn} and ϵ_{ij}, as well as over ρ_i. The expected profits for insurer n are

$$\Pi_n^J \equiv \sum_g \Lambda_{gn}(\pi_{J_n})\left(\#I_g(\pi_{J_n} - p_z) - \sum_{i\in I_g}\rho_i \sum_{j\in J_n}\sigma_{ij}^{J_n}(p_{jn}+\tau)\right), \tag{5}$$

where $\sigma_{ij}^{J_n}$ denotes the probability that, conditional on needing inpatient care, and given that the consumer's choice set consists of J_n, consumer i would choose hospital j. We assume that no consumer uses an out-of-network hospital, i.e., $\sigma_{ij}^{J\setminus j} = 0\ \forall i, j$. We assume that insurer n maximizes (5) with respect to π_{J_n}.

[7] The number of consumers in each buying group ranges from one to more than 800. (See Appendix A1.) This raises the question of how to scale the insurance choice problem by the number of consumers in the insurance buying group. We assume that the insurance choice is made by a single decision maker on behalf of the group, so that equation (2) has the same scale irrespective of the number of consumers in the group. We further assume that the decision maker weighs the preferences of each consumer in the buying group equally. (Each consumer in the group receives Z_n utils from the non-inpatient care attributes of insurer n and $-\theta\pi_{J_n}$ utils from insurer n's premium. Since every consumer has the same value of Z_n and of $\theta\pi_{J_n}$, these have the same effect regardless of the decision maker's weighting across consumers. In contrast, there is heterogeneity across consumers in the value of the hospital network $\rho_i E_\epsilon\left[\max_{j\in J_n}\{V_{ij}+\epsilon_{ij}\}\right]$, so it is for this term that the assumption that the decision maker values each consumer in the group equally is significant.) The fact that (2) has the same scale for every buying group regardless of the group's size means that the idiosyncratic term ζ_{gn} has the same distribution irrespective of the number of consumers in the insurance buying group. We do not assume that ζ_{gn} for an insurance buying group is an aggregation (e.g., a mean) of IID idiosyncratic draws for each individual consumer in the group. A mathematically equivalent approach would be to multiply the right-hand side of (2) (including ζ_{gn}) through by $\#I_g$, so that the utility of group g for insurer n would be the sum of the individual utilities of the group members.

For reasons that will be made clear below, it is necessary to solve an analogous profit-maximization problem for the case in which insurer n and hospital k do not reach an agreement, but every insurer network besides n contains all of the hospitals in J, and insurer n reaches an agreement with every hospital in $J\backslash k$. This is done for each insurer-hospital pair, so we must solve for profit-maximizing premiums for the case where all negotiations succeed (i.e., all hospitals in J are included in the network of every insurer in M), and also for each case where exactly one negotiation fails. We use $\Pi_n^{J_n}$ to denote the value of the equilibrium profit for insurer n conditional on its network J_n, assuming that all insurers other than n include all hospital in J.

3.2 Bargaining

Equilibrium prices and network configurations are determined through a set of Nash Bargain equations. Each hospital system in S has a separate negotiation with each insurer in M. Negotiations proceed under standard Nash assumptions: (i) all negotiations occur simultaneously; (ii) no party to any negotiation observes or is in any way affected by what happens in any of the other negotiations; (iii) both parties to each negotiation believe that all the other negotiations will be successful (i.e., that all other hospitals will be included in all insurers' networks), and these beliefs turn out to be correct in equilibrium; and (iv) both parties to each negotiation have beliefs, which also turn out to be correct in equilibrium, about the prices agreed to in the other negotiations. In addition, we assume that all hospital systems and insurers have beliefs, that turn out to be correct in equilibrium, about the premiums, both with and without an agreement, that would emerge from the Bertrand game played by insurers.[8]

For ease of notation, we define the expected number of patients insured by insurer n treated by hospital k under the set of network configurations $J_m = J, \forall m \in M$ as

$$q_{kn} \equiv \sum_g \Lambda_{gn}(\pi_{J_n}) \sum_{i \in I_g} \sigma_{ik}^J \rho_i.$$

Similarly, we define the expected number of patients insured by a different insurer m treated by hospital k under the set of network configurations $J_m = J, \forall m \in M\backslash n$ and $J_n = J\backslash k$ (i.e., when all hospitals are in each insurer's network except that insurer n and hospital k fail to reach an agreement) as

$$q_{k(m\backslash n)} \equiv \sum_g \Lambda_{gn}(\pi_{J_m=J}, \pi_{J_n=J\backslash k}) \sum_{i \in I_g} \sigma_{ik}^J \rho_i.$$

[8]Note that when prices and premiums are determined, hospitals and insurers are taking expectations over three sources of uncertainty: which consumers will purchase insurance from insurer n (ζ_{gn}); which consumers will seek inpatient care (ρ_i); and which hospitals those consumers will choose (ϵ_{ij}).

Let $\vec{p_n}$ denote the vector of prices negotiated by insurer n with the $\#J$ hospitals $\{p_{1n}, ..., p_{\#Jn}\}$. Given this notation, the Nash bargaining objective function between hospital system t (which could be comprised of a single hospital) and insurer n is

$$NB_{tn} \equiv \left(\sum_{m} \sum_{k \in t} q_{km} \left(p_{km} - c_k\right) - \sum_{m \in M \setminus n} \sum_{k \in t} q_{k(m \setminus n)} \left(p_{km} - c_k\right) \right)^{\alpha} \left(\Pi_n^J \left(\vec{p_n}\right) - \Pi_n^{J \setminus s} \left(\vec{p_n}\right) \right)^{1-\alpha} . \quad (6)$$

The payoff of hospital system t if an agreement is reached with insurer n, given the outcomes of the other bargaining games with other insurers, is given by $\sum_m \sum_{k \in t} q_{km} \left(p_{km} - c_k\right)$. The disagreement payoff of hospital system t is given by $\sum_{m \in M \setminus n} \sum_{k \in t} q_{k(m \setminus n)} \left(p_{km} - c_k\right)$. In the special case of a monopoly insurer, the disagreement payoff of hospital system t is zero.

Note that if no agreement is reached, system t would expect to recapture some of the patients it would have treated under an agreement with insurer n because some consumers will switch insurers as a result of the exclusion. ($q_{k(m \setminus n)} \geq q_{km}$ must be true.) This highlights the important point discussed in Balan and Brand (2014), Peters (2014), and Ho and Lee (2017), that when a hospital fails to reach an agreement with a given insurer, it does not necessarily lose all of the patients that it was receiving from that insurer. The hospital only loses those patients who do not value it enough to switch insurers to retain access to it.

We employ the standard Nash-in-Nash solution concept, meaning a Nash equilibrium of a set Nash Bargaining equations.[9] Specifically, the equilibrium negotiated price for hospital k maximizes the weighted product of the increase in hospital k's payoff (compared to no agreement) and the increase in the insurer's payoff (compared to no agreement) if an agreement is reached. The weighting is defined by the parameter $\alpha \in (0, 1)$, which denotes the share of joint surplus that is captured by hospitals. This parameter could capture, for example, different rates of time preference or the relative skill of the negotiators involved in the bargaining.

3.3 Hospital Mergers

We now turn to the question of how mergers between hospitals (or hospital systems) affect equilibrium prices. To make the intuition as clear as possible, we begin by discussing a merger between two independent hospitals k and k'. However, everything in this discussion applies generally to mergers between hospital systems. We begin our discussion with a stylized intuitive explanation of the basic mechanism by which hospital mergers affect prices. We follow this with a discussion of some additional effects.

[9]See Collard-Wexler et al. (2017) for a discussion of the theoretical justifications for using the Nash-in-Nash solution concept.

Assume that the merged entity bargains on an all-or-nothing basis, meaning that the insurer either will have both of the merged entity's hospitals in its network or will have neither of them.[10] In the negotiation between a hospital and an insurer, each side has some bargaining leverage. By leverage, we refer to how much each side will lose if an agreement is not reached, which is measured by the difference between its equilibrium payoff and its disagreement payoff.[11] The leverage of the insurer comes from the fact that hospitals want access to that insurer's enrollees, and is greater when the insurer has more enrollees. The leverage of the hospital comes from the fact that its absence from the insurer's network makes that network less attractive to potential enrollees, which reduces the insurer's gross profit. This is greater when the hospital is strongly preferred by many enrollees. The effect of a merger between two hospitals k and k' will depend on how the merger changes the *relative* bargaining leverage of the two sides.

First suppose that k and k' are not substitutes at all (i.e., the diversion ratios between them are zero). After the merger, failure to reach a deal is more damaging to the insurer than it was before, as it means losing both hospitals from its network instead of one. Failure to reach a deal is also more damaging to the hospitals than it was before, as it means that they both will lose access to that insurer's subscribers instead of just one of them. But when the hospitals are not substitutes, this increase in damage is symmetric. The stakes have increased by the same proportion for both sides, so the relative bargaining leverage, and hence the negotiated prices, are unchanged.

Now suppose instead that k and k' are substitutes (i.e., the diversion ratios between them are positive). In this case, some patients whose first choice is k will have k' as their second choice, and vice-versa. This means that, before the merger, the unattractiveness of an insurance network that lacks one the hospitals, and hence the damage to the insurer's gross profits, is mitigated by the inclusion of the other. This mitigation is larger when the hospitals are closer substitutes and when non-merging hospitals are more distant substitutes.

After the merger, failure to reach an agreement means losing *both* hospitals from the insurer's network. Absent an agreement with the merged entity, patients whose first and second choices are k and k' will have to use their (less desirable) *third* choice hospital instead. The reduction to the insurer's gross profits from losing the merged entity from its network will be greater than the sum

[10] As discussed in Farrell et al. (2011), Balan and Brand (2014), and Gowrisankaran et al. (2015), hospital mergers may increase prices if hospitals within systems bargain separately, and the circumstances under which any particular merger is likely increase prices (e.g., high diversion ratios and high pre-merger hospital gross margins) are similar under either bargaining mode. Under separate bargaining, the source of the price effect is the familiar recapture of lost profits concept. After the merger, each hospital takes into account the fact that its merger partner will recapture some of its lost patients, and the associated profits, if it fails to reach an agreement. Balan and Brand (2014) show that the effect of a merger under separate bargaining can be larger or smaller than the effect under all-or-nothing bargaining. We assume all-or-nothing bargaining here because it appears to be the more commonly adopted bargaining mode in the real world.

[11] To be clear, the notion of "leverage" discussed here is distinct from the division of the joint surplus from an agreement, which is governed by the parameter α in our theoretical model. Throughout, we assume that mergers have no effect on this parameter. The possibility that mergers may have an effect on this parameter is examined in Lewis and Pflum (2017) and Lewis and Pflum (2015).

of the pre-merger reductions from losing the hospitals individually. In contrast, the reductions in gross profit to the hospitals from failing to reach an agreement will be the same as before; the reduction in profit for the merged entity from not having access to that insurer's patients is still equal to the sum of the reductions in profits for the hospitals individually. Since one effect is larger and the other is the same, the relative bargaining position has shifted in favor of the hospitals, and so the negotiated price will increase.

This intuition is reflected in the post-merger bargaining problem between insurer n and the merged entity $\{k, k'\}$, which is analogous to the pre-merger bargaining problem in (6)

$$\max_{\{p_{kn}, p_{k'n}\}} \left[\left(\sum_{j \in \{k,k'\}} \left(q_{jn} \left(p_{jn} - c_j \right) - \sum_{m \in M \backslash n} \left(q_{j(m \backslash n)} - q_{jm} \right) \left(p_{jm} - c_j \right) \right) \right)^{\alpha} \left(\Pi_n^J - \Pi_n^{J \backslash k,k'} \right)^{1-\alpha} \right]. \tag{7}$$

As noted above, the effect of the merger on the equilibrium values of $\{p_{kn}, p_{k'n}\}$ is manifested in changes in the disagreement payoff of the insurer. Specifically, the sign of the merger's effect on price will be the same as the sign of the difference between the reduction in profit to the insurer from failing to reach an agreement with $\{k, k'\}$ versus the sum of the reductions in profits from failing to reach an agreement with k' and k' individually,

$$\Pi_n^J - \Pi_n^{J \backslash k,k'} - \left(\Pi_n^J - \Pi_n^{J \backslash k} \right) - \left(\Pi_n^J - \Pi_n^{J \backslash k'} \right). \tag{8}$$

Rearranging terms, we see that the condition for a price increase resulting from the merger is

$$\Pi_n^J - \Pi_n^{J \backslash k} < \Pi_n^{J \backslash k'} - \Pi_n^{J \backslash k,k'}. \tag{9}$$

This expression defines a concavity condition, which captures the above intuition that losing two substitute hospitals reduces the insurer's profits by more than the sum of the individual reductions. Put another way, the presence of k' in the network of insurer n reduces the value-added of k to the network of n and vice-versa. Hence, an agreement between k' and insurer n creates a negative externality in the bargaining between k and insurer n. A merger between k and k' eliminates that externality and, therefore, will cause a price increase.

The above discussion was simplified in order to articulate the basic mechanism by which a merger of competing hospitals causes equilibrium negotiated prices to increase. However, there are a number of additional effects, to which we now turn.

The discussion above implicitly assumed that each insurer has a fixed pool of subscribers, and that the exclusion of a hospital from that insurer's network would deprive that hospital of all of those enrollees. But it is possible that failure to reach an agreement with a particular hospital will cause some subscribers to switch to an insurer that does have that hospital in its network, so some of the patients that the hospital loses from failing to reach an agreement with that insurer will be recaptured via another insurer. This affects the bargaining between hospitals and insurers, as now failure to reach a deal with an insurer does not deprive that hospital of access to all of that insurer's patients, but only to those patients who will not switch insurers in order to retain access to it.

The possibility of switching insurers can introduce additional merger effects. As discussed in Peters (2014), when insurer switching is possible, a merger can affect the hospital payoffs as well as the insurer payoffs. Specifically, if some patients switch insurers in response to a hospital exclusion, then the hospital will recapture some of the patients that it would otherwise have lost. Peters (2014) shows that, all else equal, a merger tends to increase the number of recaptured patients, which amplifies the price effect of the merger.[12]

The possibility of switching insurers can also introduce a complements effect that works in the opposite direction. This effect dampens the price effects of mergers, and can even make them negative, even when the merging hospitals are substitutes for individual patients. (Note though, that in our model, a merger that reduced prices would also reduce the profits of the merging hospital systems.) Peters (2014) shows that in the context of Nash Bargaining, this effect arises from the presence of enrollees who will switch insurers if *either* of the merging hospitals is excluded from that insurer's network. While we cannot decompose the true price effects into the price-increasing substitutes effect and the price-decreasing complements effect, the complements effect is seldom the dominant one as long as the merging hospitals are at least moderately close substitutes. For example, of mergers with a weighted mean diversion ratio that exceeds 10%, only 2.2% result in a price decrease. Of these, the price effect is less than 1% in magnitude in 86% of the cases. For substantially higher diversion ratios, the percentage of mergers with a negative price effect becomes extremely small.[13]

Vistnes and Sarafidis (2013) and Dafny et al. (2017) point out that group purchasers of insurance and/or common insurers across many purchasers of insurance can may cause the negative externality defined in (9) to be greater than what direct substitution at the patient level would suggest. If so, this would tend to amplify price effects and also to allow for the possibility of a positive price effect

[12] Note that all else will generally not be equal. Forces that tend to increase insurance switching, such as greater insurer competition, also affect the insurer payoffs. In our simulations, the net effect of greater insurance competition on price effects is generally negative, not positive.

[13] There are other theoretically possible sources of complements effects. One is if the exclusion of both merging hospitals, but not either of them alone, would drive an insurer out of business. This outcome does not occur in our simulations, as all insurers have positive margins even when the most valuable system is excluded. Another is the mechanism discussed in Katz (2011), namely that losing one hospital from the first-choice insurer's network may cause some enrollees to drop insurance altogether rather than switching to another insurer. This imposes a negative externality on substitute hospitals, because those lost enrollees had some positive probability of using the substitute hospital had they remained insured. The merger eliminates this externality, which tends to reduce prices. This effect is present in our model, but is minimal for anything other than the monopoly insurer case, as in our results very few people are uninsured when there is more than one insurer. See Balan (2017) for a more complete discussion of complements effects.

even for a merger of hospitals that are not substitutes for any individual patient (i.e., with diversion ratios of zero between the merging hospitals).

While we do not focus our analysis on these additional effects, some of the key features discussed in this literature (e.g., recapture through switching insurers and the group purchase of insurance) are included in our theoretical model. It would be possible to modify our theoretical model to further explore these effects. We did not make these modifications, since that is not the purpose of this paper.

We assume a non-linear parametric function (specifically Logit) for insurance demand, which must have a convex and a concave region. Since insurance demand is derived by summing, across each purchaser of insurance, the relationship between the utility derived from the insurers network and the probability of purchasing from that insurer, these relationships must each have a convex and a concave region as well. Several of the effects discussed above operate by influencing the sizes and shapes of these regions, making some portions more or less convex or concave. Moreover, the functional form restriction itself can magnify or dampen these effects. For example, an effect that makes the relationship more concave in one region may mechanically make it more convex in another, and this can tend to dampen or amplify the effects discussed above.

3.4 Simplifications

We make two simplifying assumptions in our model in order to reduce computational expense. First, we assume that each hospital system negotiates a single price for all of its member hospitals. This is in addition to our assumption discussed above that hospital systems negotiate on an all-or-nothing basis.

Second, we assume symmetric competition in the health insurance market. We do so only because computing the equilibrium with an asymmetric M-firm oligopoly given the population size we use in our simulations is computationally expensive, and the symmetry assumption greatly reduces the burden. While this is a departure from what is commonly observed in the real world, we still capture the effect of differing levels of competition in the health insurance market on the bargaining incentives of hospitals and insurers. As discussed below, we do this by varying the number of (symmetric) insurers in the market.

Given these simplifications, we define the equilibrium price vector $\vec{p}^*$ as the set of $\#S$ prices that simultaneously solves the system of equations

$$
\begin{aligned}
\left.\frac{\partial \ln NB_1(\vec{p})}{\partial p_1}\right|_{\vec{p}^*} &= 0 \\
\left.\frac{\partial \ln NB_2(\vec{p})}{\partial p_2}\right|_{\vec{p}^*} &= 0 \\
&\vdots \\
\left.\frac{\partial \ln NB_{\#S}(\vec{p})}{\partial p_{\#S}}\right|_{\vec{p}^*} &= 0.
\end{aligned} \tag{10}
$$

This equilibrium price vector is common across all insurers because we assume symmetric competition in the insurance market. Note that this definition conditions on the insurer's (common) equilibrium profit maximizing premium π^*, and also on the off-equilibrium profit maximizing premiums under hypothetical exclusions. There are $\#S$ such premiums if there is a monopoly insurer, one for each excluded hospital system. There are $2(\#S)$ such premiums if there is an oligopoly in the insurer market, one for each excluded hospital system for the insurer that excludes, and another (common) premium for each of the other insurers, all of which do not exclude. All of these premiums, together with hospital prices, are solved for simultaneously. See Appendix A6 for details on computing the equilibrium price vector $\vec{p}^*$ and equilibrium premium π^*.

4 Parameterization

In this section, we provide a brief summary of our parameterization of the theoretical model. We provide additional detail in Appendix A1.

We create 9,000 simulated hospital markets. We chose this number of markets because it is large enough to generate a rich set of parameterizations while still being computationally feasible. Each market consists of 500,000 consumers, twelve hospitals, and specific values of the model parameters. Each of the 500,000 consumers is characterized by a randomly generated location, risk type ρ_i, and assignment into one of 60,000 insurance buying groups. Each of the twelve hospitals is characterized by a randomly generated location, and by a quality η_j and a marginal cost c_j which are generated as discussed in Appendix A1. For each market, we randomly draw a number of hospital systems $\#S$ which we fix to be in the set $\{5, 6, ..., 10\}$ and randomly assign each of the twelve hospitals into one of the $\#S$ systems.

The parameters of the model include the Nash bargaining split parameter α from (6); the travel cost parameters[14]; the parameters governing consumer preferences over insurers from (2): θ, λ, and Z_m; the insurers' administrative cost τ from (5); the mean and variance of the hospital quality

[14] See equation (A2).

distribution; and the number of insurers. In each simulation, we randomly assign the value of each of these parameters from a set of three possible values, except for the number of insurers, which is drawn from the set {1,3,5,7,9}.

The primary criterion used in selecting the range of values for these parameters is that they generate output that corresponds to real-world levels for important metrics. One such metric is the pseudo-R^2 values from the estimation of the discrete choice model, which is used to calculate diversion ratios and WTP. We select the travel cost parameters, the variances of the distributions determining the locations of consumers and hospitals, and the variance of the distribution of hospital quality so that the pseudo-R^2 matches the values commonly found in real-world experience using hospital discharge data, which are typically in the range (0.40, 0.55). That is, we choose parameter values so as to ensure that the extent to which consumer choices of hospitals are determined by idiosyncratic component of preferences (as opposed to systematic attributes of hospitals or consumers) in our simulations roughly matches that in real- world analyses. Given the values of these parameters, we select the mean of the hospital quality distribution to ensure that, even under the highest values of the travel cost parameters, almost all consumers place positive valuation on the hospital network. (Recall that travel cost creates negative utility over hospitals for consumers.) This ensures that higher risk consumers (those with larger values of ρ_i) are more likely to purchase health insurance than are lower risk consumers. (See equation (2).)

The other key metrics are hospital costs, prices, and gross margins. We set the values of the remaining parameters so that, on average, these match real-world data. We base our price and margin benchmarks on two sources. First, Health Care Cost Institute (2015) reports that the average hospital reimbursement for patients with employer sponsored health insurance in 2014 was \$18,338. Second, Ramanarayanan (2014) reports that hospital contribution margins, which are analogous to our definition of gross margin, are typically around 50%. Given this information, we set the mean value of hospitals' marginal cost c_j to \$8,000 and select values of the remaining model parameters to produce wide variation: (i) in hospital prices about a mean in the \$18,000-\$19,000 range; and (ii) in hospital gross margins about a mean of 50%.

While useful, these metrics provide only rough guidance for our choice of parameter values. There are several reasons for this. First, available data do not allow us to measure those real-world metrics with certainty. Second, different combinations of parameter values can generate similar values for those metrics, and our results may be sensitive to different combinations of parameter values that yield similar values of them. Third, because we are uncertain which parameterizations correspond most closely to the real world, we use a broad range of parameter value in order to make it more likely that our analysis covers the most relevant parameterizations. Fourth, real-world markets may exhibit significant heterogeneity.

For these reasons, we use a wide range of parameter values across our different simulations, which causes many of our markets to have mean hospital gross margins that are well above or below 50%. Table 1 lists percentiles of the mean (within market) hospital gross margin across our 9,000 markets. The mean hospital gross margin is 0.492.

Table 1: Percentiles of Within-Market Mean Hospital Gross Margins

10^{th}	25^{th}	50^{th}	75^{th}	90^{th}
0.260	0.362	0.499	0.624	0.710

In our main results, we aggregate our performance across all of these parameterizations. However, we provide results broken down by specific parameter values in Appendix A5.2 in order examine how the performance of the simulation methods varies across different values of the model parameter. Table A2 in Appendix A1 provides a list and description of each of the model parameters.

5 The Merger Simulation Methods

In this section, we detail the merger simulation methods. We begin with *Willingness-to-Pay* (WTP) as described in CDS.[15] WTP is a measure of the value-added of a hospital or hospital system to the provider network of an insurer. It is straightforward to compute using standard methods developed in the discrete choice literature. To understand the intuition, consider again the general model of consumer preferences over hospitals in (1). As noted above, WTP measures the difference in expected utility of consumers, prior to the realization of $\{\epsilon_{ij}\}_{j\in J}$, between the provider network of the consumer's insurer and that same network but excluding one hospital system. Given the assumptions that: (i) the consumer chooses the hospital from among their insurer's provider network that provides the greatest utility given the realization of $\{\epsilon_{ij}\}_{j\in J}$; and (ii) $\{\epsilon_{ij}\}_{j\in J}$ are IID draws from the Extreme Value distribution, the expected utility of consumer i for provider network J_n has the familiar closed form

$$E_\epsilon\left[\max_{j\in J_n}\{V_{ij}+\epsilon_{ij}\}\right]=\kappa+\ln\sum_{j\in J_n}\exp\{V_{ij}\},$$

where κ denotes Euler's constant. Given this definition, the value-added of hospital system t for consumer i, assuming that insurer n has each of the other hospital systems in its provider network, is

[15] Although CDS were the first to apply the term WTP in this context, the measure developed by Town and Vistnes (2001) is very similar. The differences between the two models are irrelevant for our study. Here, we focus on the CDS exposition.

$$
\begin{aligned}
WTP_{it} &= E_{\epsilon}\left[\max_{j\in J}\{V_{ij}+\epsilon_{ij}\}\right] - E_{\epsilon}\left[\max_{j\in J\setminus t}\{V_{ij}+\epsilon_{ij}\}\right] \\
&= \ln\left(\frac{1}{1-\sigma_{it}}\right), \qquad (11)
\end{aligned}
$$

where $\sigma_{it} \equiv \sum_{j\in t}\exp\{V_{ij}\} / \sum_{j\in J}\exp\{V_{ij}\}$. This defines the probability that consumer i will choose one of the hospitals in system t.

As defined in CDS, the total WTP for hospital system s is evaluated by integrating WTP_{it} over the joint distribution of consumer characteristics (demographic and clinical) and multiplying by the sample size. This may be approximated by summing (11) across individuals. Hence, the WTP for hospital system t is

$$
WTP_t = \sum_i \ln\left(\frac{1}{1-\sigma_{it}}\right). \qquad (12)
$$

CDS define the change in WTP due to a merger as the difference between the WTP of the merged entity and the sum the pre-merger values of WTP. Hence, for a merger between hospital systems t and t', the change in WTP is

$$
\Delta WTP_{t+t'} = \sum_i \left[\ln\left(\frac{1}{1-\sigma_{it}-\sigma_{it'}}\right) - \ln\left(\frac{1}{1-\sigma_{it}}\right) - \ln\left(\frac{1}{1-\sigma_{it'}}\right)\right]. \qquad (13)
$$

This has the property that the change in market power due to the merger is close to zero if consumers do not view t and t' as substitutes. Specifically, (13) can be made arbitrarily small if, $\forall i$, either σ_{it} or $\sigma_{it'}$ is sufficiently small. This implies that changes in WTP are increasing in the extent to which consumers view the merging hospital systems as substitutes and that changes in WTP necessarily approach zero as this substitutability approaches zero.

We test two merger simulation methods based on least squares regressions in which WTP is the key explanatory variable. First, we apply the regression model presented in Farrell et al. (2011), which is a modified version of the regression model presented in CDS. Based on intuition derived from the Nash bargaining framework, CDS hypothesize that the WTP of a hospital or system is proportional to the incremental gross profit (gross of payments to hospitals) of the insurer under the agreement with the hospital or system. Given this, CDS regress hospital profits on WTP. However, a regression framework that uses price, as opposed to hospital profits, as the dependent variable may be preferable under some circumstances.[16] As summarized in Farrell et al. (2011),

[16]For example, the researcher may have access to insurer claims data, which can be used to generate reliable measures of price. However, the available financial data may be insufficient to generate a measure of incremental

an appropriate modification of the CDS regression model in this circumstance would be to regress prices on WTP on a per expected discharge basis and marginal cost. Hence, the first simulation method we evaluate is based on the least squares regression model

$$p_t^* = \beta_0 + \beta_1 WTP_t/q_t + \beta_2 c_t + \nu_t, \tag{14}$$

where p_t^* denotes the equilibrium price of hospital system t (the t^{th} element of $\vec{p}^*$ defined in (10)), and c_t denotes volume-weighted marginal cost of system t, respectively. q_t denotes the expected volume of system t, and ν_t denotes an econometric error. (p_t^*, c_t, and q_t are data that would be observed by a real-world analyst.) β_0, β_1, and β_2 are reduced-form parameters to be estimated. We refer to this regression model as the WTP/Q simulation method.

Second, we test a merger simulation method developed in Brand (2013) that extends the CDS WTP framework by incorporating additional components of bargaining theory. Among other things, this alternative approach predicts the change in equilibrium prices due to a merger accounting for feedback effects between the merging hospitals and through third party hospitals. Specifically, it incorporates the intuition that since hospital prices are determined jointly in equilibrium, the price for each hospital system should reflect not just its own cost and WTP, but also the cost and WTP of each hospital with which it competes. For example, all else equal, a hospital that faces high priced rivals will have a higher equilibrium price than if it faced lower priced rivals, and vice versa. In principle, the empirical model derived from this approach should provide a better approximation to (6) compared to the WTP/Q simulation method.

We develop this method by considering a simplified bargaining framework in which, as assumed in the CDS framework, insurers are not explicitly modeled as profit maximizers. Rather, the payoff for each insurer in bargaining with hospitals is simply proportional to WTP minus payments to hospitals. Also as assumed in the CDS framework (an in contrast to our theoretical model), each insurer's enrollees are "captured" in the sense that if an insurer fails to reach an agreement with a given hospital, its enrollees cannot switch to a competing insurer. This assumption implies that the disagreement payoff for each hospital system is zero. The key distinction between this alternative approach and the CDS WTP framework is that this approach accounts for the fact that if an insurer fails to reach an agreement with a given hospital, its enrollees will be diverted to competing hospitals. We write this simplified bargaining problem between insurer n and hospital system t as

$$\left[q_{nt}\left(p_{nt} - c_t\right) - 0\right]^a \left[\Gamma_1 WTP_{nt} - \sum_{s \in S} q_{ns} p_{ns} + \sum_{s \in S \setminus t} q_{ns(t)} p_{ns}\right]^{1-a},$$

profit for specific hospital/insurer combinations. Moreover, while credible direct measures of the incremental cost of patient care may be very difficult to obtain, other variables that reliably proxy for cost may be available. In such a case, prices, as opposed to incremental profits, may be the preferable dependent variable.

where Γ_1 denotes the constant transformation from utils (as measured by WTP) into dollars for the insurer, and $q_{ns(t)}$ denotes the expected volume at system s from insurer n if n fails to reach an agreement with t. The parameter a denotes the division of the joint surplus in this simplified bargaining problem.[17] Maximizing with respect to p_{nt} yields

$$p_{nt} - c_t = \frac{a}{1-a}\left[\frac{\Gamma_1 WTP_{nt}}{q_{nt}} - p_{nt} + \sum_{s \in S \setminus t} d_{nts} p_{ns}\right],$$

where d_{nts} denotes the diversion ratio from system t to system s for insurer n.[18] (Since we assume symmetric competition in the insurance market, d_{nts} is the same across all insurers.) Stacking these equations across all hospital systems for a given insurer and solving for the price vector yields the system of equations

$$\vec{p} = D(a)^{-1}\left[\Gamma_1 \overrightarrow{WTP/q} + \frac{1-a}{a}\vec{c}\right], \tag{15}$$

where $\vec{p}$, $\overrightarrow{WTP/q}$, and $\vec{c}$ denote $\#S$ vectors of system-level prices, WTP divided by expected volume, and marginal cost, respectively. $D(a)$ denotes a $\#S$ x $\#S$ matrix in which $D(a)_{ss} = \frac{1}{a}, \forall s$ and $D(a)_{ts} = -d_{ts}, \forall s \neq t$.

Of course, the price vector on the left hand side of (15) is not equivalent to the equilibrium price vector $\vec{p}^*$ from the theoretical model defined in (10). Therefore, the right hand side of (15) will fit $\vec{p}^*$ with some error. This motivates the least squares regression of our second simulation method

$$\vec{p}^* = \Gamma_0 + \Gamma_1 D(a)^{-1}\overrightarrow{WTP/q} + \Gamma_2 D(a)^{-1}\vec{c} + \vec{\nu}, \tag{16}$$

where $\vec{\nu}$ denotes a vector of errors, and Γ_0, Γ_1, and Γ_2 denote parameters to be estimated.

We refer to this simulation method as the diversion-weighted WTP/Q method, or $DWTP/Q$. Note that changes in WTP or cost of any hospital system affects the prices of all hospital systems through the matrix $D(a)^{-1}$. That is, unlike the WTP/Q method or the regression model applied in CDS, the $DWTP/Q$ method captures feedback effects resulting from mergers between hospitals. In addition, the pre-merger prices of the merging hospitals and the magnitude of the price effect of the merger are influenced by the distribution of pre-merger prices across all hospital systems. Of the three simulation methods, only the $DWTP/Q$ method can account for these effects. Note that

[17] We use a here to avoid confusion with the parameter α which denotes the division of the joint surplus in our theoretical model (6).

[18] All else equal, price effects are larger when the merging firms' products are closer substitutes. Diversion ratios are an important and widely-used measure of the closeness of substitution. See, for example, the 2010 DOJ/FTC Horizontal Merger Guidelines (p. 21). The diversion ratio from hospital system t to hospital system s is the fraction of t's patients from a particular insurer that would choose s if t were excluded from that insurer's network. Hence, $d_{nts} \equiv \frac{q_{ns(t)} - q_{ns}}{q_{nt}}$.

the WTP/Q method can be recovered from the $DWTP/Q$ method under the assumption that the off-diagonal elements of $D(a)^{-1}$ are zero.

The bargaining weight parameter a is separately identified in the $DWTP/Q$ method, although non-linear estimation methods are required. However, our initial results suggested that the non-linear least squares estimator of a is highly unreliable. Hence, rather than estimating a in (16) using non-linear methods, we fix the value of a at $\frac{1}{2}$, and then estimate Γ_0, Γ_1, and Γ_2 using OLS. We assume $a = \frac{1}{2}$ because it seems to be a reasonable assumption absent any direct evidence about the true value of a. We maintain the assumption $a = \frac{1}{2}$ in (16) irrespective of the true value of α in our theoretical model, which as discussed in A1, we allow to take on values of 0.4, 0.5, or 0.6. That is, we assume that the real-world analyst applying the simulation method may make an incorrect assumption regarding the value of this parameter. We do this because this may be the most plausible assumption for the real-world analyst given the information available.

Finally, we turn to UPP. As described in Haas-Wilson and Garmon (2009) and Garmon (2017), the first order price effect of a merger between hospitals k and k' can be derived from a Nash bargaining model under the assumption that the merging hospitals do not bargain with insurers on an all-or-nothing basis post-merger, but rather each of the merged hospitals bargains separately with insurers. Then the first order effect of the merger on the equilibrium price of hospital k is given by

$$(1 - a)d_{kk'}(p_{k'} - c_{k'}), \tag{17}$$

where $d_{kk'}$ denotes the diversion ratio from k to k'. Similarly, the first order effect of the merger on the equilibrium price of hospital k' is given by

$$(1 - a)d_{k'k}(p_k - c_k). \tag{18}$$

As detailed below, we define the predicted price effect of merger based on the UPP method as the volume-weighted mean of these two terms. As with $DWTP/Q$, we assume that the analyst cannot estimate the true bargaining parameter a. Hence, in evaluating UPP, we assume $a = \frac{1}{2}$ irrespective of the true value of α in our simulations.

5.1 Predicted Price Effects of the Simulation Methods

After computing the pre- and post-merger equilibria in our theoretical model, we generate the simulation methods' predicted price effects of mergers as follows. We carry out the following steps for each simulated market. First, we resolve the three sources of uncertainty in our theoretical model: (i) which consumers will purchase health insurance (ζ_g); (ii) which consumers will seek inpatient care (ρ_i); and (iii) which hospitals will treat those consumers (ϵ_{ij}). This produces data on individual-level inpatient events that identifies the location of the patient and of the hospital that treated the patient. These data to which we apply the merger simulation methods are one realization from the joint distribution of possible outcomes. Such individual-level inpatient data, together with data on pre-merger hospital prices and marginal costs, comprise the data that would be available to a real-world analyst evaluating a proposed merger.

Next, we use the individual-level inpatient data generated in the first step to estimate a conditional logit model. This provides estimates of consumer preferences over hospitals. With the output of the conditional logit model, we construct WTP for each hospital system and the diversion ratios between all pairwise combinations of hospital systems.[19]

Finally, given the pre-merger prices and marginal costs from our theoretical model, and the values of WTP and diversion ratios, we estimate (14) and (16) for each of our simulated hospital markets. Using the output of these regression models, we apply the fitted relationship to the changes in WTP/Q and $D\,(1/2)^{-1}\,WTP/Q$ for each possible pairwise merger to generate the simulation methods' predicted price effects. We also calculate the predicted price effect for each possible pairwise merger for the UPP method using the estimated diversion ratios and the data on prices and marginal costs.

For the simulation methods WTP/Q, $DWTP/Q$, and UPP, the predicted price effect of a merger between hospital systems t and t' is defined as follows.

Predicted Price Effect of Simulation Method 1: *WTP/Q*

$$\widehat{\Delta p_{tt'}} = \widehat{\beta}_1 \frac{WTP_{tt'} - WTP_t - WTP_{t'}}{q_t + q_{t'}}, \tag{19}$$

where $\widehat{\beta}_1$ denotes the estimated coefficient on WTP/Q in (14).

Predicted Price Effect of Simulation Method 2: *DWTP/Q*

$$\widehat{\Delta p_{tt'}} = \widehat{\Gamma}_1 \frac{\sum_{s=1}^{\#S-1} D_{post}(a)^{-1}_{(tt')s} WTP_s - \sum_{s=1}^{\#S} \left(D_{pre}(a)^{-1}_{ts} + D_{pre}(a)^{-1}_{t's}\right) WTP_s}{q_t + q_{t'}}, \tag{20}$$

where $\widehat{\Gamma}_1$ denotes the estimated coefficient on $D(a)^{-1}WTP/Q$ in (16), $D(a)^{-1}_{t.}$ denotes the tth row of the diversion ratio matrix $D(a)^{-1}$, and $\#S$ denotes the pre-merger number of hospital systems in the market.

[19] Because we can evaluate the choice probabilities directly from the theoretical model, we could calculate WTP and diversion ratios directly without estimating a conditional logit model. However, this option is not available in a real-world application of the simulation methods. For this reason, we estimate WTP and the diversion ratios using the fitted probabilities, as a real-world analyst applying the simulation methods would do. This is a realistic source of error in predicting the price effects of mergers, and is therefore appropriately incorporated into our Monte Carlo experiment.

Note that in this expression, the diversion ratio matrix $D(a)$ differs pre- and post-merger. The rank of $D(a)$ is equal to the number of hospital systems in the market and so is reduced from $\#S$ to $\#S - 1$ under a merger between two hospital systems. This also implies that the off-diagonal elements of $D(a)$ must be re-evaluated for each merger in order to compute the predicted price effect of the merger.[20]

Predicted Price Effect of Simulation Method 3: *UPP*

$$\widehat{\Delta p_{tt'}} = \frac{q_t d_{tt'}(p_{t'} - c_{t'}) + q_{t'} d_{t't}(p_t - c_t)}{2(q_t + q_{t'})}, \tag{21}$$

where $d_{tt'}$ denotes the diversion ratio from t to t'.

We make two assumptions on the information possessed by our hypothetical analyst. First, we assume that the analyst observes hospital prices and marginal costs without error.[21] Second, we assume that the analyst knows the correct specification of the discrete choice model of consumer preferences over hospitals, though not the parameter values. In making these assumptions, our intention is not to downplay the potential importance of these topics in assessing the performance of the simulation methods in predicting real-world mergers. Rather, we make these assumptions simply to focus attention on our central question which is the closeness of approximation of the merger simulation methods to the theoretical model under a wide range of parameterizations.

6 Results

We begin by setting notation. Let r index a merger between a particular pair of hospital systems in a particular simulated market. Let d_r denote the volume-weighted mean diversion ratio between the merging hospital systems, and let p_r denote the volume-weighted mean pre-merger price of these hospital systems. Let Δp_r denote the price effect of merger r generated by our theoretical model, and let $\widehat{\Delta p_r}$ denote the predicted price effect of the same merger r generated by any one of the three simulation methods that we evaluate in this paper. These predicted price effects are defined in (19), (20), and (21).[22]

[20]Note that in (19) and (20) we ignore the change in cost due to the merger. This is because we define the cost of the merged system to equal the volume-weighted mean of the pre-merger systems' costs. Hence, there would be no effect of the merger on price through the cost terms. This definition assumes that there are no marginal cost efficiencies associated with any merger.

[21]In practice, hospital prices are typically estimated using data sources such as claims-level data and adjusted to account for varying casemix distributions across hospitals. These data may be measured with error. In Appendix A5.3, we test the robustness of our results to measurement error in prices and costs.

[22]While the focus of our discussion is on the predicted price effects generated by the simulation methods, here we briefly summarize the estimation results from the regression models underlying WTP/Q (see (14)) and $DWTP/Q$ (see

We present descriptive statistics of our simulated hospital markets for four categories of mergers grouped by the mean diversion ratio d_r. These categories are [0%,5%), [5%,10%), [10%,20%), [20%,30%), and [30%,100%]. We use the diversion ratio as our metric for categorizing mergers for two reasons. First, as discussed above, theory predicts that, all else equal, price effects of mergers are increasing in the diversion ratios between the merging hospitals, so categorizing mergers by diversion ratios is a rough way of categorizing them by degree of competitive concern. Second, diversion ratios are typically straightforward to estimate in real-world applications, as the necessary data are commonly available.[23]

Table 2 presents a summary of the merger price effects generated by our theoretical model expressed as a percentage of the pre-merger price $\frac{\Delta p_r}{p_r}$. We refer to this percentage change as the true price effect. These results are from our full set of 231,925 simulated mergers. The table includes the following summary statistics of the true price effect broken down by diversion ratio category: mean, standard deviation, and 10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th} percentiles. The mean true price effect across all mergers is 1.7%, while the median true price effect is 0.4%. The 10^{th}, 25^{th}, 75^{th}, and 90^{th} percentiles are -0.1%, 0.1%, 1.8%, and 5.0%, respectively.

As expected, the mean true price effect of mergers increases with d_r. For mergers such that $d_r < 5\%$, which constitute 52.5% of the mergers in our analysis, the mean true price effect is just 0.1% (median = 0.1%). The 10^{th}, 25^{th}, 75^{th}, and 90^{th} percentiles are -0.3%, 0.0%, 0.3%, and 0.5%, respectively. In contrast, for mergers such that $d_r \in [30\%, 100\%]$, which constitute 9.2% of the mergers in our analysis, the mean true price effect is 10.1% (median = 8.4%). The 10^{th}, 25^{th}, 75^{th}, and 90^{th} percentiles are 3.8%, 5.6%, 12.7%, and 18.5%, respectively.[24]

(16)). (Recall that UPP is not based on such a regression model.) For both methods, we find considerable variation across simulated markets in the regression coefficient of interest. For WTP/Q, the mean estimated value of β_1 is 2.54 and the standard deviation is 1.75. The 25th, 50th, and 75th percentiles are 1.22, 2.14, and 3.45. For $DWTP/Q$, the mean estimated value of Γ_1 is 5.87 and the standard deviation is 3.66. The 25th, 50th, and 75th percentiles are 3.11, 4.97, and 7.83. As we would expect, we find that the estimated values of β_1 and Γ_1 are higher when the value of α is higher, the value of λ is higher, or the value of θ is lower.

[23] As discussed in Section 5, the predictions of the merger simulation methods are, in part, determined by the diversion ratios between the hospitals. We calculate diversion ratios the way they would be calculated in real-world applications of those methods, using patient-level inpatient discharge data. Note that diversion ratios calculated in this way do not account for the possibility that some patients will switch insurers in order to retain access to Hospital A, or that they will drop their insurance entirely if Hospital A goes out of their preferred insurer's network. Our theoretical model does account for these possibilities, which is one important reason why the simulation methods are not *a priori* certain to closely approximate the theoretical model.

[24] The distribution of true price effects in our simulations is a result of a number of assumptions. One is that we generate the true price effect for each possible pairwise merger between hospital system. A different rule for determining the set of of hospital mergers would generate a different distribution of true price effects. Another assumption is the particular distribution of parameter values in our theoretical model. For example, we assume that key parameters, such as travel costs, λ, θ, and the number of insurers, are independently and uniformly distributed across markets. These assumptions are, of course, arbitrary. Hence, the distribution of true price effects in our simulations does not necessarily reflect the distribution of price effects resulting from real-world mergers. We present these results only to illustrate that our theoretical model produces the intuitive result that mergers between hospitals that are closer substitutes are likely to cause larger price effects. We emphasize that this problem is less severe in the context of evaluating the performance of the simulation methods, which is the primary purpose of this paper. The reason is that we condition on narrow categories of mergers defined by the true price effect (e.g., mergers in which the true

Table 2: True Price Effects of Mergers $\frac{\Delta p_r}{p_r}$ from the Theoretical Model

Mergers	N	Mean	Stan Dev	Percentiles				
s.t. $d_r \in$				10^{th}	25^{th}	50^{th}	75^{th}	90^{th}
All	231,925	0.017	0.037	-0.001	0.001	0.004	0.018	0.050
[0%,5%)	121,848	0.001	0.005	-0.003	0.000	0.001	0.003	0.005
[5%,10%)	35,167	0.008	0.007	0.003	0.005	0.008	0.012	0.015
[10%,20%)	35,091	0.019	0.013	0.007	0.012	0.018	0.026	0.034
[20%,30%)	18,415	0.041	0.022	0.018	0.027	0.039	0.053	0.068
[30%,100%]	21,404	0.101	0.070	0.038	0.056	0.084	0.127	0.185

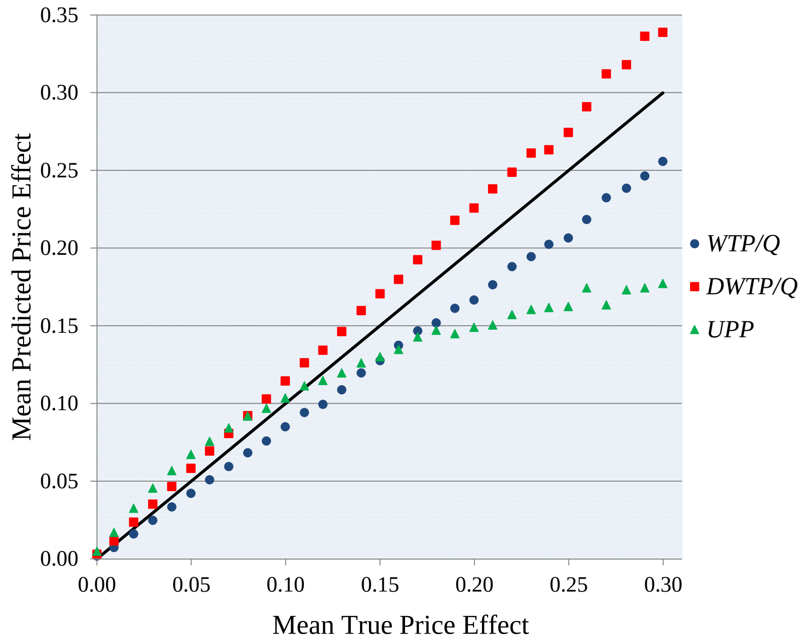
6.1 Bias of the Simulation Methods

We begin our evaluation of the simulation methods with an examination of their bias. Instead of grouping mergers by diversion ratio categories as in Table 2 above, we proceed by grouping our 231,925 mergers into 31 categories defined by one percentage point increments of the true price effect $\frac{\Delta p_r}{p_r}$ (i.e., $(\leq 0.5\%), (0.5\%, 1.5\%), (1.5\%, 2.5\%), ..., (29.5\%, 30.5\%), (\geq 30.5\%)$). We then compare, within each of these categories, the mean of the true price effect $\frac{\Delta p_r}{p_r}$ with the mean of the predicted price effect $\widehat{\frac{\Delta p_r}{p_r}}$ generated by WTP/Q, $DWTP/Q$, and UPP.

Figure 1 contains the scatter plot of the results. The x-axis indicates the true price effect, and the y-axis indicates the predicted price effect. For each of the 31 categories, a perfect simulation method would generate a dot on the solid 45^o line. The vertical distance between that line and the dots on the three colored curves represent the bias of the three simulation methods for that category. The figure indicates that WTP/Q exhibits a bias towards under-predicting the true price effects, while $DWTP/Q$ exhibits a bias towards over-predicting. UPP exhibits a bias towards over-predicting when the mean true price effect is low, but an increasing bias towards under-predicting as the true price effect increases. For example, in the category of mergers for which the true price effect is in (4.5%,5.5%), the mean predicted price effect is 4.2% for WTP/Q, 5.8% for $DWTP/Q$, and 6.7% for UPP. In the category of mergers for which the true price effect is in (14.5%,15.5%), the mean predicted price effect is 12.8% for WTP/Q, 17.1% for $DWTP/Q$, and 13.0% for UPP. We view these differences as indicative of only a moderate amount of bias. That is, for mergers that are similar to each other based on their true price effects, the simulation methods provide predicted price effects that, on average, are reasonably close to the true price effects.

price effect is between 4.5% and 5.5%). Changes in the distribution of parameter values may substantially affect the distribution of true price effects *across* these categories, but will likely have a smaller effect on the performance of the simulation methods *within* each category. However, any distribution of true price effects generated by an alternative distribution of model parameters must still satisfy the benchmark values of the metrics discussed in Section 4, namely the psuedo-R^2 from the discrete choice model and hospital gross margins.

Figure 1: Mean True and Predicted Price Effects

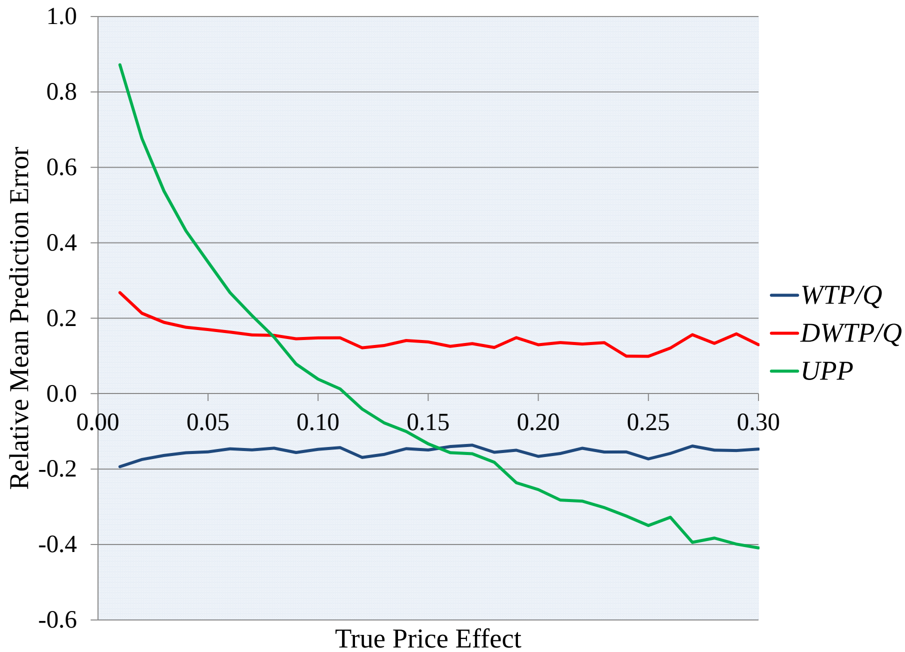


Using mean prediction errors in levels to measure the performance of the simulation methods, as we did above, ignores the fact that an acceptable magnitude for a prediction error may depend on the magnitude of the true price effect. For example, a 2% mean prediction error may be more acceptable for a merger with a true price effect of 10% than for one with a true price effect of 5%. For this reason, we next evaluate the mean prediction errors of the simulation methods as a percentage of the mean true price effect. We refer to this measure as the relative mean prediction error.

Figure 2 plots the relative mean prediction errors for WTP/Q, $DWTP/Q$, and UPP within the same 31 merger categories used in Figure 1, namely grouping mergers by one percentage point increments of the true price effect. The results indicate that the relative mean prediction error for WTP/Q and $DWTP/Q$ is quite stable across categories of mergers, particularly for categories in which the mean true price effect is at least 5%. The relative mean prediction error for WTP/Q is steady at around -15%, and the relative mean prediction error for $DWTP/Q$ is steady at around 14%. The relative mean prediction error for UPP does not stabilize and exhibits a consistent decline as the true price effect increases, crossing the horizontal axis (i.e., crossing from a positive to a negative mean prediction error) when the mean true price effect is about 12%. These results are broadly consistent with those in Figure 1.

Table 3 gives the mean prediction error, the standard deviation of the prediction errors, and relative mean prediction error for each of five categories of mergers, namely those for which the

Figure 2: Relative Mean Prediction Error by True Price Effects



true price effect is contained in the following increments: (0.5%,1.5%), (4.5%,5.5%), (9.5%,10.5%), (14.5%,15.5%), and (19.5%,20.5%).[25] Columns (1), (4), and (7) give the mean prediction error; columns (2), (5), and (8) give the standard deviation of the prediction errors; and columns (3), (6), and (9) give the relative mean prediction errors for WTP/Q, $DWTP/Q$, and UPP, respectively. Consistent with Figure 1, columns (1) and (4) of Table 3 indicate that the magnitude of the mean prediction error for WTP/Q and $DWTP/Q$ increases with the true price effect. Across our five categories of mergers, the mean prediction error for WTP/Q increases in magnitude from -0.002 in the (0.5%,1.5%) category to -0.033 in the (19.5%,20.5%) category. Similarly for $DWTP/Q$, the mean prediction error increases from 0.002 in the (0.5%,1.5%) category to 0.026 in the (19.5%,20.5%) category. Also consistent with Figure 1, column (7) of Table 3 indicates that UPP exhibits a different pattern. For UPP, the mean prediction error is 0.008 in the (0.5%,1.5%) category, rises to 0.017 in the (4.5%,5.5%) category, and then falls to -0.051 in the (19.5%,20.5%) category.

Consistent with Figure 2, column (3) of Table 3 indicates that the relative mean prediction error for WTP/Q is largely unchanged for categories of mergers such that the mean true price effects exceeds 5%, ranging in magnitude from -0.148 in the (9.5%,10.5%) category to -0.166 in the (19.5%,20.5%) category. But the relative mean prediction error for $DWTP/Q$ (column (6)) exhibits

[25]In what follows, it will prove convenient to break down mergers into categories fine enough that the true price effect of each merger is very close to the mean true effect of all of the mergers in its category. We chose to use categories that are one percentage point wide (e.g. 4.5% - 5.5%), and to present the five categories listed in the text. Presenting all of the categories would be cumbersome and would not yield additional insight. Appendix A4 contains the full set of results.

Table 3: Descriptive Statistics of Prediction Errors

Prediction Error Defined as a Percentage of Pre-Merger Price $\frac{\widehat{\Delta p_r} - \Delta p_r}{p_r}$

		Method 1: *WTP/Q*			Method 2: *DWTP/Q*			Method 3: *UPP*		
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mergers $s.t.\ \frac{\Delta p_r}{p_r} \in$	N	Mean	St Dev	Relative Mean	Mean	St Dev	Relative Mean	Mean	St Dev	Relative Mean
(0.5%,1.5%)	45,907	-0.002	0.005	-0.194	0.002	0.005	0.268	0.008	0.010	0.872
(4.5%,5.5%)	5,479	-0.008	0.016	-0.154	0.008	0.014	0.170	0.017	0.022	0.349
(9.5%,10.5%)	1,581	-0.015	0.025	-0.148	0.015	0.020	0.148	0.004	0.028	0.038
(14.5%,15.5%)	578	-0.022	0.041	-0.149	0.021	0.029	0.137	-0.020	0.033	-0.133
(19.5%,20.5%)	239	-0.033	0.043	-0.166	0.026	0.035	0.130	-0.051	0.036	-0.254

Relative Mean Prediction Error: Mean Prediction Error/Mean True Price Effect

a somewhat more meaningful improvement, declining from 0.170 in the (4.5%,5.5%) category to 0.130 in the (19.5%,20.5%) category. *UPP* (column (9)) appears to exhibit the greatest variation, declining from 0.349 in the (4.5%,5.5%) category to 0.038 in the (9.5%,10.5%) category. However, it continues to decline to -0.254 in the (19.5%,20.5%) category.

Finally, Table 3 indicates that the relative mean prediction errors of all three of the simulation methods in the (0.5%,1.5%) category are relatively high. However, given that the mean true price effects in this category of mergers is so low, the relative mean prediction errors in this category provide little useful information regarding the performance of the simulation methods.

Overall, these results indicate only modest bias for *WTP/Q* and *DWTP/Q*, particularly for mergers with price effects large enough that they are likely to pose a significant antitrust concern. The former exhibits some tendency to under-predict the true price effect, while the latter exhibits some tendency to over-predict it. *UPP* performs less well in that the magnitude of the bias is significantly greater for the categories of mergers in which the true price effect is lower than 8% or greater than 15%. However, as discussed in Section 7 below, in real-world cases *UPP* may have some practical advantages that the *WTP/Q* and *DWTP/Q* lack.

In Appendix A2, we provide an examination of the mechanisms underlying the biases exhibited by the simulation methods described here.

6.2 Dispersion of the Predicted Price Effects

Measures of bias alone are not sufficient to evaluate the performance of the simulation methods. Even if the prediction errors of a simulation method exhibit only a moderate amount of bias, the method can still be highly unreliable (i.e., may frequently be far away from the true price effect) if the prediction errors are large in magnitude but have opposing signs. For this reason, we now turn to two measures of performance that evaluate the dispersion of the predicted price effects of the simulation methods about the true price effects.

For the first measure, we calculate the frequency with which the predicted price effects are within a given proportion of the true price effects. Specifically, we calculate the following for each of the three simulation methods: (i) the frequency with which predicted price effect is less than 50% of the true price effect; (ii) the frequency with which predicted price effect is within 50% (in magnitude) of the true price effect; and (iii) the frequency with which predicted price effect is greater than 150% of the true price effect. The results are given in Table 4 for the five categories of mergers described above. See Appendix A4 for a full set of results.

Table 4 indicates that, at least for the categories of mergers such that the mean true price effects exceeds 5%, WTP/Q and $DWTP/Q$ perform quite well, and their performance improves as the true price effects increase. $DWTP/Q$ performs better than does WTP/Q, with predicted price effects that are within 50% (in magnitude) of the true price effects for 93.3% of mergers in the (4.5%,5.5%) category and 97.1% of mergers in the (19.5%,20.5%) category. WTP/Q performs somewhat less well, with predicted price effects that are within 50% (in magnitude) of the true price effects for 89.8% of mergers in the (4.5%,5.5%) category and 95.4% of mergers in the (19.5%,20.5%) category. UPP performs meaningfully less well, with predicted price effects that are within 50% (in magnitude) of the true price effects for 73.7% of mergers in the (4.5%,5.5%) category and 92.1% of mergers in the (19.5%,20.5%) category. Table 4 is consistent with the results in Table 3 in that WTP/Q is more likely to under-predict the true price effects by more than 50% than to over-predict by more than 50%, while the opposite is true for $DWTP/Q$. Also consistent with the results in Table 3 is that UPP is more likely to over-predict when the true price effects are relatively small and more likely to under-predict when the true price effects are relatively large.

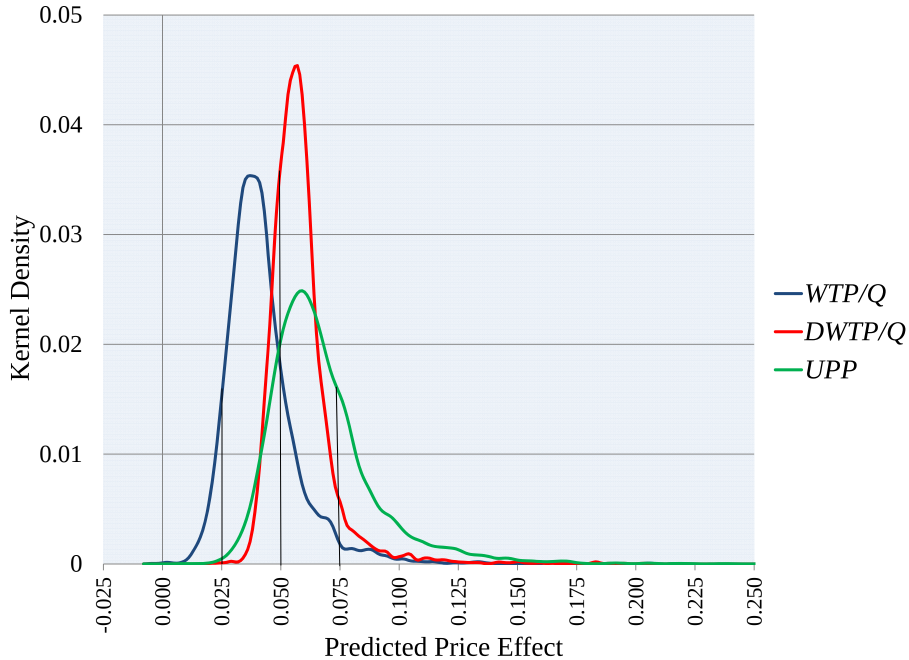
Table 4: Dispersion of Predicted Price Effects

Prediction Error Defined as a Percentage of Pre-Merger Price, $\frac{\widehat{\Delta p_r} - \Delta p_r}{p_r}$

Mergers $s.t.$ $\frac{\Delta p_r}{p_r} \in$	Method 1: WTP/Q			Method 2: $DWTP/Q$			Method 3: UPP		
	(1) $\frac{\widehat{\Delta p_r}}{p_r} \leq \frac{\Delta p_r}{2p_r}$	(2) $\lvert\frac{\widehat{\Delta p_r} - \Delta p_r}{p_r}\rvert < \frac{\Delta p_r}{2p_r}$	(3) $\frac{\widehat{\Delta p_r}}{p_r} \geq \frac{3\Delta p_r}{2p_r}$	(4) $\frac{\widehat{\Delta p_r}}{p_r} \leq \frac{\Delta p_r}{2p_r}$	(5) $\lvert\frac{\widehat{\Delta p_r} - \Delta p_r}{p_r}\rvert < \frac{\Delta p_r}{2p_r}$	(6) $\frac{\widehat{\Delta p_r}}{p_r} \geq \frac{3\Delta p_r}{2p_r}$	(7) $\frac{\widehat{\Delta p_r}}{p_r} \leq \frac{\Delta p_r}{2p_r}$	(8) $\lvert\frac{\widehat{\Delta p_r} - \Delta p_r}{p_r}\rvert < \frac{\Delta p_r}{2p_r}$	(9) $\frac{\widehat{\Delta p_r}}{p_r} \geq \frac{3\Delta p_r}{2p_r}$
(0.5%,1.5%)	0.180	0.760	0.059	0.001	0.844	0.155	0.016	0.398	0.586
(4.5%,5.5%)	0.065	0.898	0.037	0.001	0.933	0.066	0.001	0.737	0.262
(9.5%,10.5%)	0.046	0.937	0.017	0.000	0.956	0.044	0.009	0.936	0.054
(14.5%,15.5%)	0.043	0.941	0.016	0.002	0.962	0.036	0.024	0.964	0.012
(19.5%,20.5%)	0.033	0.954	0.013	0.000	0.971	0.029	0.079	0.921	0.000

Figure 3 depicts the kernel densities of the predicted price effects of the three simulation methods for mergers in the (4.5%,5.5%) category (i.e., when the true price effect is in that range). The figure illustrates: (i) the positive bias exhibited by $DWTP/Q$ and UPP and the negative bias exhibited by WTP/Q detailed in Table 3; and (ii) the relatively low (high) dispersion exhibited by $DWTP/Q$ (UPP) detailed in Table 4.

Figure 3: Kernel Densities of Predicted Price Effects for Mergers $r : \frac{\Delta p_r}{p_r} \in (4.5\%, 5.5\%)$



For our second measure of performance that evaluates the dispersion of the predicted price effects, we follow Miller et al. (2016) by calculating the Median Absolute Prediction Error (MAPE). As the name suggests, the MAPE is calculated by taking the absolute value of the prediction error for each simulated merger, and then taking the median of those absolute values. A lower MAPE corresponds to better performance.[26]

[26]We measure dispersion of the predicted price effects about the true price effects using the MAPE ratio, which has two potentially important limitations. One is that the MAPE ratio utilizes the absolute loss function (i.e., the absolute difference between the true and predicted price effects), and our results may differ under alternative loss functions such as quadratic. The other is that our measurement of the MAPE utilizes just the point estimates of the predicted price effects of the simulation methods. It does not account for the sampling distributions of the predicted price effects generated in the least squares regression models underlying the merger simulation methods, (14) and (16). We address both of these points by constructing an alternative performance metric based on Root Mean Squared Error, which, like the MAPE ratio, accounts for dispersion of the predicted price effects about the true price effects, but also penalizes the simulation methods for providing larger standard errors with its predictions. The results were similar to our baseline results, so we do not report them here. They are available from the authors upon request.

We evaluate the MAPE within each of the 31 merger categories defined above, and express the MAPE as a percentage of the mid-point true price effect (e.g., a true price effect of 5% in the 4.5%-5.5% category). We refer to this metric as the MAPE ratio. Hence, defining the prediction error as a percentage of the pre-merger price p_r, we evaluate the MAPE ratio for each of *WTP/Q*, *DWTP/Q*, and *UPP* as

$$\frac{\text{med}\left\{\left|\frac{\widehat{\Delta p_r}-\Delta p_r}{p_r}\right|\right\}_{r:\left|\frac{\Delta p_r}{p_r}-x\right|<0.005}}{x}, \text{ for } x \in \{0.01, 0.02, ..., 0.30\}. \tag{22}$$

For example, if a simulation method had a MAPE ratio of 0.2 for mergers in the 4.5%-5.5% category, that would mean that half of the predicted price effects generated by that method would be within one percentage point of the true effect, and half would be outside that range.

The results are given in Table 5. For *WTP/Q*, the MAPE ratio decreases from 0.290 in the (0.5%,1.5%) category to 0.194 in the (19.5%,20.5%) category. The MAPE ratio for *DWTP/Q* is relatively constant, decreasing from 0.141 in the (0.5%,1.5%) category to 0.135 in the (19.5%,20.5%) category. Consistent with the bias patterns for *UPP* described above, the MAPE ratio for *UPP* decreases from 0.534 in the (0.5%,1.5%) category to 0.165 in the (9.5%,10.5%) category but then increases to 0.246 in the (19.5%,20.5%) category.

Table 5: MAPE Ratios

Mergers *s.t.* $\frac{\Delta p_r}{p_r} \in$	Method 1: *WTP/Q*	Method 2: *DWTP/Q*	Method 3: *UPP*
(0.5%,1.5%)	0.290	0.141	0.534
(4.5%,5.5%)	0.246	0.144	0.278
(9.5%,10.5%)	0.209	0.138	0.165
(14.5%,15.5%)	0.212	0.127	0.197
(19.5%,20.5%)	0.194	0.135	0.246

We are not aware of any objective benchmark by which to evaluate whether these MAPE ratios indicate "good" or "poor" performance in predicting the true price effects. Our primary approach is to present the results in full detail, and leave it to the reader to form his or her own opinion. However, our own standard, which we will apply in our characterization of our results, is as follows. A predictor with a MAPE ratio of less than 0.15 (e.g., half of predictions would be within 0.75 percentage points for mergers with a true price effect of 5%) is a highly reliable predictor of the true price effects. A predictor with a MAPE ratio in the (0.15,0.25) range is less reliable but still likely to be highly informative of the true price effects. A predictor with a MAPE ratio greater than 0.25 (e.g., half of predictions would be within 1.25 percentage points for mergers with a true price effect of 5%) is significantly less informative of the true price effects, but nevertheless may be worthy of consideration in analyzing a merger. A predictor with a MAPE ratio above 0.4 is likely to be of limited usefulness in predicting the price effects of mergers. Of course, these thresholds are arbitrary. For example, we do not view MAPE ratios of 0.148 and 0.152 as meaningfully different.

Two important conclusions can be drawn from the MAPE ratios. First, while the relative mean prediction error of WTP/Q changes very little for categories of mergers such that the mean true price effect exceeds 5%, the MAPE ratio of WTP/Q declines significantly as the mean true price effect increases. For example, the MAPE ratio results indicate that WTP/Q is a much more reliable predictor of the true price effects in the (19.5%,20.5%) category than in the (4.5%,5.5%) category (0.194 v. 0.246) even though the relative mean prediction error of WTP/Q is about the same (-0.166 v. -0.154).

Second, $DWTP/Q$ has a significantly lower MAPE ratio than WTP/Q in each of the five categories of mergers. This indicates that $DWTP/Q$ is the more reliable predictor of the true price effects even though the magnitude of its bias about the same as that of WTP/Q. This is consistent with the fact that, as illustrated in Table 3, the prediction errors for WTP/Q exhibit significantly greater variance compared to $DWTP/Q$. For example, in the (4.5%,5.5%) and (19.5%,20.5%) categories, the standard deviation of the prediction errors of WTP/Q is larger than the standard deviation of the prediction errors of $DWTP/Q$ (0.016 v. 0.014 and 0.043 v. 0.035, respectively).

To summarize our main findings, WTP/Q and $DWTP/Q$ exhibit a moderate amount of bias that is persistent in sign across all mergers. WTP/Q exhibits a tendency to under-predict the true merger price effects, while $DWTP/Q$ exhibits a tendency to over-predict the true merger price effects. *UPP* exhibits a tendency to over-predict the true price effects when the true price effects are low but an increasing tendency to under-predict the true price effects when the true price effects are high.

We also find that $DWTP/Q$ performs well in predicting the price effects of mergers in our simulations for all categories of mergers. The MAPE ratio for $DWTP/Q$ is consistently below 0.15, which we view as very good. WTP/Q performs reasonably well in predicting the true price effects in our simulations for mergers in the categories with the highest true price effects ((9.5%,10.5%) and greater). The MAPE ratio for WTP/Q in these categories of mergers is consistently around 0.20, which we view as reasonably good. However, WTP/Q performs significantly less well in the (0.5%,1.5%) and (4.5%,5.5%) categories of mergers, in which the MAPE ratios are 0.290 and 0.246, respectively. *UPP* performs reasonably well in predicting the true price effects of mergers in our simulations for mergers in the (9.5%,10.5%) and (14.5%,15.5%) categories, with MAPE ratios of 0.165 and 0.197, respectively. However, *UPP* performs significantly less well in the (4.5%,5.5%) and (19.5%,20.5%) categories of mergers, in which the MAPE ratios are 0.278 and 0.246, respectively.

In Appendix A5, we present present a series of robustness tests. We examine performance of the simulation methods: (i) under different competitive conditions in the hospital and insurance markets (A5.1); (ii) conditional on each possible parameter value in our theoretical model (A5.2); and (iii) under seventeen modifications to our baseline parameterizations and assumptions (A5.3).

6.3 Application as Screen in Prospective Merger Analysis

To this point, our results have been about how closely the predictions of the merger simulation methods correspond to the true price effects from the theoretical model. We now address a related question that may be of particular interest to antitrust practitioners, namely how effectively a screen that is based on the simulation methods (i.e., challenge a merger if the predicted price effect is greater than some threshold) flags mergers with true effects above the threshold and avoids flagging mergers with true effects below the threshold. Following Miller et al. (2016), we adopt a threshold of 5% in this analysis.[27,28]

We proceed by using the same 31 merger categories as before (i.e., one percentage point increments of the true price effect $\frac{\Delta p_r}{p_r}$, such as (≤0.5%), (0.5%,1.5%), (1.5%,2.5%), etc. Within each of these merger categories, we calculate the frequency with which the predicted price effect exceeds 5%.

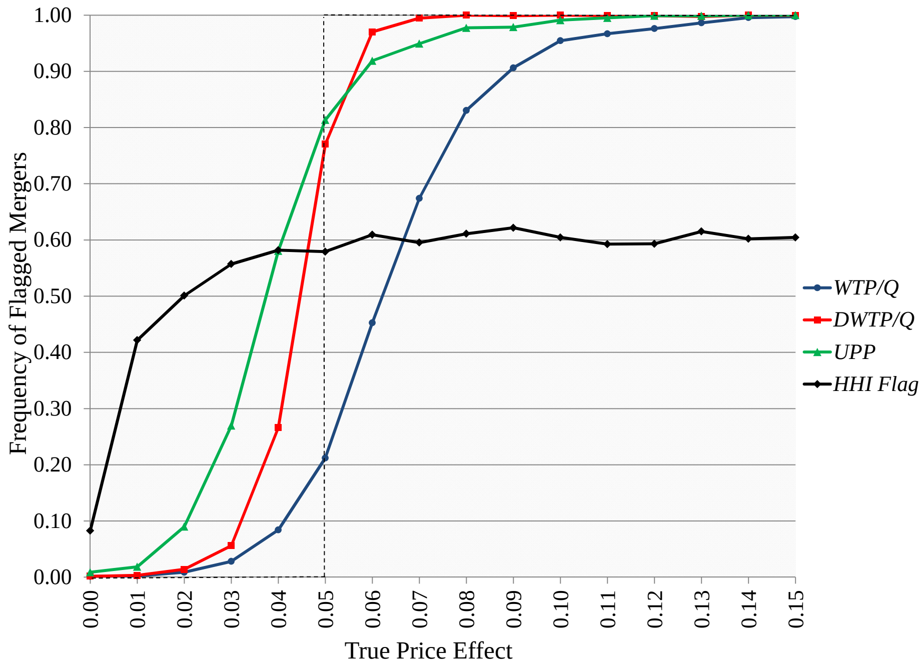
The results are given in Figure 4. A hypothetical perfect predictor is represented by the dashed line. Such a predictor would flag 100% of mergers for which the true effect is greater than 5%, and 0% of mergers for which the true price effect is less than 5%. For any imperfect predictor, when the true price effect is at least 5%, the absolute difference between this frequency and unity gives the rate of false negatives. Similarly, when the true price effect is less than 5%, the difference between this frequency and zero gives the rate of false positives. For example, among the mergers with true effects in the (6.5%,7.5%) category, WTP/Q predicts a price increase of at least 5% in 67.4% of mergers, giving a false negative error rate in that category of 32.6%. In contrast, $DWTP/Q$ and UPP predict a price increase of at least 5% in 97.0% and 92.9% of mergers, respectively. This give much lower false negative rates in the category of mergers for $DWTP$ (3.0%) and UPP (7.1%). As the true price effects become larger, the rates of false negatives go to zero for each of the simulation methods. That is, the rate of very large false negatives (e.g., failing to flag a merger using a 5% screen when the true price effect is 10% or greater) is small for all three methods.

A similar comparison indicates that DWP/Q and UPP have higher false positive rates than does WTP/Q. For example, in the (3.5%,4.5% categories of true price effects), WTP, $DWTP$, and

[27] Note that this choice of threshold does not mean that mergers that cause price increases of less than 5% are permissible. There are a number of reasons why a relatively high threshold might be chosen that are beyond the scope of this paper. We have also performed a similar analysis using a 2% threshold. The results are broadly similar.

[28] Of course, even if such a screen were to be used in the real world, it would be only one element of the full array of theory and evidence, both quantitative and qualitative, on which decisions on whether to challenge a merger are based.

Figure 4: Error Rates for 5% Price Effect Threshold



UPP predict price increases of at least 5% in 8.4%, 26.6%, and 58.1% of mergers, respectively. As the true price effects become smaller, the rates of false positives go to zero for each of the simulation methods. The rate of very large false positives is small for all three methods.

These results are broadly consistent with our earlier results. For mergers in the (4.5%,5.5%) *WTP/Q* tends to under-predict the true effects, and therefore has a relatively high rate of false negatives and a low rate of false positives. The reverse is true for *DWTP* and *UPP*. See Figure 1.

The 2010 DOJ/FTC *Horizontal Merger Guidelines* lays out a screen that is based on market concentration. Specifically, it a market is classifies as "highly concentrated" market if the Hirfindal-Hirschman Index (HHI) is at least 2,500.[29] A merger is presumed to likely enhance market power if the post-merger HHI exceeds 2,500 and the change in the HHI is at least 200. We apply this screen in Figure 4 as well.

In constructing the HHI in this analysis, we construct hospital system shares using the expected volume of each system in each simulated market. That is, we assume that each of the twelve hospital systems in each market is included in the relevant antitrust market.[30]

[29]In our simulations, the mean market-level HHI is 2,996. This is somewhat lower than the mean MSA-level HHI of 3,261 in the United States for 2006 reported in Gaynor et al. (2015).

[30]Each of our 9,000 markets is assumed to be a "market" for the purposes of calculating the HHIs, which means that each market contains twelve hospitals. Consistent with previous work by Miller et al. (2016) and Garmon (2017), we do not perform a market definition exercise using the Hypothetical Monopolist Test as described in the DOJ/FTC Horizontal Merger Guidelines (https://www.justice.gov/atr/horizontal-merger-guidelines-08192010). Had we done

We find that the HHI flag performs very poorly relative to the merger simulation methods. Using a 5% threshold, the HHI flag generally has higher rates of both false positives and false negatives. For all mergers in the (3.5%,4.5%) category and above, the HHI flags mergers as likely enhancing market power with a frequency of about 60%. Hence, the HHI flag has a false negative rate of about 40% irrespective of the true price effect. It also has much greater false positive rates: about 58.2% in the (3.5%,4.5%) category and 42.5% in the (0.5%,1.5%) category.

Table 6 summarizes these results in a manner similar to that in Table 4 of Garmon (2017). In columns 1 through 5, Table 6 contains the number of flagged mergers (using a 5% screen), correct positives, correct negatives, false positives, and false negatives for each simulation method and for the HHI flag. Column 6 contains the mean true price effect for the flagged mergers. Columns 7, 8, and 9 give the results of performance metrics that are commonly applied in machine learning algorithms. *Markedness* (column 7) measures how frequently the predictions (positive and negative) are correct.[31] *Informedness* (column 8) measures how frequently the true outcomes (positive and negative) are correctly predicted by the prediction method.[32] *Markedness* and *Informedness* are scaled from -1 to 1, with 1 indicating perfectly correct predictions, -1 indicating perfectly incorrect predictions, and 0 indicating that the predictions are random. The Matthews Correlation Coefficient (column 9) is the geometric mean of *Markedness* and *Informedness*. Consistent with the earlier results, $DWTP/Q$ has the highest Matthews Correlation Coefficient while the HHI flag has by far the lowest.

7 Discussion

We now address the question of what inferences can be validly drawn from our results. The question of interest is whether the simulation methods predict real-world price effects well. More specifically, it is whether they predict real-world price effects well enough to merit receiving substantial weight in real-world merger analysis. There are two possible reasons why they might not. First, the methods might not accurately predict the price effects from the theoretical model. Second, the methods might accurately predict the theoretical model, but the model might not closely correspond to the real world. Our experiment can be thought of as a test of the first reason. In Bayesian terms, a negative result from that test (i.e., a finding that the simulation methods are poor predictors of

so, the HHI-based simulation might have performed better in flagging problematic mergers. On the other hand, performing this type of market definition would require using one of the simulation methods to determine whether a hypothetical monopolist could profitably increase price. Moreover, market definition has the well-known problem that it treats every hospital as either completely in the market or completely outside, rather than allowing hospitals to vary in their degree of competitive significance. In contrast, none of the three simulation methods evaluated in this paper require any market definition, which is an important advantage.

[31] Specifically, *Markedness* is defined as the ratio of correct positive predictions to all positive predictions plus the ratio of correct negative predictions to all negative predictions minus 1.

[32] Specifically, *Informedness* is defined as the ratio of correct positive predictions to all true positive outcomes plus the ratio of correct negative predictions to all true negative outcomes minus 1.

Table 6: Correct and False Predictions Based on a 5% Price Effect Threshold

Method	Flagged Mergers	Correct Positive	Correct Negative	False Positive	False Negative	Mean True Price Effect for Flagged Mergers	Mark-edness	Inform-edness	Matthews Corr Coeff
WTP/Q	19,248	17,326	206,801	1,922	5,876	0.110	0.873	0.738	0.802
DWTP/Q	27,943	22,702	203,482	5,241	500	0.093	0.810	0.953	0.879
UPP	35,530	22,233	195,426	13,297	969	0.078	0.621	0.895	0.745
HHI Flag	53,895	13,882	168,710	40,013	9,320	0.038	0.205	0.407	0.289

There are 231,925 mergers in the sample, 23,202 of which result in a true price effect of at least 5%.

the true effects from the theoretical model) would lead to a very low posterior probability that the simulation methods predict real-world price effects well, regardless of the prior probability. However, if the test is passed, that may justify a meaningful positive updating of the probability that the simulation methods do predict real-world price effects well. See Appendix A3 for a discussion of the factors that influence the magnitude of the Bayesian update.

Our approach has a number of important limitations, both conceptual and practical. The most obvious conceptual limitation is that our experiment is not based on real-world data. So even if our theoretical model is a good representation of the real world, we cannot be certain that it is calibrated correctly, though we can partially address this by using some sources of real-world data to guide our parameterizations. Another conceptual limitation is that while that our theoretical model appears to capture important features of reality, that is far from constituting a proof that it close enough to reality to generate reliable results. The model does not incorporate some other factors in real-world bargaining between hospitals and insurers that may be important. For example, the model assumes simultaneous bargaining between hospital and insurers and symmetric competition in the insurance market, neither of which is certain to obtain in the real world. In addition, our model is set up so that all model hospital-insurer combinations will reach an agreement in equilibrium. It does not account for the possibility of equilibrium network exclusions. It also does not allow for tiering or other steering arrangements, or "most-favored nation" clauses, or co-insurance (as opposed to co-pays), which have the effect of making patients pay different out-of-pocket prices for different hospitals in their insurer's network. We leave an examination of the how well the merger simulation methods perform in the presence of such exclusions or tiering for future research.

Our theoretical model also assumes that consumers can experience only one type of health condition that requires inpatient treatment. In the real world, of course, there are many types of health conditions that result in inpatient events. This is important because consumers' rate of exchange between their valuation of an insurer's network and premium, governed by the parameters θ and λ in our theoretical model, may vary considerably across health conditions. Since our theoretical model allows only one type of health condition, it cannot capture such variation. Both Capps et al. (2003) and Gowrisankaran et al. (2015) also assume away variation in the rate of exchange between valuation of an insurer's network and prices paid to hospitals across health conditions in their empirical models. If such variation is important in the real world, it would likely be a significant source of prediction error and one that our analysis does not address. We view this as a potentially important area for future research.

Finally, we assume that the division of the joint surplus in bargaining between hospital systems and insurers, governed by the parameter α in our theoretical model, is the same for each hospital system-insurer combination. Meaningful variation in the value of this parameter across hospital system-insurer combinations would likely be another source of prediction error that our analysis does not address.

In addition to these conceptual issues, our approach makes a practical assumption about what is available to the hypothetical analyst in our experiment that is unlikely to obtain in the real world. Specifically, we assume that the analyst has the correct model of consumer preferences over hospitals in estimating the discrete choice model (including the correct distribution of the idiosyncratic component) and either has all of the relevant data (distance in our theoretical model) or can make reliable inferences on what is unobserved (hospital quality in our theoretical model). In practice, errors in modeling consumer preferences or data limitations will affect the estimation of the discrete choice model underlying diversion ratios and WTP.

In addition to the limitations of our approach, there are other practical limitations to applying the WTP/Q and $DWTP/Q$ methods. As described in Brand and Garmon (2014) and Farrell et al. (2011), in a given hospital market, there may be a small number of observations or insufficient variation in the data (i.e., the hospital systems in the analysis may have similar values of WTP/Q or $DWTP/Q$). In this case, the relationship between price and WTP/Q or $DWTP/Q$ cannot be reliably estimated. Under these circumstances, UPP may be the more reliable method. The severity of these problems, and hence the appropriateness of applying WTP/Q or $DWTP/Q$, or the weight that the results should be given if the simulation methods are applied, is likely to depend on case-specific circumstances.

In sum, we find evidence that the simulation methods do a good job of predicting the true price effects of our theoretical model. This result, combined with some reason to believe that the model is a reasonable approximation of the real world, is sufficient to justify a positive updating of the prior probability that the simulation methods predict real-world price effects well enough for them to receive substantial weight in real-world merger analysis.

Given this generally positive result, it remains to discuss the relative merits of the three simulation methods that we analyze: *WTP/Q*, *DWTP/Q*, and *UPP*. Based on our results, *DWTP/Q* outperforms *WTP/Q* in our simulations. This is not surprising given the fact that the *DWTP/Q* extension is more closely connected to our theoretical model. However, *DWTP/Q* is more sensitive to variation in some of the key parameters. This suggests that the more "reduced-form" approach of *WTP/Q* has at least some merit.

Both *WTP/Q* and *DWTP/Q* substantially outperform *UPP* in our simulations. However, *UPP* has some important practical advantages. It is much easier to calculate and apply, and it is free from at least some of the practical problems associated with *WTP/Q* and *DWTP/Q*. For example, *UPP* does not require price and cost data for third party hospital systems, as do *WTP/Q* and *DWTP/Q*. The more severe these practical problems prove to be in a particular case, the stronger the justification for using *UPP*, and vice-versa. In addition, as discussed in Appendix A5.3, our results suggest that *UPP* may be less sensitive to measurement error in prices compared to *WTP/Q* and *DWTP/Q*. For this reason, there may be good justification to use *UPP* in merger analysis.

We close by contrasting our approach to evaluating the accuracy of these simulation methods to an alternative approach using event studies. Under this alternative approach, the price effect of mergers is estimated by performing retrospective analyses of a number of hospital mergers, applying the merger simulation methods to the pre-merger data from those mergers, and comparing the predictions of the simulation methods to the estimates from the retrospective analyses. This is the general approach taken by Fournier and Gai (2007), May and Noether (2014), and Garmon (2017) in the hospital industry, and by Peters (2006), Ashenfelter and Hosken (2010), Weinberg (2011), and Weinberg and Hosken (2013) in other industries. While clearly valuable, this approach comes with some difficulties. Perhaps the biggest difficulty is the limited power of the test; each retrospective analysis and each merger simulation analysis is a formidable undertaking, and it is costly to perform enough of them to generate sufficient powerful. This problem is compounded by the fact that the retrospective analyses may measure the true price effects with considerable error. This is partly because of the difficulty in defining valid control groups for the difference-in-differences analyses, and partly because the researcher generally does not have information on the timing of contract renewals. This latter point is important; the effect of a merger on hospital-insurer bargaining is only registered at the next contract negotiation. Until then, there may be no price effect, or there may be an effect that arises if the acquiring hospital has a higher price than the acquired one, and the acquiring hospital is allowed to fold the acquired hospital into its existing contracts until the next negotiation. This fact, combined with the fact that the merged hospitals' negotiations may take place at different dates from those of the control hospitals, can lead to spurious findings of a price effect, or to spurious findings of a lack of an effect. More generally, mergers may cause changes in equilibrium prices for reasons other than the loss of horizontal competition. Retrospective approaches typically cannot disentangle price changes due to the elimination of competition or merger-specific efficiencies from other changes that may be caused by a merger. Our approach does not suffer from these difficulties.

8 Conclusion

In recent years, researchers have developed new methods for predicting the price effects of hospital mergers. A natural question to ask is how well these methods work. The purpose of this paper is to make a contribution to answering this question. We do this by laying out a rich model of hospital and insurer competition, and then performing numerous Monte Carlo simulations using this model. We run these simulations under a variety of assumed ownership configurations, which generates "true" price effects for a large number of simulated mergers. We then compare these true price effects to the effects predicted by the simulation methods. While the performance varies somewhat, both across the simulation methods that we evaluate and across different parameterizations of the model, for the most part the simulation methods perform quite well.

Another question that could be addressed using an approach broadly similar to that used in this paper is the accuracy of very simple metrics, such as diversion ratios or percent changes in WTP (which are considered in Garmon (2017)) as screening tools for identifying anti-competitive mergers. This is important because those simple measures are very easy to compute and do not require data on prices, which are often difficult to obtain or may be measured with significant error. In addition, there are some instances in which WTP/Q and $DWTP/Q$ cannot be reliably applied but the simpler methods can be. For example, the regression models used in these two simulation methods may give unreliable results when there are only a small number of hospitals that face conditions similar to those faced by the merging hospitals, or where there is not enough heterogeneity across hospitals to obtain precise estimates in a merger simulation regression model. These simple measures may be useful under such circumstances since they do not require estimating a regression model. The UPP simulation method discussed in this paper has this advantage as well.

An additional benefit of this research is that it may provide guidance for future research in merger analysis in healthcare markets. Our results indicate that the merger simulation methods are sensitive to variation in some of the key parameters of our theoretical model. Most notably, $DWTP/Q$ performs less well for higher values of λ, which measures sensitivity to variation in consumers' valuations of insurers' hospital networks in the insurance choice problem. λ is a key parameter in determining the market power of hospitals. The merger simulation methods developed to date are not able to identify this important parameter absent a model of insurance demand. Hence, an empirical approach that is able to do so may significantly improve the predictive power of hospital merger simulations.

Appendices

A1 Parameterization

In this appendix, we provide a complete discussion of our parameterizations of the theoretical model. As discussed in Section 4, most of the model parameters for each simulation take on one of three possible values, which are randomly assigned with equal probability. We determine the set of possible values by benchmarking the pseudo-R^2 values from the conditional logit model (used to construct WTP and diversion ratios to real-world values) as well as hospital prices, costs, and gross margins, against real-world values.[33]

The parameters that determine the pseudo-R^2 values from the conditional logit model can be benchmarked without reference to hospital gross margins. These include the parameters governing the distributions of consumer and hospital locations and the variance of hospital quality, as well as the parameters governing the preferences of consumers over hospitals as defined in (1). Hence, we first determine the sets of values for these parameters and then determine the sets of values for the remaining parameters by benchmarking against hospital prices, costs, and gross margins.

A1.1 Hospital and Consumer Attributes

Each hospital j is characterized by a location draw $(x_j, y_j) \sim F^j_{xy}$, a quality draw $\eta_j \sim F_\eta$, a constant marginal cost c_j (that is common to all hospitals), and a system affiliation. Each patient i is characterized by a location draw $(x_i, y_i) \sim F^i_{xy}$ and a draw defining the probability of needing inpatient care $\rho_i \sim F_\rho$.

For each simulation, every hospital and every consumer has a randomly assigned location. These locations are characterized by their position relative to the origin. The variance of F^j_{xy} (dispersion of hospital locations) is set to be somewhat less than that of F^i_{xy} (dispersion of consumer locations). This is in order to make it unlikely that a hospital will be located at the edge of the population of consumers.

Each simulation is randomly assigned one of two distributions for F^j_{xy} and F^i_{xy}: Normal, to replicate a densely populated city center with thinly populated surrounding areas; and Uniform, to replicate a large suburban area where the population is evenly distributed. We use the following Normal and Uniform distributions for the locations of consumers and hospitals:

$$(F^i_{xy}, F^j_{xy}) \in \left\{ \left(N(0,9)^2, N(0,8)^2\right), \left(U[-16,16]^2, U[-14,14]^2\right) \right\}. \tag{A1}$$

For a draw of hospital locations in a given simulated market, we center the hospital locations at the origin.

[33] We also evaluate the gross margins and market shares of insurers, as well as pass-through rates of changes in hospital prices through insurance premiums in determining the set of possible model parameters.

We assume a normal distribution for F_η. To benchmark the standard deviation of F_η, we examined the distribution of hospital fixed-effects estimated in previous analyses using real-world patient-level discharge data. Hospital fixed-effects are often used to control for unobserved attributes such as quality, so variation in real-world fixed effects estimates provides a reasonable proxy for the variation in hospital quality. In examining the output of several previous analyses, we found that the standard deviation of the estimated hospital fixed-effects typically lies in the interval [1.4, 1.8].[34] Therefore, for each simulation we draw a value of the standard deviation of F_η from the set {1.4, 1.6, 1.8}. For a draw of $\{\eta_j\}_{j \in J}$ in a given simulated market, we do not rescale the draws to ensure that the sample standard deviation equals the population analog. Hence, given the small number of hospitals in our model, the variation in quality across hospitals varies significantly across our simulated markets. We discuss the mean of F_η below.

We assume that hospital marginal cost c_j is perfectly correlated with hospital quality η_j. Hence, quality variation is the only source of cost variation in our simulations. Specifically, we assume

$$c_j = c + 0.2(\eta_j - E[\eta_j]),$$

where c denotes the expected hospital marginal cost. In our simulations, this specification generates somewhat less within-market variation in hospital marginal cost as there is within-market variation in WTP/Q and somewhat more within-market variation in hospital marginal cost as there is within-market variation in $DWTP/Q$.[35]

Quality, which is perfectly correlated with cost, is also positively correlated with both WTP and hospital volume Q. Quality is also positively correlated WTP/Q because Q in linear in the probability that a given consumer will choose that hospital, but WTP is convex in the probability that a given consumer will choose that hospital. This correlation can introduce collinearity such that the effects of WTP/Q on price are confounded with the effects of cost. This collinearity tends to degrade the performance of the two WTP-based simulation methods, but as discussed in Section 6 the methods generally perform well despite this. In the real world, the correlation between cost and quality is less than unity, so the collinearity problem is likely to be smaller. That is, the assumption of perfect correlation between hospital cost and quality is conservative in that it tends to decrease the performance of the simulation methods in our Monte Carlo experiment.

This collinearity problem can result in a negative estimated relationship between price and WTP/Q (and between price and $DWTP/Q$). But the estimated value of β_1 is negative in only six of our 9,000 simulated hospital markets, and in only three of those six markets (and in no others) is the estimated value of Γ_1 also negative. However, even in these six markets, the raw

[34]For example, the standard deviation of the hospital fixed-effects reported in Gowrisankaran et al. (2015) is 1.75.

[35]The median (across simulated markets) standard deviations of hospital marginal cost, WTP/Q, and $DWTP/Q$ are 0.285, 0.400, and 0.179, respectively. We have explored different marginal cost scalings such as $c_j = c + 0.5(\eta_j - E[\eta_j])$. The results are very similar to our baseline results.

correlation between price and WTP/Q (and between price and $DWTP/Q$) is always positive, so the negative coefficient is likely the result of collinearity. That is, a negative estimated relationship between price and WTP/Q is extremely rare in our simulations even given an assumption (perfect correlation between cost and quality) that would tend to make it more likely.

A1.2 Consumer Preferences over Hospitals

We specify the utility of consumer i for hospital j in (1) as

$$U_{ij} = -\gamma_1 dist_{ij} - \gamma_2 dist_{ij}^2 + \eta_j + \epsilon_{ij}, \tag{A2}$$

where $dist_{ij}$ denotes the straight-line distance from consumer i to hospital j, γ_1 and γ_2 measure the effect of distance on utility, and ϵ_{ij} is an IID Type I Extreme Value draw.[36]

Given the variation in η_j, ϵ_{ij}, and the location distributions, we select parameter values for the utility cost of travel, (γ_1, γ_2), so that the resulting pseudo-R^2 values from our discrete choice model estimation are similar to those found in practice, which are usually in the range of (0.40, 0.55). For each simulated market, we randomly assign values of (γ_1, γ_2) from the set {(0.1,0.001),(0.3,0.003), (0.5,0.005)}.

Table A1 gives percentiles of the distribution of the pseudo-R^2 values across our simulated markets. The range 0.40-0.55, which is most consistent with real-world experience, is roughly covered by the 25th and 50th percentiles. For reasons discussed in Section 6, we include parameterizations that generate pseudo-R^2 values that go well beyond this range. We select parameterizations so as to produce more pseudo-R^2 that are greater than 0.55 than ones that are less than 0.40. This is conservative in the sense that the simulation methods tend to perform less well in simulated markets with higher pseudo-R^2 values; the pseudo-R^2 values greater than 0.55 generally occur when travel costs are high, $(\gamma_1, \gamma_2) = (0.5, 0.005)$, and, as discussed below, our results show that the simulation methods generally perform less well when travel costs are high.

Table A1: Percentiles of Pseudo-R^2 Values

10^{th}	25^{th}	50^{th}	75^{th}	90^{th}
0.296	0.422	0.558	0.652	0.698

A1.3 Bargaining Game

The bargaining parameter α defines that share of the joint surplus that is captured by hospitals. Hence, it is a key parameter in determining hospital gross margins and the price effects of mergers. We assume that hospitals and insurers either split the joint surplus 50-50 or that there is a modest

[36] In practice, driving distances or average drive-times would be used instead of straight-line distances.

deviation from an even split in either direction. Specifically, for each simulated market, we randomly assign the value of α from the set $\{0.4, 0.5, 0.6\}$.

A1.4 Insurance Market Parameters

There are several parameters that govern preferences over insurers. These are defined in (2), and include λ, θ, Z, and the parameters of F_η. Given the set of values for the parameters governing the consumer and hospital attributes, consumer preferences over hospitals, and the split of the joint surplus in the bargaining game, and for the reasons discussed in Section 6, we choose these parameters so that equilibrium hospital gross margins cover a wide distribution centered at 0.50.

The parameter λ, plays a particularly important role in the model. It scales the consumer's expected utility of the insurer's hospital network (i.e., it governs how much consumers care about the exclusion of a hospital from an insurer's network, and hence how likely they are to switch to a competing insurer if a particular hospital is excluded from their insurer), and so it plays a key role in determining how much market power hospitals have. Higher values of λ imply lower disagreement payoffs of insurers but, importantly, do not affect the disagreement payoffs of hospitals. Since higher λ means less insurer bargaining leverage, it causes higher hospital margins and larger price effects.

One objective in choosing values of λ is to generate meaningful variation in the curvature of the demand faced by insurer with respect to consumers' expected utility of its hospital network,

$$\frac{1}{\#I_g} \sum_{i \in I_g} \rho_i \ln \sum_{j \in J^m} \exp\{V_{ij}\}.$$

As defined in (3), the probability that a consumer will choose to buy insurance from a given insurer is a non-linear function of this expected utility.

It is important to choose parameter values such that this function exhibits meaningful departures from linearity. The reason is that, as can be observed from (12), (14), and (16), the merger simulation methods assume that hospital prices are linear in the differences, under hypothetical exclusions, in consumers' expected utility of the insurer's hospital network (in the case of WTP/Q), or linear in a linear combination of these differences (in the case of $DWTP/Q$). This represents a meaningful difference between the theoretical model and the simulation methods, and it is important to test the performance of those methods when that difference is substantial in magnitude. For each simulated market, we randomly assign a value of λ from the set $\{2, 5, 8\}$.

In our theoretical model, the probability that a given consumer will purchase insurance from a given insurer will exhibit greater curvature in the consumer's expected utility of the insurer's hospital network larger values of λ. Hence, *a priori*, it seems likely that the merger simulation methods will perform less well under parameterizations with larger values of λ. But as seen in Appendix A5.2, the methods perform quite well even under relatively high values of λ.

Like λ, the parameter θ, which measures the sensitivity of consumers to insurance premiums, plays a key role in determining how much market power hospitals have. Under lower values of θ, consumers are less sensitive to changes in insurance premiums, and, therefore, are less likely to switch to the outside option (no insurance) under a premium increase. Lower values of θ also imply lower disagreement payoffs for insurers because it is more difficult for insurers to compensate consumers for a hypothetical network exclusion by offering a lower premium. Therefore, hospital gross margins and merger price effects due to mergers are generally decreasing in θ. For each simulated market, we randomly assign the value of θ from the set $\{0.5, 0.8, 1.1\}$.

We set the value of the mean of the hospital quality distribution, F_η, so that the value of $\ln \sum_{j \in J^n} \exp\{V_{ij}\}$ (i.e., the value of the insurance network) is positive for almost all consumers. This is to ensure that consumers with a higher value of ρ_i (i.e., sicker consumers) are more likely to buy health insurance than are those with a lower value, ceteris paribus. Given the aforementioned parameter values of the location distributions and travel costs, we draw values of mean hospital quality from the set $\{14, 16, 18\}$.

The parameter Z measures consumers' valuation of non-inpatient healthcare services covered by the insurer. Ideally, values of Z should be selected to reflect how consumer's weigh the relative values of expected inpatient and non-inpatient healthcare services in their health insurance purchasing decisions. Since we do not have empirical evidence on which to base this evaluation, we choose values of Z so that in some simulated markets, consumers value expected inpatient and non-inpatient healthcare roughly equally, on average, and in other simulated markets, consumers systematically place greater weight on one or the other. By happenstance, we find that the distribution of consumers' expected utility of the insurer's hospital network is usually centered around one. For each simulated market, we randomly assign the value of Z from the same set as λ, $\{2, 5, 8\}$.

Our theoretical model contains two additional insurer cost parameters: p_z, which denotes healthcare expenditures for non-inpatient services, and τ, which denotes a per inpatient event administrative cost. We set the value of p_z buy again referring to Health Care Cost Institute (2015), which notes that 2014 per capita non-inpatient expenditures in the commercial sector were \$3,969, with per capita out-of-pockets expenditures of \$759. Given this information, we set the value of p_z to \$3,200. To set the value of τ, we select values that, based on average hospital prices, represent a small, but not trivial, added cost for the insurer to administer inpatient claims. For each simulated market, we randomly assign the value of τ from the set $\{\$500, \$750, \$1000\}$.

Finally, we randomly assign the number of insurers in each simulated market. As discussed above, variation in the number of competing insurers has a significant effect on the disagreement payoffs of both insurers and hospitals, and, therefore, may have a significant effect on pre-merger margins and on the price effects of mergers. This represents an important difference between the theoretical model and the simulation methods, as the methods ignore the effect of insurer competition in determining the equilibrium of the bargaining game. Specifically, the simulation methods assume that hospitals cannot recapture patients through competing insurers under a hypothetical network exclusion with a given insurer. This makes the simulation methods more similar to markets with a single (or a dominant) insurer, than to markets with numerous insurers. Hence, *a priori*, it seems likely that the merger simulation methods will perform less well in markets with a large number of insurers. To cover a reasonable range in the extent of insurer competition, we randomly assign the number of insurers in each market from the set {1, 3, 5, 7, 9}. As discussed in Section A5.1 and illustrated in Table A6, this expected pattern of worse performance as the number of insurers increases is exhibited only by $DWTP/Q$, though it performs quite well even in markets with nine insurers.

Table A2 summarizes the parameters of our theoretical model.

Table A2: List of Parameters

Parameter	Description	Set of Values
α	Hospitals' Share of Joint Surplus in Nash Bargaining Objective Function	0.4, 0.5, 0.6
γ_1, γ_2	Travel Cost Parameters in Consumer Preferences over Hospitals	(0.1,0.001), (0.3,0.003), (0.5,0.005)
θ	Price Sensitivity in Consumer Preferences over Insurers	0.5, 0.8, 1.1
λ	Hospital Network Sensitivity in Consumer Preferences over Insurers	2, 5, 8
$\#S$	The Number of Hospital Systems	6, 7, 8, 9, 10
$\#M$	The Number of Insurers	1, 3, 5, 7, 9
Z	Value of Non-inpatient Attributes in Consumer Preferences over Insurers	2, 5, 8
$\mathrm{E}[\eta_j]$	Expected Hospital Quality	14, 15, 16
$\mathrm{sd}[\eta_j]$	Population Standard Deviation of Hospital Quality	1.4, 1.6, 1.8
c	Expected Hospital Per Inpatient Event Cost	$8,000
p_z	Insurer Per Enrollee Expenses on Non-Inpatient Services	$3,200
τ	Insurer Administrative Cost per Inpatient Event	$500, $750, $1000
	Distribution of Consumer and Hospital Locations, Normal	N(0,9), N(0,8)
	Distribution of Consumer and Hospital Locations, Uniform	U[-16,16], U[-14,14]

A1.5 Insurance Buying Groups

We randomly assign the 500,000 consumers into 60,000 insurance buying groups. Specifically, we assign consumers into buying groups by drawing $u_g \sim U[0,1]$ for each group g and sequentially evaluating

$$\#I_g = \min\left\{\lfloor \exp\left\{0.75 + 6u_g^6\right\}\rfloor, 440,000 + g - \sum_{k=1}^{g-1} \#I_k\right\}.$$

That is, we assign the first $\#I_1$ consumers to buying group 1, the next $\#I_2$ to buying group 2, and so forth. Under this parameterization, roughly 9% of the consumers in our model buy insurance as individuals. Of those assigned to a buying group, the mean group size is typically around 30 and the maximum group size is typically around 850.

A1.6 Deriving the Distribution of Risk Types, $F(\rho)$

As discussed in Section 3, each consumer is randomly assigned a risk type, which captures their probability of needing inpatient hospital care, drawn from a parametric distribution, F_ρ. To benchmark the parameters of this distribution, we fit the density function to an empirical density defined on the frequency of inpatient events within discrete categories of consumers. We use the 2012 NHIS Public Use data to create the empirical distribution. We limit the NHIS sample to consumers covered by private insurance,[37] and use the *phospyr2* field as an indicator of whether the consumer had an inpatient event during that year, dropping any observation for which $phospyr2 > 2$ (don't know or refused). We aggregate the remaining data into 36 bins defined on gender and 5-year age categories, and use the frequency of $phospyr2 = 1$ to define the type, i.e., the probability of an inpatient event, for that bin. We define the empirical distribution of types by the distribution of NHIS data across the 36 bins.

We fit a logistic distribution by searching for location and scale parameters, a and b, respectively, to minimize the distance between moments and percentiles of the logistic and empirical distribution. Specifically, we minimize the distance between the means, standard deviations, and the 25th, 50th, and 75th percentiles. Based on the observed probabilities in the empirical distribution, we truncate the logistic distribution at 0.01 and 0.30. Given values of a and b, ρ_i is drawn as

$$\rho_i = a - b \ln\left(\left[u_i\left(\frac{1}{1+e^{-R}} - \frac{1}{1+e^{-L}}\right) + \frac{1}{1+e^{-L}}\right]^{-1} - 1\right) \tag{A3}$$

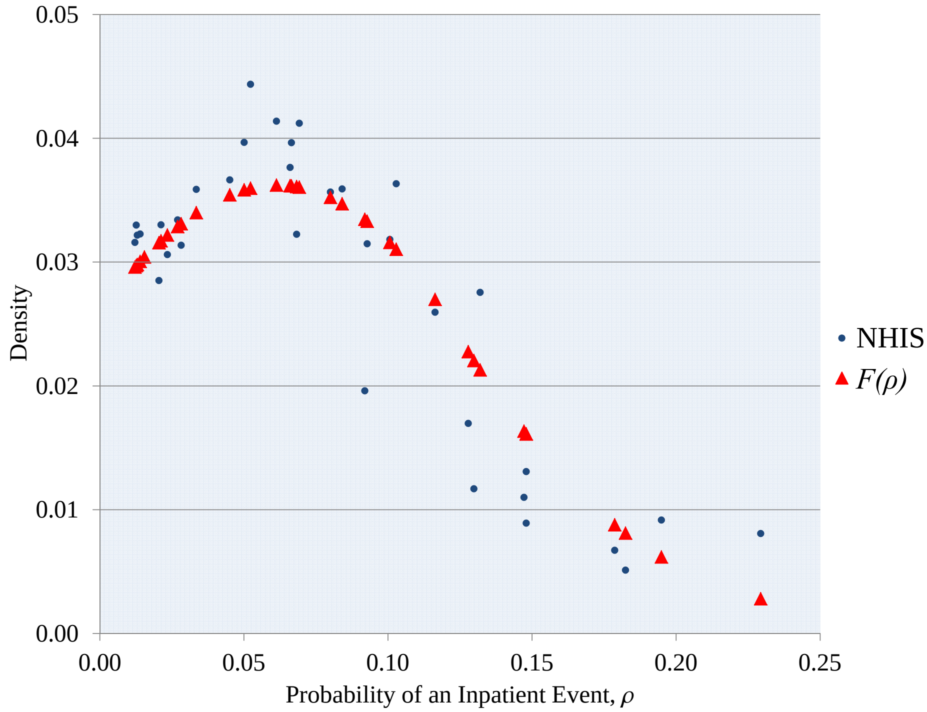
where $R \equiv \frac{0.30-a}{b}$, $L \equiv \frac{0.01-a}{b}$, and $u_i \sim U[0,1]$. Our minimum distance estimator produced the estimates $\widehat{a} = 0.01115$ and $\widehat{b} = 0.04096$. Figure A1 plots the empirical distribution of types from the NHIS and a kernel density of $F(\rho)$. Table A3 gives descriptive statistics.

Table A3: Descriptive Statistics of Type Distributions

	NHIS, 2012	$F(\rho)$
Mean	0.0643	0.0669
Standard Deviation	0.0450	0.0468
25th Percentile	0.0234	0.0312
50th Percentile	0.0613	0.0556
75th Percentile	0.0919	0.0904

[37]Specifically, we drop any observation for which the *private* field is greater than two.

Figure A1: $F(\rho)$ and the Empirical Distribution of Risk Types using NHIS 2012



A2 Sources of the Biases Exhibited by the Simulation Methods

In this Appendix, we provide an examination of the mechanisms underlying the bias exhibited by the simulation methods described in Section 6.1. As illustrated in Figure 1, WTP/Q exhibits a tendency to under-predict the true price effects, $DWTP/Q$ exhibits a tendency to over-predict the true price effects, and UPP exhibits a tendency to over-predict the true price effects when the true price effects are low, but exhibits an increasing tendency to under-predict the true price effects as the true price effects increase.

To explain these patterns, we make the following observations. First, the true price effects are convex in the diversion ratios between the merging hospitals. Similarly, changes in WTP are convex in the diversion ratios between the merging hospitals. In contrast, the predicted price effects of UPP are linear in the diversion ratios between the merging hospitals. Hence, it seems reasonable that UPP should be increasingly likely to under-predict the true price effects as the true price effects increase, while WTP/Q and $DWTP/Q$ would not necessarily exhibit this pattern. This is consistent with the finding that UPP follows a curved path in Figure 1, while WTP/Q and $DWTP/Q$ follow linear paths. That is, the biases of WTP/Q and $DWTP/Q$ are roughly constant fractions of the true price effect, but the bias of UPP becomes less positive or more negative as the true price effect increases.

Second, we note that a key distinction among the simulation methods is that only $DWTP/Q$ accounts for second order, or "feedback", effects through competing (non-merging) hospitals in estimating the post-merger price equilibrium. That is, only $DWTP/Q$ takes into account the fact that the first-order price increase for the merging hospitals will increase the prices of competing hospitals not involved in the merger, which in turn will feed back into additional (second order) pricing pressure for the merging hospitals.[38] This likely explains why the predicted price effects of $DWTP/Q$ are systematically higher than those of WTP/Q. It is also a source of negative bias for WTP/Q and UPP; the theoretical model incorporates these feedback effects, while WTP/Q and UPP do not.

Third, one notable feature of our theoretical model that is not accounted for by any of the simulation methods is that, in our theoretical model, insurers can adjust the profit maximizing premium under hypothetical exclusions of hospital systems. Specifically, the insurer's premiums are *not* constrained to be the same in the equilibrium payoff Π_n^J and the $\#S$ payoffs under which the insurer fails to reach an agreement with one of the $\#S$ hospital systems $\Pi_n^{J\setminus s}$. This ability to re-optimize the premium under an off-the-equilibrium-path exclusion of a given hospital system tends to reduce the system's bargaining leverage, both before and after the merger, because it allows insurers to mitigate the damage from exclusions. It also tends to reduce the price effects of mergers. Since none of the simulation methods account for this mechanism, it seems reasonable that the bias of all three simulation methods would be less positive or more negative if insurers were not able to re-optimize premiums under hypothetical exclusions.

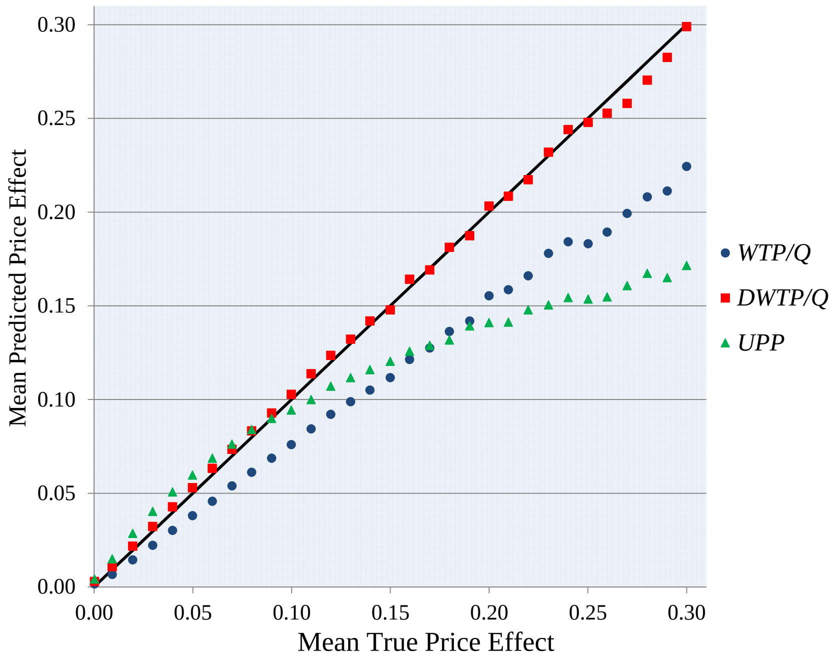
To demonstrate that this is a key source of bias, we computed the pre- and post-merger price equilibria for all 231,925 mergers in our 9,000 markets under the assumption that the insurers cannot re-optimize premiums under hypothetical exclusions. That is to say, we adopt an equilibrium concept under which $\pi_J^* = \pi_{J\setminus 1}^* = ... = \pi_{J\setminus \#S}^*$. Figure A2 gives the analog of Figure 1 under this restricted equilibrium concept. The figure shows that, as predicted, the paths of all three simulation methods are rotated towards the horizontal axis, indicating that the bias is lower (less positive or more negative). Moreover, the upward bias exhibited by $DWTP/Q$ in our baseline equilibrium concept is eliminated under this restricted equilibrium concept suggesting that this is the principal source of upward bias.

A3 Bayesian Inference

As discussed in Section 7, there are two possible ways that the merger simulation methods might fail to accurately predict real-world merger effects. The first is if the methods do not accurately

[38] A price increase at the merging hospitals will reduce the disagreement payoff for the insurer in bargaining with any competing (non-merging) hospital. This leads to an increase in the equilibrium price for competing (non-merging) hospitals.

Figure A2: Mean True and Predicted Price Effects
(Assuming Insurers cannot Re-Optimize Premiums under Hypothetical Exclusions)



predict the results of our theoretical model.[39] The second is if the theoretical model does not match the real world. The main purpose of this paper is to test the first one, and as discussed in Section 6 we find that the methods generally perform well.

The extent to which this finding increases the posterior probability that the simulation methods accurately predict real-world merger effects can be expressed in Bayesian terms as follows. Define A as "simulation methods predict theoretical model merger effects well," B as "theoretical model closely matches the real world," and C as "simulation methods predict real-world price effects well." Since each of these is binary (Y=Yes, N=No), there are eight possible combinations of ABC. Of these, only four have non-trivial probability of occurring (YYY YNN NYN NNN), so for convenience we set the other four probabilities to zero. By a straightforward application of Bayes' Rule, the prior probability P(C) = P(YYY)/(P(YYY)+P(YNN)+P(NYN)+P(NNN)), and the posterior probability P(C|A)=P(YYY)/(P(YYY)+P(YNN)). That is, a finding that A has occurred transfers probability mass from NYN and NNN to YYY and YNN.

It is easy to see that the posterior probability P(C|A), and the magnitude of the updating (P(C|A)- P(C)), both depend on the relative magnitudes of P(YYY) and P(YNN). P(C|A) can be anywhere from zero to unity depending on these relative magnitudes. That is, how much it matters that A occurred depends crucially on the probability of B.

[39]It is possible that the simulation methods would accurately predict real-world effect even if they predicted the effects of the theoretical model poorly, but there is no reason to believe that this would be the case.

While there is no decisive proof, there is reason to believe that our model is at least a reasonable approximation of reality (i.e., that B occurs with fairly high probability), justifying a relatively large updating and posterior probability conditional on the simulation methods accurately predicting the theoretical model. As noted above, our model, like other recent models, contains a number of features that are designed to capture the structure of real-world hospital markets in the United States. All of these models make the common and intuitive assumptions that insurance premiums and hospitals prices are simultaneously set in a differentiated Bertrand premium-setting game played by the insurers and via Nash-in-Nash Bargaining between hospitals and insurers. In addition, the parameterizations are set, to the extent possible, to match real-world metrics.

The posterior probability P(C|A) and the magnitude of the updating (P(C|A)- P(C)) also depend on the magnitude of P(NNN)+P(NYN). That is, the effect of a result that the simulation methods accurately predict real-world price effects depends on the prior probability that the simulation methods would accurately predict the theoretical model. If the model and the methods were so similar that this result was nearly guaranteed (i.e., if P(NNN)+P(NYN) was very small), then the magnitude of the update would be very small. For example, suppose that P(NNN)+P(NYN)=0.1. In that case, the update would be from P(YYY) to P(YYY)/.9=1.11P(Y), so a finding that A has occurred cause an update of only 11.1% relative to the prior probability. In contrast, suppose that P(NNN)+P(NYN)=0.9. In that case, the update would be from P(YYY) to P(YYY)/.1, so a finding that A has occurred would cause an update of 1000% relative to the (initially very small) prior probability. As discussed above, the simulation methods can be thought of as an approximation to the theoretical model. If the simulation methods were constructed so that this approximation was *necessarily* a very close one (i.e., if it was constructed so that P(NNN)+P(NYN) was very small), then it would be no surprise that they predicted the model's merger effects well, and then passing our test would generate a posterior probability that the simulation methods predict real-world price effects well that is only slightly higher than the prior probability. However, this is not the case. Though both our theoretical model and the simulation methods derive their basic intuition from bilateral bargaining theory (compare Section 3 and Section 5), they are dissimilar enough that the closeness of the approximation is not obvious, and therefore a finding that the approximation is in fact close justifies a positive updating in favor of the simulation methods' real-world usefulness.

There are a number of important features that are included in the theoretical model, but are not directly accounted for by the simulation methods. The absence of these features from the simulation methods is precisely what makes them relatively easy (and in the case of *UPP* very easy) to apply in real-world cases. These differences are numerous and substantial enough that this result was not guaranteed, and so finding the result constitutes meaningful evidence on which to update.

A list of the differences between the simulation methods and the theoretical model is as follows.

- First and most important, the simulation methods do not directly account for the role of insurers. Consumers decide whether to buy insurance, which insurer to buy it from, and whether that choice would be different if a particular hospital was missing from an otherwise preferred insurer's network. The insurers play a premium-setting game the outcome of which depends on the degree of insurer competition. These insurer-related factors affect the bargaining incentives of both the insurers and the hospitals, and hence they affect equilibrium hospital prices. But they are not directly reflected in the merger simulation methods, and this may be a source of prediction error. The magnitude of this prediction error may be a function of the insurer market structure, which as discussed above, we allow to range from one insurer to nine in the theoretical model. One manifestation of this is that, in the simulation methods, the predicted price effects of mergers necessarily go to zero as the diversion ratios between the merging hospitals approach zero. This is not necessarily the case in our theoretical model.

- Second, if the objective of insurers is to maximize profits (as is assumed in our theoretical model), then the regression model underlying the *WTP*-based simulation methods is misspecified, and so might not closely approximate the theoretical model. Formally, WTP/Q and $DWTP/Q$ assume that, gross of payments to hospitals, the insurer's payoff is simply proportional to the value consumers place on its provider network. The reasoning behind this is that a measure of the reduction in consumer valuation of an insurer's provider network due to the exclusion of a given hospital system may be a good proxy for the reduction in the insurer's gross profits, and hence effectively reflects the bargaining position of the insurer. We view this as a reasonable assumption, but the *WTP* metric is not guaranteed to be linearly related to the difference in insurer profits, as is assumed by the WTP/Q and $DWTP/Q$ methods.

- Third, the methods do not account for group purchases of health insurance. In the U.S., most private insurance is group insurance organized through an employer, and, therefore, reflects some aggregation of the preferences of the employees. The simulation methods, in contrast, implicitly assume individual health insurance choices are based on individual preferences.

- Fourth, the methods do not account for the role of non-inpatient healthcare services, and expenditures on those services, in consumers choice of whether to purchase insurance and which insurer to choose. This non-inpatient care (captured in our theoretical by the parameters Z and p_z) affects insurance demand and profits, which in turn affects equilibrium hospital prices.

- Fifth, the WTP/Q method and *UPP* do not account for the fact that the price responses of non-merging firms, and hence the post-merger equilibrium prices of the merging firms, will differ across markets, even holding constant the diversion ratios and gross margins between the merging firms. The $DWTP/Q$ method does account for this. This matters because such price responses tend to increase the predicted price effects in $DWTP/Q$ and, as discussed above, because $DWTP/Q$ takes into account the fact that hospitals that have higher priced rivals will themselves have higher prices, all else equal.

- Sixth, in our theoretical model, hospital prices are determined under three sources of uncertainty: (i) which consumers will buy insurance; (ii) which of the consumers who buy insurance will require inpatient care; and (iii) which hospital those patients will choose. In contrast, the simulation methods are applied to ex-post data on observed hospital discharges, which represents one realization of these uncertainties. If that realization happens to be unrepresentative, then the predictions of the simulation methods would not closely approximate the true price effects generated in the theoretical model.

- Seventh, the methods do not account for the possibility that, as discussed in Section 3.3, a merger between two hospitals has a complements effect as well as a substitutes effect, which in the theoretical model tends to push price effects downwards. However, the fact that few mergers in our analysis have true price effects that are negative, and that almost all of those that do have negative price effects also have extremely low diversion ratios, suggests that the complements effect is generally small, so this factor is likely not very important.

Given that it is not obvious a priori that our test *must* be passed, the fact that it *was* passed may justify a substantial updating of the probability that the simulation methods predict real-world price effects well enough to be considered in merger analysis. As noted above, the magnitude of this updating will also depend on one's priors regarding the probability that the model closely matches the real world. If one has strong priors that the model does not capture the real world well, or alternatively that our parameterizations of the model are highly inaccurate, then the magnitude of the updating will be small, and vice-versa.

A4 Full Dispersion Results

In this Appendix, we give the full set of results on the dispersion of the predicted price effects of the simulation methods. As discussed in Section 6.1, we group our 231,925 mergers into 31 categories defined by one percentage point increments of the true price effect $\frac{\Delta p_r}{p_r}$ (i.e., $\leq 0.5\%, (0.5\%, 1.5\%), (1.5\%, 2.5\%),\ ..., (29.5\%, 30.5\%)$). Following Table 4, we calculate the frequency in each category with which the merger simulation methods under- and over-predict the true price effect by more than 50% of the true price effect. Following Table 5, we give the MAPE ratio in each category for each of the merger simulation methods.

The results are given in Table A4. Columns (1), (4), and (7) give the frequency with which the merger simulation methods under-predict the true price effect by more than 50% of the true price effect. Columns (2), (5), and (8) give the frequency with which the merger simulation methods over-predict the true price effect by more than 50% of the true price effect. Columns (3), (6), and (9) give the MAPE rations. We find that each of the simulation methods perform poorly in the $< 0.5\%$ category, but the performance of all three improves rapidly as the true price effects increase. $DWTP/Q$ performs the best. It's MAPE ratio is consistently in the 10%-15% range for all categories of mergers above the $< 0.5\%$ category. The predicted price effects of $DWTP/Q$ are within 50% of the true price effect for 84.4% of the mergers in the $(0.5\%, 1.5\%)$ category, and this percentage increases to about 95% for mergers in the $(6.5\%, 7.5\%)$ category and above. WTP/Q also performs reasonably well. It's MAPE ratio gradually declines from about 0.29 in the $(0.5\%, 1.5\%)$ category, stabilizing in the 0.17-0.20 range for mergers in the $(9.5\%, 10.5\%)$ category and above. The predicted price effects of the WTP/Q are within 50% of the true price effect for 76.0% of the mergers in the $(0.5\%, 1.5\%)$ category, and this percentage increases to about 90%-95% for mergers in the $(4.5\%, 5.5\%)$ category and above. UPP performs less well overall and exhibits the pattern of significant upward bias when the true price effects are low and significant downward bias when the true price effects are high. The MAPE ratio of UPP declines from about 0.534 in the $(0.5\%, 1.5\%)$ category to 0.156 in the $(11.5\%, 12.5\%)$ category and above, and then increases to about 0.40 for mergers in the $(26.5\%, 27.5\%)$ category and above. The predicted price effects of the UPP are within 50% of the true price effect for only 39.8% of mergers in the $(0.5\%, 1.5\%)$ category. This percentage increases to 97.1% in the $(15.5\%, 16.5\%)$ category but then decreases to 70.0% in the $(29.5\%, 30.5\%)$ category. Consistent with results in Figure 2 on relative bias, UPP is far more likely to over-predict than under-predict the true price effects when the true price effects are low and vice versa when the true price effects are high.

Table A4: Dispersion of Prediction Price Effects and MAPE Ratios

$\frac{\Delta p_r}{p_r} \in$	*WTP/Q* (1) $\frac{\widehat{\Delta p_r}}{p_r} \leq \frac{\Delta p_r}{2p_r}$	(2) $\frac{\widehat{\Delta p_r}}{p_r} \geq \frac{3\Delta p_r}{2p_r}$	(3) MAPE	*DWTP/Q* (4) $\frac{\widehat{\Delta p_r}}{p_r} \leq \frac{\Delta p_r}{2p_r}$	(5) $\frac{\widehat{\Delta p_r}}{p_r} \geq \frac{3\Delta p_r}{2p_r}$	(6) MAPE	*UPP* (7) $\frac{\widehat{\Delta p_r}}{p_r} \leq \frac{\Delta p_r}{2p_r}$	(8) $\frac{\widehat{\Delta p_r}}{p_r} \geq \frac{3\Delta p_r}{2p_r}$	(9) MAPE
< 0.5%	0.196	0.402	0.873	0.011	0.524	0.652	0.034	0.775	1.647
(0.5%,1.5%)	0.180	0.059	0.290	0.001	0.155	0.141	0.016	0.586	0.534
(1.5%,2.5%)	0.113	0.050	0.282	0.001	0.110	0.151	0.005	0.497	0.483
(2.5%,3.5%)	0.093	0.042	0.267	0.001	0.088	0.148	0.001	0.409	0.404
(3.5%,4.5%)	0.075	0.039	0.252	0.001	0.074	0.146	0.002	0.331	0.327
(4.5%,5.5%)	0.065	0.037	0.246	0.001	0.066	0.144	0.001	0.262	0.278
(5.5%,6.5%)	0.062	0.030	0.233	0.000	0.061	0.139	0.001	0.207	0.235
(6.5%,7.5%)	0.060	0.028	0.230	0.001	0.056	0.140	0.002	0.163	0.213
(7.5%,8.5%)	0.050	0.029	0.230	0.000	0.051	0.135	0.004	0.125	0.188
(8.5%,9.5%)	0.058	0.019	0.215	0.001	0.040	0.139	0.008	0.088	0.181
(9.5%,10.5%)	0.046	0.017	0.209	0.000	0.044	0.138	0.009	0.054	0.165
(10.5%,11.5%)	0.055	0.025	0.219	0.001	0.048	0.128	0.015	0.055	0.165
(11.5%,12.5%)	0.065	0.011	0.207	0.001	0.032	0.122	0.011	0.026	0.156
(12.5%,13.5%)	0.049	0.009	0.207	0.002	0.038	0.123	0.029	0.026	0.181
(13.5%,14.5%)	0.039	0.016	0.203	0.000	0.040	0.135	0.018	0.012	0.173
(14.5%,15.5%)	0.043	0.016	0.212	0.002	0.036	0.127	0.024	0.012	0.197
(15.5%,16.5%)	0.050	0.010	0.200	0.000	0.033	0.135	0.025	0.004	0.197
(16.5%,17.5%)	0.048	0.020	0.204	0.004	0.046	0.128	0.029	0.002	0.200
(17.5%,18.5%)	0.029	0.013	0.200	0.000	0.051	0.123	0.051	0.003	0.210
(18.5%,19.5%)	0.051	0.010	0.195	0.000	0.048	0.125	0.065	0.000	0.259
(19.5%,20.5%)	0.033	0.013	0.194	0.000	0.029	0.135	0.079	0.000	0.246
(20.5%,21.5%)	0.017	0.013	0.169	0.004	0.030	0.113	0.051	0.000	0.303
(21.5%,22.5%)	0.049	0.000	0.172	0.000	0.032	0.132	0.135	0.000	0.300
(22.5%,23.5%)	0.027	0.014	0.176	0.000	0.041	0.110	0.082	0.000	0.311
(23.5%,24.5%)	0.031	0.000	0.196	0.000	0.016	0.106	0.116	0.000	0.325
(24.5%,25.5%)	0.079	0.000	0.206	0.009	0.035	0.121	0.132	0.000	0.360
(25.5%,26.5%)	0.052	0.013	0.203	0.000	0.052	0.110	0.156	0.000	0.339
(26.5%,27.5%)	0.029	0.010	0.143	0.000	0.029	0.134	0.216	0.000	0.407
(27.5%,28.5%)	0.048	0.000	0.175	0.012	0.024	0.117	0.214	0.000	0.373
(28.5%,29.5%)	0.016	0.000	0.173	0.000	0.048	0.120	0.194	0.000	0.413
(29.5%,30.5%)	0.030	0.015	0.171	0.000	0.045	0.125	0.303	0.000	0.405

A5 Robustness of the Results

A natural question is whether the performance of the simulation methods varies by competitive conditions in the hospital and insurance markets. In Appendix A5.1, we examine the sensitivity of our baseline results to such variation. To explore variation in hospital competition, we evaluate the MAPE ratios within categories of hospital mergers based on the pre-merger gross margin of the hospitals. To explore variation in insurer competition, we evaluate the MAPE ratios within categories of hospital mergers based on the number of insurers in the market. The results indicate that the merger simulation methods generally perform modestly less well under parameterizations in which hospitals have higher gross margins and when there is greater competition in the insurance market.

As noted above, the results presented in Section 6 are highly aggregated across the thousands of possible parameterizations discussed in Section 4. We chose those parameterizations in order to replicate the real world in some key metrics, including mean hospital gross margins and prices. At the same time, we included some parameterizations that may be considered too extreme to be plausible, in order to create a high probability that the parameters that correspond most closely to the real world would be included among them and to assess the performance of the simulation methods under what may be implausible parameterizations.[40]

A finding that the simulation methods perform well across most of this broad range of parameterizations does not imply that they perform well in the real world because, among other things, we do not know which sets of parameter values correspond most closely to the real world. Good performance in a large number of irrelevant parameterizations may be masking poor performance in a small number of relevant ones. To address this, in Appendix A5.2 we report more refined MAPE ratio results broken down by: (i) each possible value for each parameter in our model; and (ii) each of the categories of mergers based on the true price effects discussed in Table 3. Also, as discussed in footnote 24, the fact that we analyze the performance of the simulation methods within narrow categories of mergers based on the true price effect serves to mitigate this concern.

Overall, these refined results are very similar to the aggregate ones. In Appendix A5.2 we do not find that, conditional on any specific parameter value, the simulation methods perform poorly other than for mergers for which the true price effects is in the (0.5%,1.5%) category. That said, we do find some sensitivity of the results based on variation in some of the key model parameters, most notably, the insurance demand parameter λ. Consistent with our results by hospital gross margin quartiles, we find that the simulation methods, particularly $DWTP/Q$, perform less well when λ is high. So exactly how well the simulation methods perform does depend somewhat on where the real world lies in parameter space. But, other than for mergers for which the true price effects is in

[40]For example, as noted above, many of our parameterizations result in within-market mean hospital gross margins in excess of 0.7, which is likely to be unrealistically high.

(0.5%,1.5%), the simulation methods do not perform poorly conditional on any specific parameter value.

To further test the robustness of our results, we present in Appendix A5.3 seventeen additional sets of results under various modifications to our baseline parameterizations and assumptions. These include the alternative equilibrium concept discussed in Section A2, alternative values for the insurance demand parameters θ and λ, alternative assumptions on how consumers are aggregated into insurance buying groups, fewer hospitals and hospital systems, and measurement error in hospital system prices and costs. Broadly speaking, we find that our results are robust to these modifications. One noteworthy result is that while measurement error in prices modestly degrades the performance (as measured by the MAPE ratio) of WTP/Q and $DWTP/Q$, it does not degrade the performance of UPP.

A5.1 Performance by Level of Hospital and Insurer Competition

In this Appendix, we examine the sensitivity of our baseline results to such variation. To explore variation in hospital competition, we evaluate the MAPE ratios within categories of hospital mergers based on the level of pre-merger market power of the hospitals. To explore variation in insurer competition, we evaluate the MAPE ratios within categories of hospital mergers based on the number of insurers in the market.

Turning first to variation in pre-merger competitive conditions in the hospital market, we group mergers into the same five categories as above and divide each category into quartiles based on the volume-weighted pre-merger gross margins of the hospitals. We evaluate the MAPE ratio for each true price effect category-gross margin quartile combination.

The results are given in Table A5. The results indicate that the merger simulation methods generally perform less well under parameterizations in which hospitals have greater market power, though this is not uniformly the case. This pattern is most clearly exhibited by $DWTP/Q$. In the $(0.5\%, 1.5\%)$ category, the MAPE ratio of $DWTP/Q$ increases from 0.105 in the bottom quartile to 0.209 in the top quartile. This pattern is replicated in the $(4.5\%, 5.5\%)$, $(9.5\%, 10.5\%)$, and $(14.5\%, 15.5\%)$ categories, though the increases are more modest. This pattern is not replicated in the $(19.5\%, 20.5\%)$ category.

The MAPE ratio of WTP/Q is less sensitive to variation in the gross margins of hospitals than is the MAPE ratio of $DWTP/Q$. The pattern of higher MAPE ratios when gross margins are higher is exhibited in the $(0.5\%, 1.5\%)$ category (increasing from 0.273 in the lowest quartile to 0.355 in the highest quartile), and in the $(4.5\%, 5.5\%)$ category (from 0.227 in the lowest to 0.297 in the highest), but there is little systematic relationship between the MAPE ratio of WTP/Q and hospital gross margins in the higher true price effect categories.

UPP exhibits a pattern of increasing MAPE as hospital gross margins increase when the true price effects are relatively low but decreasing MAPE as hospital gross margins increase when the true price effects are relatively high. For example, in the (4.5%, 5.5%) category, the MAPE ratio of *UPP* increases from 0.137 in the bottom quartile to 0.515 in the top quartile. But in the (14.5%, 15.5%) category, the MAPE ratio of *UPP* decreases from 0.417 in the bottom quartile to 0.138 in the top quartile. As shown in Figure 2, *UPP* exhibits a negative bias when the true price effects are less than approximately 11% and a positive bias when the true price effects are greater than that. Table A5 shows that in the category of mergers in which *UPP* is closest to being unbiased (the 9.5%-10.5% category), the MAPE ratio of *UPP* is much less sensitive to variation in the gross margins of hospitals than it is in the other categories.

We note that the mean hospital gross margin in the top quartile is greater than 0.7, which seems very high. Therefore, it is likely that many of the parameterizations in this quartile are not representative of the real world.

The most likely reason why the simulation methods perform less well when hospital gross margins are higher lies in variation of the parameter λ. As discussed above, higher values of λ imply a greater loss in value for consumers from an exclusion of a given hospital system, and hence greater market power for hospitals, which is reflected in higher gross margins. Larger values of λ also increase the curvature in insurance demand (see equation (3)) with respect to the $EMAX$ terms that define the util value of the provider network. (See equation (4).) Since price is assumed to be linear in these $EMAX$ terms in both WTP/Q and $DWTP/Q$, greater curvature in insurance demand (3) with respect to the $EMAX$ term in the theoretical model should increase the prediction errors. (We note, however, that the reduction in performance as hospital gross margins increase is even greater for *UPP*, which does not directly rely on the $EMAX$ terms.) See Appendix A5.2 for results broken down by value of λ.

To test the sensitivity of our results to variation in competitive conditions in the insurance market, we evaluate the MAPE ratios in the five categories of mergers defined above and by the number of insurers in the market. One may expect our results to be sensitive to the number of insurers. This is because the theoretical model allows for consumers to switch insurers in response to the exclusion of a hospital system from an insurer's provider network, but the simulation methods do not. While this is generally a potential source of prediction error, the problem may be greater when there are more insurers. This is because more choices means that each consumer likely has a smaller gap between the first vs. second choice insurer, and so has a higher probability of switching insurers in response to an exclusion of a hospital system. On the other hand, the effect of variation in the level of competition in the insurance market may be captured by the simulation methods indirectly, e.g., through the gross margins of hospitals. So we have no clear prediction regarding

Table A5: MAPE Ratios by Hospital Gross Margin Quartiles

Mergers *s.t.* $\frac{\Delta p_r}{p_r} \in$	Quartile	N	Mean Hosp Gr Margin	WTP/Q	$DWTP/Q$	UPP
(0.5%,1.5%)	1^{st}	12,888	0.266	0.273	0.105	0.357
	2^{nd}	12,686	0.432	0.273	0.132	0.486
	3^{rd}	10,975	0.560	0.294	0.174	0.653
	4^{th}	9,358	0.700	0.355	0.209	0.940
(4.5%,5.5%)	1^{st}	947	0.289	0.227	0.129	0.137
	2^{nd}	1,436	0.441	0.231	0.141	0.216
	3^{rd}	1,577	0.566	0.236	0.142	0.323
	4^{th}	1,519	0.706	0.297	0.161	0.515
(9.5%,10.5%)	1^{st}	155	0.295	0.208	0.105	0.263
	2^{nd}	385	0.442	0.193	0.141	0.134
	3^{rd}	496	0.566	0.211	0.130	0.133
	4^{th}	545	0.706	0.217	0.153	0.206
(14.5%,15.5%)	1^{st}	45	0.308	0.220	0.117	0.417
	2^{nd}	115	0.441	0.190	0.117	0.253
	3^{rd}	194	0.565	0.199	0.126	0.184
	4^{th}	224	0.706	0.247	0.148	0.138
(19.5%,20.5%)	1^{st}	12	0.322	0.258	0.113	0.456
	2^{nd}	45	0.450	0.154	0.152	0.409
	3^{rd}	77	0.574	0.204	0.138	0.255
	4^{th}	105	0.714	0.229	0.123	0.177

how performance of the simulation methods will vary with the number of insurers. And as discussed below, the results were mixed in this regard.

The results are given in Table A6. For $DWTP/Q$, the MAPE ratio increases in the number of insurers within each of the five merger categories. For example, within the 4.5%-5.5% category, the MAPE increases from 0.094 for a single insurer to 0.168 for nine insurers. Even given this variation, the MAPE ratio for $DWTP/Q$ is quite low across all categories of mergers.

In contrast, WTP/Q does not exhibit a pattern of performing relatively less well when the number of insurers is large. Overall, the MAPE ratio of WTP/Q exhibits somewhat less sensitivity to the number of insurers (compared to $DWTP/Q$) and typically decreases in the number of insurers. For example, in the $(4.5\%, 5.5\%)$ category, the MAPE ratio of WTP/Q decreases from 0.280 when there is one insurer to 0.241 when there are nine.

UPP exhibits the pattern of performing less well when the number of insurers is large in the $(0.5\%, 1.5\%)$ and $(4.5\%, 5.5\%)$ categories, and to a lesser extent in the $(9.5\%, 10.5\%)$ category. But we find little evidence of a systematic relationship between the MAPE ratio of *UPP* and the number of insurers in the $(14.5\%, 15.5\%)$ and $(19.5\%, 20.5\%)$ categories of mergers.

Table A6: MAPE Ratios by Number of Insurers

Mergers *s.t.* $\frac{\Delta p_r}{p_r} \in$	# Insurers	N	Mean Hosp Gr Margin	*WTP/Q*	*DWTP/Q*	*UPP*
	1	10,675	0.495	0.326	0.105	0.461
	3	9,157	0.457	0.268	0.147	0.491
(0.5%,1.5%)	5	9,195	0.468	0.283	0.155	0.574
	7	8,394	0.464	0.284	0.168	0.606
	9	8,486	0.463	0.284	0.162	0.618
	1	1,246	0.550	0.280	0.094	0.248
	3	1,083	0.512	0.238	0.151	0.249
(4.5%,5.5%)	5	1,066	0.521	0.228	0.165	0.293
	7	982	0.519	0.245	0.171	0.307
	9	1,102	0.514	0.241	0.168	0.319
	1	372	0.577	0.231	0.085	0.146
	3	305	0.540	0.196	0.130	0.151
(9.5%,10.5%)	5	328	0.566	0.207	0.163	0.179
	7	304	0.545	0.203	0.183	0.171
	9	272	0.551	0.194	0.176	0.180
	1	138	0.575	0.251	0.095	0.203
	3	116	0.586	0.198	0.119	0.202
(14.5%,15.5%)	5	120	0.576	0.194	0.142	0.178
	7	108	0.576	0.241	0.150	0.211
	9	96	0.560	0.192	0.167	0.198
	1	67	0.627	0.252	0.086	0.222
	3	47	0.569	0.216	0.135	0.323
(19.5%,20.5%)	5	48	0.597	0.187	0.156	0.281
	7	35	0.604	0.173	0.152	0.211
	9	42	0.588	0.153	0.166	0.254

A5.2 Relative Bias and MAPE Ratios by Parameter Values

In this Appendix, we give the relative bias and MAPE ratio results conditional on specific values of the parameters in our theoretical model. We provide these results for the five categories of mergers discussed in Section 6.1. Table A7 gives the results for the category of mergers such that the true price effects lies in $(0.5\%, 1.5\%)$. Tables A8 through A11 give comparable results for mergers in the $(4.5\%, 5.5\%)$, $(9.5\%, 10.5\%)$, $(14.5\%, 15.5\%)$ and $(19/5\%, 20.5\%)$ categories, respectively. Throughout, we use the MAPE ratio, which measures the dispersion of the predicted price effects about the true price effects (or equivalently, the dispersion of the prediction errors about zero), as the main metric of performance.

With respect to the travel cost parameters (γ_1, γ_2), we find little variation in the performance of $DWTP/Q$ based on variation in these parameters for mergers in the $(4.5\%, 5.5\%)$ category and higher. For mergers in the $(0.5\%, 1.5\%)$ category, $DWTP/Q$ does perform less well in markets in which (γ_1, γ_2) are higher. WTP/Q exhibits the opposite pattern in that it's performance is not monotonically related to the values of (γ_1, γ_2) for mergers in the $(0.5\%, 1.5\%)$ category, but WTP/Q performs better when (γ_1, γ_2) take on their higher values in the $(4.5\%, 5.5\%)$ category and higher. UPP performs worse when (γ_1, γ_2) take on their higher values in the $(14.5\%, 15.5\%)$ and $(14.5\%, 15.5\%)$ categories only.

We find little variation in the performance of all three simulation methods based on variation in the value of the price sensitivity parameter θ. We view this result as significant because, in practice, little is known about the price sensitivity of consumers in the insurance market.

As discussed in Section A5.1, intuition suggests that the simulation methods should perform less well in markets in which the value of λ is high. However, the results indicate that this pattern is consistently manifested in $DWTP/Q$ only. In contrast, the performance of WTP/Q is largely invariant to variation in the value of λ. The performance of UPP exhibits the curious pattern of performing less well when λ is high and the true price effects are relatively low (see Tables A7 and A8), but performing better when λ is high and the true price effects are relatively high (see Tables A10 and A11).

We find the WTP/Q and $DWTP/Q$ perform better as the number of hospital systems in the market increases, but the performance of UPP is largely invariant to the number of hospital systems. This could be explained by the fact that an additional hospital system adds another degree of freedom in the regression models underlying WTP/Q and $DWTP/Q$. Ceteris paribus, this additional observation would increase the precision of the predicted price effects of WTP/Q and $DWTP/Q$ but is irrelevant for UPP. However, it seems unlikely that this consideration is the only meaningful explanation since the magnitudes of the bias of WTP/Q and $DWTP/Q$ also decrease as the number of hospital systems increase. While additional degrees of freedom should increase the precision of the predicted price effects, it is unclear why additional degrees of freedom would affect bias.

We discuss the results based on variation in the number of insurers in Section A5.1.

We find little variation in the performance of all three simulation methods based on variation in the value of non-inpatient care attributes of insurance Z. We view this result as significant because, in practice, little is known about the relative value consumers place on inpatient care versus non-inpatient care attributes in their insurance choices.

Finally, we find little variation in the performance of all three simulation methods based on variation in the values of: the mean of the hospital quality distribution $E[\eta_j]$, the standard deviation of the hospital quality distribution $sd[\eta_j]$, the type of location distribution (Uniform or Normal), and the administrative cost incurred by insurers τ.

Table A7: Relative Bias and MAPE Ratios by Parameter Values

Mergers $s.t.$ $\frac{\Delta p_r}{p_r} \in (0.5\%, 1.5\%)$

		N	Mean Hosp Gr Margin	WTP/Q Rel. Bias	WTP/Q MAPE	$DWTP/Q$ Rel. Bias	$DWTP/Q$ MAPE	UPP Rel. Bias	UPP MAPE
α	0.4	16,399	0.393	-0.183	0.243	0.248	0.168	0.392	0.300
	0.5	15,137	0.473	-0.169	0.292	0.253	0.110	0.770	0.578
	0.6	14,371	0.557	-0.176	0.359	0.229	0.136	1.275	0.930
γ_1, γ_2	0.1,0.001	19,767	0.441	-0.298	0.291	0.082	0.099	0.753	0.540
	0.3,0.003	14,675	0.472	-0.115	0.263	0.297	0.172	0.793	0.515
	0.5,0.005	11,465	0.519	-0.044	0.331	0.454	0.259	0.861	0.549
θ	0.5	15,664	0.567	-0.184	0.298	0.245	0.140	0.838	0.558
	0.8	15,398	0.456	-0.177	0.286	0.244	0.138	0.776	0.523
	1.1	14,845	0.384	-0.168	0.287	0.241	0.145	0.764	0.522
λ	2	15,891	0.313	-0.261	0.301	0.103	0.102	0.526	0.426
	5	15,757	0.506	-0.169	0.281	0.258	0.150	0.813	0.576
	8	14,259	0.606	-0.089	0.287	0.385	0.194	1.068	0.656
# Hospital Systems	5	3,003	0.470	-0.247	0.403	0.342	0.172	0.894	0.496
	6	4,437	0.469	-0.217	0.358	0.315	0.166	0.832	0.507
	7	6,420	0.467	-0.201	0.320	0.257	0.146	0.819	0.515
	8	8,857	0.468	-0.176	0.292	0.240	0.142	0.766	0.525
	9	10,661	0.480	-0.162	0.274	0.218	0.132	0.802	0.562
	10	12,529	0.467	-0.144	0.249	0.213	0.131	0.754	0.541
# Insurers	1	10,675	0.495	-0.304	0.326	0.066	0.105	0.504	0.461
	3	9,157	0.457	-0.172	0.268	0.248	0.147	0.703	0.491
	5	9,195	0.468	-0.146	0.283	0.291	0.155	0.877	0.574
	7	8,394	0.464	-0.118	0.284	0.321	0.168	0.952	0.606
	9	8,486	0.463	-0.110	0.284	0.335	0.162	1.006	0.618
Z	2	14,845	0.477	-0.144	0.283	0.283	0.158	0.883	0.608
	5	15,506	0.467	-0.187	0.292	0.231	0.135	0.767	0.511
	8	15,556	0.467	-0.196	0.295	0.218	0.133	0.733	0.492
$E[\eta_j]$	14	14,924	0.470	-0.180	0.286	0.241	0.144	0.797	0.539
	15	15,280	0.473	-0.179	0.291	0.243	0.140	0.791	0.541
	16	15,703	0.468	-0.170	0.293	0.246	0.139	0.791	0.525
$sd[\eta_j]$	1.4	16,110	0.466	-0.177	0.278	0.226	0.138	0.789	0.544
	1.6	15,571	0.471	-0.171	0.289	0.243	0.138	0.803	0.545
	1.8	14,226	0.476	-0.180	0.306	0.264	0.149	0.786	0.509
Location Distribution	Uniform	22,153	0.471	-0.198	0.300	0.233	0.138	0.791	0.533
	Normal	23,754	0.470	-0.155	0.282	0.253	0.144	0.795	0.536
τ	0.50	15,320	0.467	-0.180	0.290	0.243	0.141	0.767	0.529
	0.75	15,584	0.473	-0.177	0.287	0.242	0.140	0.803	0.539
	1.00	15,003	0.471	-0.171	0.293	0.246	0.142	0.809	0.536

Table A8: Relative Bias and MAPE Ratios by Parameter Value

Mergers $s.t.$ $\frac{\Delta p_r}{p_r} \in (4.5\%, 5.5\%)$

			Mean Hosp	*WTP/Q*		*DWTP/Q*		*UPP*	
		N	Gr Margin	Rel. Bias	MAPE	Rel. Bias	MAPE	Rel. Bias	MAPE
	0.4	1,963	0.450	-0.137	0.191	0.194	0.180	0.085	0.142
α	0.5	1,805	0.533	-0.151	0.255	0.173	0.118	0.347	0.306
	0.6	1,711	0.600	-0.175	0.333	0.137	0.115	0.651	0.533
	0.1,0.001	1,284	0.496	-0.294	0.296	0.096	0.136	0.461	0.362
γ_1, γ_2	0.3,0.003	1,972	0.512	-0.161	0.225	0.161	0.143	0.301	0.245
	0.5,0.005	2,223	0.550	-0.066	0.231	0.219	0.153	0.324	0.260
	0.5	2,149	0.600	-0.158	0.248	0.163	0.139	0.408	0.318
θ	0.8	1,725	0.503	-0.157	0.246	0.168	0.142	0.330	0.261
	1.1	1,605	0.444	-0.144	0.241	0.179	0.152	0.287	0.252
	2	1,414	0.357	-0.219	0.259	0.078	0.100	0.126	0.194
λ	5	2,072	0.538	-0.154	0.241	0.164	0.147	0.367	0.297
	8	1,993	0.628	-0.107	0.243	0.240	0.182	0.486	0.354
	5	666	0.523	-0.237	0.352	0.205	0.165	0.400	0.262
	6	804	0.518	-0.180	0.287	0.215	0.174	0.369	0.288
# Hospital	7	884	0.524	-0.166	0.260	0.179	0.151	0.367	0.272
Systems	8	969	0.511	-0.141	0.232	0.158	0.140	0.307	0.271
	9	1,024	0.534	-0.127	0.225	0.146	0.125	0.359	0.304
	10	1,132	0.529	-0.112	0.198	0.138	0.125	0.312	0.274
	1	1,246	0.550	-0.240	0.280	0.051	0.094	0.255	0.248
	3	1,083	0.512	-0.168	0.238	0.161	0.151	0.288	0.249
# Insurers	5	1,066	0.521	-0.127	0.228	0.205	0.165	0.377	0.293
	7	982	0.519	-0.103	0.245	0.228	0.171	0.418	0.307
	9	1,102	0.514	-0.113	0.241	0.224	0.168	0.422	0.319
	2	1,781	0.531	-0.133	0.238	0.189	0.158	0.384	0.309
Z	5	1,797	0.525	-0.160	0.248	0.161	0.136	0.350	0.275
	8	1,901	0.517	-0.168	0.252	0.159	0.139	0.313	0.255
	14	1,806	0.523	-0.158	0.250	0.167	0.140	0.361	0.288
E$[\eta_j]$	15	1,786	0.528	-0.147	0.243	0.173	0.145	0.353	0.281
	16	1,887	0.522	-0.157	0.247	0.168	0.146	0.331	0.267
	1.4	1,932	0.523	-0.159	0.243	0.156	0.140	0.338	0.277
sd$[\eta_j]$	1.6	1,871	0.524	-0.149	0.241	0.172	0.145	0.345	0.277
	1.8	1,676	0.526	-0.153	0.263	0.182	0.149	0.362	0.281
Location	Uniform	2,664	0.525	-0.172	0.264	0.159	0.138	0.351	0.280
Distribution	Normal	2,815	0.523	-0.136	0.234	0.179	0.149	0.346	0.276
	0.50	1,921	0.523	-0.158	0.248	0.168	0.140	0.345	0.268
τ	0.75	1,809	0.525	-0.159	0.240	0.165	0.146	0.338	0.277
	1.00	1,749	0.524	-0.144	0.250	0.176	0.145	0.362	0.289

Table A9: Relative Bias and MAPE Ratios by Parameter Value

Mergers *s.t.* $\frac{\Delta p_r}{p_r} \in (9.5\%, 10.5\%)$

			Mean Hosp	*WTP/Q*		*DWTP/Q*		*UPP*	
		N	Gr Margin	Rel. Bias	MAPE	Rel. Bias	MAPE	Rel. Bias	MAPE
α	0.4	544	0.495	-0.125	0.171	0.184	0.180	-0.121	0.152
	0.5	548	0.558	-0.138	0.214	0.153	0.115	0.045	0.147
	0.6	489	0.626	-0.182	0.288	0.101	0.109	0.208	0.204
γ_1, γ_2	0.1,0.001	210	0.511	-0.289	0.301	0.090	0.146	0.182	0.155
	0.3,0.003	552	0.553	-0.158	0.199	0.153	0.135	0.040	0.162
	0.5,0.005	819	0.572	-0.104	0.190	0.159	0.133	0.001	0.172
θ	0.5	648	0.621	-0.146	0.207	0.131	0.130	0.083	0.156
	0.8	516	0.540	-0.153	0.228	0.151	0.135	0.026	0.174
	1.1	417	0.480	-0.142	0.190	0.168	0.147	-0.015	0.163
λ	2	332	0.391	-0.213	0.240	0.059	0.081	-0.146	0.182
	5	605	0.562	-0.154	0.205	0.141	0.139	0.052	0.150
	8	644	0.638	-0.107	0.189	0.199	0.176	0.120	0.171
# Hospital Systems	5	243	0.549	-0.189	0.258	0.208	0.220	0.063	0.197
	6	249	0.556	-0.191	0.240	0.177	0.171	0.042	0.157
	7	256	0.549	-0.151	0.223	0.155	0.140	0.015	0.169
	8	279	0.556	-0.144	0.204	0.131	0.133	0.013	0.151
	9	283	0.565	-0.136	0.179	0.108	0.103	0.038	0.156
	10	271	0.568	-0.081	0.162	0.117	0.111	0.061	0.172
# Insurers	1	372	0.577	-0.208	0.231	0.038	0.085	0.011	0.146
	3	305	0.540	-0.175	0.196	0.140	0.130	0.015	0.151
	5	328	0.566	-0.125	0.207	0.190	0.163	0.055	0.179
	7	304	0.545	-0.123	0.203	0.184	0.183	0.058	0.171
	9	272	0.551	-0.087	0.194	0.214	0.176	0.060	0.180
Z	2	512	0.564	-0.132	0.219	0.166	0.149	0.071	0.156
	5	515	0.567	-0.151	0.202	0.134	0.130	0.040	0.163
	8	554	0.541	-0.158	0.206	0.143	0.134	0.007	0.172
$E[\eta_j]$	14	493	0.554	-0.156	0.211	0.147	0.133	0.045	0.157
	15	548	0.560	-0.150	0.203	0.144	0.136	0.032	0.175
	16	540	0.556	-0.136	0.210	0.151	0.141	0.038	0.162
$sd[\eta_j]$	1.4	520	0.565	-0.131	0.178	0.147	0.134	0.038	0.165
	1.6	526	0.557	-0.152	0.217	0.145	0.130	0.043	0.174
	1.8	535	0.550	-0.158	0.222	0.151	0.141	0.034	0.156
Location Distribution	Uniform	814	0.553	-0.177	0.228	0.126	0.125	0.027	0.159
	Normal	767	0.562	-0.116	0.187	0.170	0.153	0.050	0.171
τ	0.50	529	0.560	-0.158	0.220	0.145	0.130	0.028	0.159
	0.75	529	0.552	-0.137	0.207	0.157	0.155	0.035	0.177
	1.00	523	0.559	-0.146	0.202	0.141	0.127	0.052	0.157

Table A10: Relative Bias and MAPE Ratios by Parameter Value

Mergers *s.t.* $\frac{\Delta p_r}{p_r} \in (14.5\%, 15.5\%)$

		N	Mean Hosp Gr Margin	*WTP/Q* Rel. Bias	*WTP/Q* MAPE	*DWTP/Q* Rel. Bias	*DWTP/Q* MAPE	*UPP* Rel. Bias	*UPP* MAPE
α	0.4	216	0.517	-0.114	0.157	0.170	0.152	-0.250	0.256
	0.5	210	0.582	-0.163	0.218	0.147	0.116	-0.118	0.177
	0.6	152	0.649	-0.180	0.304	0.077	0.103	0.012	0.124
γ_1, γ_2	0.1,0.001	63	0.562	-0.313	0.329	0.108	0.136	0.047	0.142
	0.3,0.003	173	0.566	-0.169	0.200	0.151	0.156	-0.164	0.208
	0.5,0.005	342	0.582	-0.109	0.190	0.136	0.113	-0.150	0.210
θ	0.5	277	0.627	-0.153	0.224	0.141	0.133	-0.073	0.162
	0.8	182	0.545	-0.164	0.219	0.130	0.128	-0.168	0.211
	1.1	119	0.501	-0.118	0.205	0.139	0.120	-0.219	0.235
λ	2	102	0.397	-0.195	0.275	0.071	0.086	-0.274	0.300
	5	238	0.576	-0.167	0.204	0.133	0.126	-0.138	0.192
	8	238	0.650	-0.111	0.220	0.169	0.161	-0.067	0.162
# Hospital Systems	5	116	0.565	-0.195	0.289	0.193	0.187	-0.080	0.199
	6	97	0.574	-0.193	0.248	0.156	0.146	-0.142	0.211
	7	89	0.572	-0.167	0.190	0.134	0.120	-0.152	0.239
	8	91	0.574	-0.142	0.201	0.127	0.112	-0.153	0.208
	9	87	0.586	-0.088	0.189	0.098	0.116	-0.128	0.169
	10	98	0.581	-0.098	0.162	0.099	0.087	-0.157	0.155
# Insurers	1	138	0.575	-0.219	0.251	0.043	0.095	-0.157	0.203
	3	116	0.586	-0.134	0.198	0.137	0.119	-0.142	0.202
	5	120	0.576	-0.117	0.194	0.184	0.142	-0.104	0.178
	7	108	0.576	-0.145	0.241	0.175	0.150	-0.125	0.211
	9	96	0.560	-0.112	0.192	0.172	0.167	-0.133	0.198
Z	2	183	0.582	-0.148	0.206	0.143	0.138	-0.100	0.167
	5	190	0.567	-0.143	0.208	0.132	0.127	-0.155	0.208
	8	205	0.577	-0.157	0.242	0.137	0.126	-0.142	0.215
E$[\eta_j]$	14	173	0.573	-0.153	0.228	0.124	0.121	-0.125	0.214
	15	209	0.565	-0.152	0.220	0.141	0.140	-0.156	0.202
	16	196	0.588	-0.143	0.208	0.145	0.127	-0.116	0.166
sd$[\eta_j]$	1.4	196	0.576	-0.134	0.204	0.128	0.116	-0.134	0.187
	1.6	168	0.573	-0.162	0.221	0.123	0.127	-0.160	0.207
	1.8	214	0.576	-0.154	0.219	0.156	0.137	-0.112	0.198
Location Distribution	Uniform	310	0.577	-0.174	0.222	0.122	0.125	-0.144	0.202
	Normal	268	0.573	-0.121	0.205	0.154	0.128	-0.121	0.183
τ	0.50	190	0.570	-0.147	0.199	0.152	0.142	-0.137	0.204
	0.75	208	0.568	-0.136	0.215	0.133	0.113	-0.139	0.196
	1.00	180	0.588	-0.167	0.244	0.127	0.132	-0.122	0.194

Table A11: Relative Bias and MAPE Ratios by Parameter Value

Mergers $s.t.\ \frac{\Delta p_r}{p_r} \in (19.5\%, 20.5\%)$

			Mean Hosp	*WTP/Q*		*DWTP/Q*		*UPP*	
		N	Gr Margin	Rel. Bias	MAPE	Rel. Bias	MAPE	Rel. Bias	MAPE
α	0.4	69	0.514	-0.104	0.144	0.183	0.162	-0.380	0.386
	0.5	100	0.599	-0.164	0.206	0.146	0.128	-0.261	0.250
	0.6	70	0.684	-0.231	0.277	0.053	0.112	-0.121	0.159
γ_1, γ_2	0.1,0.001	15	0.550	-0.298	0.344	0.087	0.146	-0.117	0.139
	0.3,0.003	74	0.578	-0.209	0.230	0.112	0.138	-0.248	0.265
	0.5,0.005	150	0.615	-0.132	0.173	0.142	0.132	-0.271	0.240
θ	0.5	106	0.641	-0.171	0.190	0.101	0.130	-0.223	0.221
	0.8	82	0.578	-0.184	0.210	0.149	0.145	-0.278	0.271
	1.1	51	0.547	-0.126	0.220	0.156	0.131	-0.280	0.288
λ	2	31	0.413	-0.233	0.232	0.078	0.094	-0.385	0.409
	5	100	0.573	-0.153	0.179	0.142	0.147	-0.290	0.267
	8	108	0.677	-0.159	0.206	0.132	0.130	-0.183	0.200
# Hospital Systems	5	57	0.570	-0.196	0.249	0.198	0.186	-0.259	0.265
	6	46	0.614	-0.233	0.276	0.122	0.140	-0.278	0.241
	7	28	0.589	-0.115	0.122	0.162	0.190	-0.237	0.271
	8	40	0.617	-0.138	0.189	0.098	0.104	-0.229	0.184
	9	37	0.601	-0.113	0.133	0.084	0.107	-0.279	0.292
	10	31	0.615	-0.158	0.164	0.080	0.082	-0.229	0.217
# Insurers	1	67	0.627	-0.243	0.252	0.038	0.086	-0.235	0.222
	3	47	0.569	-0.183	0.216	0.132	0.135	-0.298	0.323
	5	48	0.597	-0.123	0.187	0.180	0.156	-0.255	0.281
	7	35	0.604	-0.107	0.173	0.183	0.152	-0.204	0.211
	9	42	0.588	-0.123	0.153	0.169	0.166	-0.277	0.254
Z	2	76	0.617	-0.163	0.226	0.118	0.132	-0.229	0.221
	5	87	0.584	-0.125	0.170	0.159	0.143	-0.257	0.251
	8	76	0.599	-0.217	0.232	0.108	0.126	-0.276	0.272
$\mathrm{E}[\eta_j]$	14	88	0.595	-0.164	0.191	0.131	0.146	-0.257	0.241
	15	80	0.602	-0.161	0.194	0.132	0.125	-0.269	0.276
	16	71	0.602	-0.174	0.203	0.124	0.137	-0.234	0.221
$\mathrm{sd}[\eta_j]$	1.4	62	0.604	-0.177	0.181	0.100	0.113	-0.254	0.219
	1.6	81	0.600	-0.194	0.196	0.108	0.137	-0.231	0.248
	1.8	96	0.596	-0.136	0.203	0.166	0.144	-0.274	0.251
Location Distribution	Uniform	134	0.591	-0.164	0.201	0.130	0.122	-0.273	0.265
	Normal	105	0.610	-0.169	0.183	0.129	0.155	-0.231	0.222
τ	0.50	77	0.581	-0.150	0.172	0.147	0.147	-0.298	0.274
	0.75	81	0.612	-0.177	0.232	0.121	0.113	-0.231	0.215
	1.00	81	0.604	-0.170	0.201	0.121	0.138	-0.236	0.237

A5.3 Modifications to Baseline Parameterizations and Assumptions

In this Appendix, we present relative bias and MAPE results under seventeen modifications to our baseline parameterizations and assumptions. For each modification, we replicate our results for all 231,925 mergers in our 9,000 simulated hospital markets. As in Section 6, we present these results for categories of mergers, indexed by r, for which the true price effect, denoted $\frac{\Delta p_r}{p_r}$, lies in the following ranges: (0.5%,1.5%), (4.5%,5.5%), (9.5%,10.5%), (14.5%,15.5%), and (19.5%,20.5%). We also list the mean hospital gross margin under each modification to illustrate how each modification affects, on average, the market power of hospital systems. In each of the tables below, we include the results from our baseline model in the top block to facilitate comparison.

In our first modification, denoted M1 in the Table A12, we modify the equilibrium concept by assuming that insurers cannot re-optimize premiums under hypothetical exclusions of hospital systems. This modification is discussed in Section A2. As illustrated in Figure A2, we find that the bias exhibited by each of the simulation methods becomes more negative under this restricted equilibrium concept. Of particular interest is that fact that the positive bias exhibited by $DWTP/Q$ is eliminated. The MAPE ratio of $DWTP/Q$ is also significantly lower compared to our baseline results.

In modifications M2-M6, we assume different sets of possible values of the key parameters in consumers' preferences over insurers, θ and λ. In M2 and M3, we use lower and higher values of θ, respectively, compared to our baseline parameterization. In M2, we draw of θ from {0.4,0.7,1.0} instead of {0.5,0.8,1.1}. In M3, we draw θ from {0.6,0.9,1.2}. In M4 and M5, we use higher and lower values of λ, respectively, compared to our baseline parameterization. In M4, we draw λ from {3,6,9} instead of {2,5,8} In M5, we draw λ from {1,4,7}. In M6, we draw θ from {0.6,0.9,1.2} and λ from {3,6,9}. As expected, we find that hospital gross margins are higher when consumers are less price sensitive (θ is lower), and that hospital gross margins are lower when consumers are more price sensitive (θ is higher). Similarly, we find that hospital gross margins are higher when consumers are more sensitive to reductions in the value of the provider network (λ is higher), and that hospital gross margins are lower when consumers are less sensitive to reductions in the value of the provider network (λ is lower). Generally, we find that our baseline results are robust to these alternative values of θ and λ.

In M7, we reduce the number of hospitals in our markets from 12 to 8 and the number of hospital systems from 5-10 to 4-7. We find that this modification does reduce the performance of WTP/Q and $DWTP/Q$ by a small amount but does not materially affect the performance of UPP. One possible explanation is that reducing the number of systems in each market reduces the number of observations in the regression models underlying WTP/Q and $DWTP/Q$, making the predictions of those methods less precise. This is not a relevant consideration for UPP.

Turning to Table A13, we explore the sensitivity of our results under alternative groupings of consumers into insurance buying groups in M8 and M9. (See Appendix A1.5 for a discussion of our baseline approach to defining insurance buying groups.) In M8, we assume that all consumers buy insurance as individuals. This modification is of particular interest, since none of the three simulation methods directly account for the fact that most consumers purchase health insurance through groups. Hence, one might expect the simulation methods to perform better under this modification. Surprisingly, we find the opposite result for $DWTP/Q$. While the MAPE ratios for WTP/Q under this modification are similar to our baseline results, the MAPE ratios for $DWTP/Q$ are significantly higher compared to our baseline results. The results for UPP are somewhat mixed.

In M9, we increase the extent to which consumers are aggregated into insurance buying groups by assuming that each of the 500,000 consumers is randomly assigned to one of 5,000 insurance buying groups of size 100. We find that this modification has little effect on the performance of WTP/Q and UPP, but the performance of $DWTP/Q$ is slightly better compared to our baseline results.

In M10, we test the robustness of our results to misspecification of the model of consumer preferences over hospitals. (See equation (A2).) Specifically, we assume that the true travel cost parameters (γ_1, γ_2) vary across consumers, but the analyst does nothing to account for this heterogeneity. Instead of assuming that (γ_1, γ_2) take on the values (0.1,001), (0.3,0.003), or (0.5,0.005) and are constant across consumer within a simulated market, we assume that for each consumer

$$\gamma_{1i} \sim N(0.3, 0.05) \text{ and } \gamma_{2i} = 0.001\gamma_{1i}.^{41} \tag{A4}$$

We assume that the analyst simply estimates the discrete choice model underlying WTP and the diversion ratios, ignoring the true underlying heterogeneity in travel cost parameters. We find that this misspecification does little to reduce the performance of WTP/Q and UPP. It does reduce the performance of $DWTP/Q$ by a significant amount for mergers in the (0.5%,1.5%) category but by only a small amount for the other categories of mergers. For the categories (4.5%,5.5%) and higher, the MAPE ratio of $DWTP/Q$ remains below 0.20.

In M11, we assume that travel costs are linear in the distance between the consumer and the hospitals, as opposed to quadratic. That is, we assume $\gamma_2 = 0$. We find that this modification has almost no effect on our results.

In M12, we test whether our results are sensitive to a different distribution of risk types F_ρ. Specifically, we assume that each consumer has the same probability of requiring inpatient care, and this probability is equal to the expected value of ρ_i in our baseline model. We find that this modification has almost no effect on our results.

[41] We winsorize the draws of γ_{1i} at 0.1 and 0.5. The probability of winsorization is approximately 6.33E-5 or about 32 of the 500,000 consumers.

Finally, in M13 and M14, we test whether our results are sensitive to a significant increase in consumers' valuation of healthcare not related to inpatient care Z and expenditures on that care p_z. Specifically, we increase the values of Z from {2,5,8} to {4,7,10} in M13 and the value of p_z from \$3,200 to \$5,000 in M14. We find that these modifications have almost little effect on our results. WTP/Q performs slightly worse than it does in our baseline results, $DWTP/Q$ performs slightly better than it does in our baseline results. The performance is UPP is largely unchanged.

Turning to Table A14, we explore the sensitivity of our results to measurement error in hospital system prices and costs. As noted above, we assume that hospital system prices and costs are observed without error in our baseline results. In the real world, prices and costs may be observed with meaningful measurement error. This is likely to degrade the performance of the simulation methods to at least some degree.

In (M15), we assume that hospital system prices within a given market are observed with an IID Normal mean zero error. Hence, we assume that the observed price for hospital system j is

$$p_j^{observed} = p_j + error_j^p,$$

where p_j denotes the true equilibrium price generated in our theoretical model and $error_j^p \sim N(0, v^p)$. We assume that v^p is proportional to the standard deviation of hospital system prices in the market. While we have no way to characterize how much measurement error an analyst would typically encounter in practice, we introduce what appears to us to be a reasonable amount of error by scaling this standard deviation so that, on average, the true hospital system prices in each market explain about 90% of the variation in the observed hospital system prices. The scaling that meets this standard in our simulations is to set v^p equal to 0.35 times the standard deviation of hospital system prices in the market.

In (M16), we assume that hospital system costs within a given market are observed with an IID Normal mean zero error.

$$c_j^{observed} = c_j + error_j^c,$$

where c_j denotes the true hospital system cost in our theoretical model and $error_j^c \sim N(0, v^c)$. Here, we assume that v^c equals the average standard deviation (across markets) of hospital costs c_j. Hence, we set $v^c = 0.3$. Given this assumption, the true hospital system costs explain about 52% of the variation in the observed hospital costs within each market, on average.[42]

[42] The reason we chose this level of measurement error in cost is as follows. In testing the effect of measurement in cost, we found that measurement error in costs had little effect on the performance of the simulation methods. Therefore, we chose a value of v^c that results in as much measurement error in hospitals costs as there is true variation in hospital costs. We view this as likely a high amount of measurement error.

In (M17), we assume that both hospital system prices and costs are measured with error, with v^p still set to 0.35 times the standard deviation of hospital system prices in the market and v^c still set to 0.3.

We find that measurement error in prices degrades the performance (as measured by the MAPE ratio) of WTP/Q and $DWTP/Q$. However, the degradation is not so great that the simulation methods become unreliable under the amount of measurement error we apply here. Measurement error in costs results in a smaller degradation in the performance of $DWTP/Q$, but actually improves the performance of WTP/Q. Combining measurement error in prices and costs results in about the same level performance as measurement error in price alone for both WTP/Q and $DWTP/Q$. In contrast, we find that neither measurement error in prices nor costs, or price and costs combined, degrades the performance of UPP.

Table A12: Results Summary Under Modifications to Baseline Parameterizations and Assumptions

Modification	Mean Hosp Gr Margin	$\frac{\Delta p_r}{p_r} \in$	*WTP/Q* Rel. Bias	*WTP/Q* MAPE	*DWTP/Q* Rel. Bias	*DWTP/Q* MAPE	*UPP* Rel. Bias	*UPP* MAPE
Baseline	0.492	(0.5%,1.5%)	-0.194	0.290	0.268	0.141	0.872	0.534
		(4.5%,5.5%)	-0.154	0.246	0.170	0.144	0.349	0.278
		(9.5%,10.5%)	-0.148	0.209	0.148	0.138	0.038	0.165
		(14.5%,15.5%)	-0.149	0.212	0.137	0.127	-0.133	0.197
		(19.5%,20.5%)	-0.166	0.194	0.130	0.135	-0.254	0.246
(M1) Insurers Do Not Re-Optimize Premiums	0.508	(0.5%,1.5%)	-0.269	0.294	0.184	0.142	0.641	0.427
		(4.5%,5.5%)	-0.237	0.274	0.063	0.110	0.199	0.186
		(9.5%,10.5%)	-0.239	0.262	0.029	0.101	-0.054	0.172
		(14.5%,15.5%)	-0.255	0.268	-0.014	0.083	-0.196	0.225
		(19.5%,20.5%)	-0.223	0.232	0.016	0.080	-0.295	0.307
(M2) $\theta \in \{0.4, 0.7, 1.0\}$	0.525	(0.5%,1.5%)	-0.195	0.292	0.270	0.142	0.883	0.541
		(4.5%,5.5%)	-0.153	0.249	0.170	0.141	0.368	0.287
		(9.5%,10.5%)	-0.155	0.205	0.144	0.139	0.055	0.174
		(14.5%,15.5%)	-0.151	0.207	0.131	0.128	-0.113	0.169
		(19.5%,20.5%)	-0.158	0.211	0.144	0.123	-0.224	0.254
(M3) $\theta \in \{0.6, 0.9, 1.2\}$	0.462	(0.5%,1.5%)	-0.189	0.288	0.269	0.142	0.867	0.529
		(4.5%,5.5%)	-0.151	0.243	0.172	0.146	0.330	0.268
		(9.5%,10.5%)	-0.139	0.213	0.155	0.134	0.033	0.173
		(14.5 15.5%)	-0.129	0.200	0.129	0.125	-0.155	0.199
		(19.5%,20.5%)	-0.148	0.184	0.146	0.133	-0.279	0.305
(M4) $\lambda \in \{3, 6, 9\}$	0.541	(0.5%,1.5%)	-0.156	0.289	0.323	0.158	0.976	0.581
		(4.5%,5.5%)	-0.140	0.245	0.190	0.155	0.399	0.299
		(9.5%,10.5%)	-0.133	0.207	0.160	0.147	0.072	0.170
		(14.5%,15.5%)	-0.144	0.211	0.148	0.135	-0.102	0.168
		(19.5%,20.5%)	-0.144	0.192	0.149	0.133	-0.230	0.240
(M5) $\lambda \in \{1, 4, 7\}$	0.426	(0.5%,1.5%)	-0.224	0.293	0.219	0.130	0.770	0.475
		(4.5%,5.5%)	-0.171	0.247	0.147	0.135	0.291	0.266
		(9.5%,10.5%)	-0.146	0.214	0.143	0.134	0.031	0.179
		(14.5%,15.5%)	-0.145	0.211	0.135	0.122	-0.125	0.184
		(19.5%,20.5%)	-0.176	0.195	0.106	0.108	-0.263	0.283
(M6) $\theta \in \{0.6, 0.9, 1.2\}$ $\lambda \in \{3, 6, 9\}$	0.512	(0.5%,1.5%)	-0.156	0.287	0.320	0.156	0.966	0.576
		(4.5%,5.5%)	-0.140	0.244	0.190	0.156	0.376	0.292
		(9.5%,10.5%)	-0.142	0.209	0.154	0.147	0.046	0.165
		(14.5%,15.5%)	-0.141	0.217	0.153	0.140	-0.125	0.193
		(19.5%,20.5%)	-0.157	0.192	0.150	0.137	-0.241	0.252
(M7) $\#J = 8$ $\#S \in \{4, 5, 6, 7\}$	0.511	(0.5%,1.5%)	-0.182	0.369	0.428	0.188	0.977	0.541
		(4.5%,5.5%)	-0.158	0.294	0.236	0.182	0.409	0.298
		(9.5%,10.5%)	-0.182	0.270	0.197	0.171	0.087	0.177
		(14.5%,15.5%)	-0.189	0.256	0.176	0.153	-0.117	0.187
		(19.5%,20.5%)	-0.177	0.245	0.174	0.156	-0.239	0.262

Table A13: Results Summary Under Modifications to
Baseline Parameterizations and Assumptions

Modification	Mean Hosp Gr Margin	$\frac{\Delta p_r}{p_r} \in$	*WTP/Q* Rel. Bias	*WTP/Q* MAPE	*DWTP/Q* Rel. Bias	*DWTP/Q* MAPE	*UPP* Rel. Bias	*UPP* MAPE
		(0.5%,1.5%)	-0.194	0.290	0.268	0.141	0.872	0.534
		(4.5%,5.5%)	-0.154	0.246	0.170	0.144	0.349	0.278
Baseline	0.492	(9.5%,10.5%)	-0.148	0.209	0.148	0.138	0.038	0.165
		(14.5%,15.5%)	-0.149	0.212	0.137	0.127	-0.133	0.197
		(19.5%,20.5%)	-0.166	0.194	0.130	0.135	-0.254	0.246
(M8)		(0.5%,1.5%)	-0.116	0.288	0.353	0.173	0.711	0.417
500,000 Insurance		(4.5%,5.5%)	-0.092	0.237	0.247	0.224	0.226	0.208
Buying Groups	0.419	(9.5%,10.5%)	-0.065	0.209	0.250	0.234	-0.020	0.183
of Size 1		(14.5%,15.5%)	-0.034	0.178	0.285	0.255	-0.159	0.193
		(19.5%,20.5%)	-0.040	0.211	0.259	0.240	-0.299	0.317
(M9)		(0.5%,1.5%)	-0.203	0.292	0.259	0.138	0.902	0.555
5,000 Insurance		(4.5%,5.5%)	-0.165	0.246	0.145	0.132	0.358	0.290
Buying Groups	0.500	(9.5%,10.5%)	-0.168	0.221	0.122	0.125	0.056	0.180
of Size 100		(14.5%,15.5%)	-0.169	0.207	0.124	0.119	-0.138	0.187
		(19.5%,20.5%)	-0.164	0.183	0.100	0.108	-0.251	0.249
(M10)		(0.5%,1.5%)	-0.082	0.258	0.380	0.224	0.865	0.523
Random		(4.5%,5.5%)	-0.133	0.217	0.190	0.170	0.299	0.243
Travel Cost	0.487	(9.5%,10.5%)	-0.141	0.198	0.165	0.163	0.011	0.165
Parameters		(14.5%,15.5%)	-0.144	0.196	0.152	0.146	-0.136	0.193
		(19.5%,20.5%)	-0.141	0.206	0.184	0.180	-0.258	0.272
(M11)		(0.5%,1.5%)	-0.209	0.283	0.235	0.127	0.883	0.539
		(4.5%,5.5%)	-0.175	0.245	0.152	0.136	0.370	0.295
Linear	0.481	(9.5%,10.5%)	-0.178	0.231	0.132	0.126	0.078	0.162
Travel Cost		(14.5%,15.5%)	-0.165	0.216	0.128	0.134	-0.099	0.176
		(19.5%,20.5%)	-0.202	0.221	0.116	0.119	-0.236	0.259
(M12)		(0.5%,1.5%)	-0.187	0.290	0.277	0.145	0.930	0.560
		(4.5%,5.5%)	-0.155	0.248	0.169	0.144	0.372	0.294
$\rho_i = E[\rho_i]$	0.502	(9.5%,10.5%)	-0.150	0.216	0.146	0.136	0.056	0.165
$\forall i$		(14.5%,15.5%)	-0.149	0.215	0.136	0.126	-0.115	0.193
		(19.5%,20.5%)	-0.166	0.187	0.129	0.128	-0.258	0.258
(M13)		(0.5%,1.5%)	-0.209	0.293	0.247	0.135	0.831	0.510
		(4.5%,5.5%)	-0.162	0.247	0.159	0.137	0.337	0.268
$Z \in$	0.491	(9.5%,10.5%)	-0.155	0.209	0.138	0.130	0.029	0.169
{4,7,10}		(14.5%,15.5%)	-0.164	0.215	0.124	0.126	-0.140	0.200
		(19.5%,20.5%)	-0.179	0.197	0.115	0.121	-0.253	0.251
(M14)		(0.5%,1.5%)	-0.166	0.288	0.302	0.152	0.937	0.572
		(4.5%,5.5%)	-0.136	0.245	0.190	0.153	0.366	0.289
$p_z = \$5,000$	0.491	(9.5%,10.5%)	-0.138	0.206	0.158	0.142	0.054	0.169
		(14.5%,15.5%)	-0.136	0.207	0.153	0.132	-0.127	0.199
		(19.5%,20.5%)	-0.156	0.186	0.149	0.140	-0.249	0.247

Table A14: Results Summary Under Modifications to Baseline Parameterizations and Assumptions

Modification	Mean Hosp Gr Margin	$\frac{\Delta p_r}{p_r} \in$	WTP/Q Rel. Bias	WTP/Q MAPE	$DWTP/Q$ Rel. Bias	$DWTP/Q$ MAPE	UPP Rel. Bias	UPP MAPE
		(0.5%,1.5%)	-0.194	0.290	0.268	0.141	0.872	0.534
		(4.5%,5.5%)	-0.154	0.246	0.170	0.144	0.349	0.278
Baseline	0.492	(9.5%,10.5%)	-0.148	0.209	0.148	0.138	0.038	0.165
		(14.5%,15.5%)	-0.149	0.212	0.137	0.127	-0.133	0.197
		(19.5%,20.5%)	-0.166	0.194	0.130	0.135	-0.254	0.246
(M15)		(0.5%,1.5%)	-0.198	0.319	0.261	0.247	0.872	0.536
		(4.5%,5.5%)	-0.161	0.290	0.162	0.226	0.350	0.282
Prices Measured	0.492	(9.5%,10.5%)	-0.158	0.242	0.136	0.212	0.035	0.167
with Error		(14.5%,15.5%)	-0.150	0.251	0.133	0.195	-0.134	0.188
		(19.5%,20.5%)	-0.175	0.240	0.125	0.191	-0.256	0.252
(M16)		(0.5%,1.5%)	-0.025	0.233	0.467	0.291	0.872	0.539
		(4.5%,5.5%)	-0.064	0.202	0.269	0.221	0.350	0.285
Costs Measured	0.492	(9.5%,10.5%)	-0.085	0.174	0.211	0.186	0.038	0.166
with Error		(14.5%,15.5%)	-0.100	0.176	0.190	0.179	-0.133	0.196
		(19.5%,20.5%)	-0.130	0.184	0.164	0.164	-0.255	0.246
(M17)		(0.5%,1.5%)	-0.032	0.259	0.455	0.306	0.871	0.540
Prices and Costs		(4.5%,5.5%)	-0.073	0.240	0.256	0.248	0.350	0.286
Measured	0.492	(9.5%,10.5%)	-0.095	0.205	0.200	0.213	0.035	0.170
with Error		(14.5%,15.5%)	-0.096	0.213	0.193	0.199	-0.134	0.192
		(19.5%,20.5%)	-0.140	0.233	0.150	0.204	-0.257	0.253

A6 Computation

In this appendix, we provide details on our approach to solving for the Nash-in-Nash price equilibrium in our simulated hospital markets. The equilibrium consists of two broad components: (i) maximizing the set of Nash objective functions that model the bilateral bargaining between hospitals and insurers, and (ii) maximizing the profit functions of the insurers in the Bertrand games that model competition among insurers. The terms that define the equilibrium are, respectively, the prices paid by insurers to hospitals to provide inpatient care and insurance premiums.

For each simulated market, we solve these components simultaneously using a nested search algorithm. In the outer loop of the algorithm, we solve the systems of equations defined by the insurer Bertrand games by searching for optimal premiums conditional on the current guess of hospital prices. In the inner loop, we solve the system of equations defined by the first order conditions of the Nash bargaining objective functions by searching for optimal hospital prices conditional on the current guess of optimal insurance premiums. Upon convergence in the inner loop, we resolve the insurer Bertrand games (the outer loop) given the updated prices from the Nash bargaining game. We define a set of hospital prices and insurer premiums as the equilibrium if the hospital prices satisfy the first order conditions of the Nash objective functions to a given tolerance, the premiums satisfy the first order conditions of the insurer Bertrand game to a given tolerance, and the update in optimal premiums across outer loop iterations is within a given tolerance.

Before proceeding, we remind the reader of some basic notation. J denotes the set of hospitals, and S denotes the set of hospital systems. J_s denotes the set of hospitals in system s, and, in somewhat of an abuse of notation, $J\backslash s$ denotes the set of hospitals excluding system s. M denotes the set of insurers. J_n denotes the set of hospitals included in the network of insurer n. π_{J_n} is the general notation for the premium charged by insurer n when insurer n has network J_n. However, when it is clear from the context, we use π_J and $\pi_{J\backslash s}$ to denote premiums charged by a given insurer if its network consists of J or $J\backslash s$, respectively.

A6.1 Solving the Insurer Bertrand Games

In this section, we describe the search algorithm we apply in solving the insurer Bertrand games for a given vector of hospital prices. These Bertrand games model the downstream competition among insurers in selling their insurance product to consumers, and the equilibrium profits determined by these games constitute the insurer payoffs in the upstream Nash bargaining games with hospitals. As discussed in the paper, there are two categories of insurer Bertrand games. The first models insurer competition in the equilibrium outcome under which, in our setting, all hospital-insurer combinations reach an agreement. The insurer profit from this game define the insurer payoff in the Nash bargaining game denoted Π_n^J in (6). The second models insurer competition in the hypothetical outcome under which all hospital-insurer combinations reach an agreement other than insurer n and one of the hospital systems in the market. The profit for insurer n from this game define the insurer disagreement payoff in the Nash bargaining game denoted $\Pi_n^{J\setminus s}$ in (6). We refer to these hypothetical equilibria as "exclusion equilibria" since they involve the hypothetical exclusion of one of the hospital systems. Since there are $\#S$ Nash bargaining problems, we solve this hypothetical "exclusion" Bertrand game for each of the $\#S$ hospital systems in the market.

A6.1.1 Equilibrium Premium and Insurer Profits

We begin by describing our search algorithm for solving the equilibrium profit for all insurers under which all hospital-insurer combinations reach an agreement. The expected profit of insurer n if all hospital-insurer combinations reach an agreement is defined as

$$\Pi_n^J(\pi_{J_n}) \equiv \sum_g \Lambda_{gn}(\{\pi_{J_m}\}_{m\in M}) \left(\#I_g(\pi_{J_n} - p_z) - \sum_{i\in I_g} \rho_i \sum_{j\in J_n} \sigma_{ij}^{J_n}(p_{jn} + \tau) \right), \tag{A5}$$

where the probability that buying group g chooses insurer n is given as

$$\Lambda_{gn}(\{\pi_{J_m}\}_{m\in M}) \equiv \frac{\exp\left\{Z_n - \theta\pi_{J_n} + \frac{\lambda}{\#I_g}\sum_{i\in I_g}\rho_i Emax_{iJ_n}\right\}}{1 + \sum_{m\in M}\exp\left\{Z_m - \theta\pi_{J_m} + \frac{\lambda}{\#I_g}\sum_{i\in I_g}\rho_i Emax_{iJ_m}\right\}}.$$

As noted in the paper, we assume symmetric competition among insurers. This allows us to solve the equilibrium Bertrand game by solving a single equation. Taking the derivative of (A5) with respect to π_{Jn} and then applying symmetry, we have the first order condition

$$\sum_g \#I_g\Lambda_g(\pi_J) - \theta\Lambda_g(\pi_J)(1 - \Lambda_g(\pi_J)) \left(\#I_g(\pi_J - p_z) - \sum_{i\in I_g} \rho_i \sum_{j\in J} \sigma_{ij}^{J}(p_j + \tau) \right) = 0, \tag{A6}$$

where π_J denotes the premium that is common to all insurers in the symmetric equilibrium, and

$$\Lambda_g(\pi_J) \equiv \left(\#M + \exp\left\{ \theta\pi_J - Z - \frac{\lambda}{\#I_g} \sum_{i \in I_g} \rho_i Emax_{iJ} \right\} \right)^{-1}.$$

Since this is a single variable search problem, and (A6) is monotone in π_J, we solve (A6) using bisection and apply the following convergence criteria.

Convergence Criteria C1. *Letting π^R denote the right bracket in the bisection algorithm (at which A6 < 0) and π^L denote the left bracket in the bisection algorithm (at which A6 > 0), we define convergence in solving for the equilibrium insurer premium π_J^* as values of π^R and π^L such that:*

If $\pi^R - \pi^L < 10^{-10}$, then $\pi_J^* = \frac{\pi^R + \pi^L}{2}$.

The equilibrium profit for each insurer is (A5) evaluated at π_J^*. Since prices and premiums in our simulations are scaled by \$1,000, our convergence criteria solves the optional insurance premium to the nearest \$0.0000001.

A6.1.2 Exclusion Equilibrium Premium and Insurer Profits: Monopoly Insurer Case

Next, we describe our search algorithm for solving the equilibrium profit for insurer n if all hospital-insurer combinations other than insurer n and hospital system s reach an agreement. We compute this equilibrium for each hospital system in the market, and the solutions constitute the $\#S$ "exclusion equilibria". Our approach to solving these Bertrand games depends on the number of insurers in the market. If there is a single insurer, then solving for the profit maximizing premium under the hypothetical exclusion of system s is exactly analogous to solving for the equilibrium premium under symmetry. We discuss the oligopoly insurer case in the next subsection. If insurer n is a monopolist and excludes system s, its profit function is given by

$$\Pi_n^{J\setminus s}(\pi_{J\setminus s}) \equiv \sum_g \Lambda_{gn}(\pi_{J\setminus s}) \left(\#I_g(\pi_{J\setminus s} - p_z) - \sum_{i \in I_g} \rho_i \sum_{j \in J\setminus s} \sigma_{ij}^{J\setminus s}(p_{jn} + \tau) \right), \quad \text{(A7)}$$

where the probability that buying group g chooses insurer n is given as

$$\Lambda_{gn}(\pi_{J\setminus s}) \equiv \left(1 + \exp\left\{ \theta\pi_{J\setminus s} - Z - \frac{\lambda}{\#I_g} \sum_{i \in I_g} \rho_i Emax_{iJ\setminus s} \right\} \right)^{-1}.$$

Taking the derivative of (A7) with respect to $\pi_{J\backslash s}$, we have the first order condition

$$\sum_g \#I_g \Lambda_{gn}(\pi_{J\backslash s}) - \theta \Lambda_{gn}(\pi_{J\backslash s})(1-\Lambda_{gn}(\pi_{J\backslash s})) \left(\#I_g(\pi_{J\backslash s} - p_z) - \sum_{i \in I_g} \rho_i \sum_{j \in J\backslash s} \sigma_{ij}^{J\backslash s}(p_{jn} + \tau) \right) = 0. \tag{A8}$$

As with the search for the equilibrium premium, this is a single variable search problem, and the derivative of the profit function under the exclusion of s is monotone in $\pi_{J\backslash s}$. Hence, we again solve (A8) using bisection, applying the same convergence criteria.

Convergence Criteria C2. *Letting π^R denote the right bracket in the bisection algorithm (at which A8 < 0), and π^L denote the left bracket in the bisection algorithm (at which A8 > 0), we define convergence in solving for the equilibrium exclusion insurer premium $\pi^*_{J\backslash s}$ as values of π^R and π^L such that:*

If $\pi^R - \pi^L < 10^{-10}$, then $\pi^*_{J\backslash s} = \frac{\pi^R + \pi^L}{2}$.

The exclusion equilibrium profit for the monopoly insurer under the exclusion of system s is (A7) evaluated at $\pi^*_{J\backslash s}$.

A6.1.3 Exclusion Equilibrium Premium and Insurer Profits: Oligopoly Insurer Case

If there is more than one insurer, a hypothetical exclusion of a given hospital system for one of the insurers creates asymmetric competition in the insurance market since one insurer's network is different from the others. Since competition is otherwise symmetric, the first order conditions of the Bertrand game played by insurers under the hypothetical exclusion reduces to a two-by-two system of equations: one first order condition for the insurer that is excluding the hospital system (insurer n) and one first order condition for the remaining insurers ($m \in M\backslash n$), each of which includes all hospital systems.

The profit function of insurer n under the exclusion of hospital system s is

$$\Pi_n^{J\backslash s}(\pi_{J\backslash s}) \equiv \sum_g \Lambda_{gn}(\pi_{J\backslash s}, \pi_J) \left(\#I_g(\pi_{J\backslash s} - p_z) - \sum_{i \in I_g} \rho_i \sum_{j \in J\backslash s} \sigma_{ij}^{J\backslash s}(p_{jn} + \tau) \right), \tag{A9}$$

where the probability that buying group g chooses insurer n if n excludes system s and all other insurer include all systems is

$$\Lambda_{gn}(\pi_{J\backslash s}, \pi_J) \equiv \frac{\exp\left\{Z_n - \theta\pi_{J\backslash s} + \frac{\lambda}{\#I_g}\sum_{i\in I_g}\rho_i Emax_{iJ\backslash s}\right\}}{1+\exp\left\{Z_n - \theta\pi_{J\backslash s} + \frac{\lambda}{\#I_g}\sum_{i\in I_g}\rho_i Emax_{iJ\backslash s}\right\} + \sum_{m\in M\backslash n}\exp\left\{Z_m - \theta\pi_{J_m} + \frac{\lambda}{\#I_g}\sum_{i\in I_g}\rho_i Emax_{iJ_m}\right\}}.$$

The profit function of each of remaining insurers $m \in M\backslash n$, for which system s is not excluded is

$$\Pi_m^J(\pi_J) \equiv \sum_g \Lambda_{gm}(\pi_J, \pi_{J\backslash s})\left(\#I_g(\pi_J - p_z) - \sum_{i\in I_g}\rho_i \sum_{j\in J}\sigma_{ij}^J(p_{jm}+\tau)\right), \tag{A10}$$

where the probability that buying group g chooses insurer m if n excludes system s and all other insurer include all systems is

$$\Lambda_{gm}(\pi_J, \pi_{J\backslash s}) \equiv \frac{\exp\left\{Z_m - \theta\pi_{J_m} + \frac{\lambda}{\#I_g}\sum_{i\in I_g}\rho_i Emax_{iJ_m}\right\}}{1+\exp\left\{Z_n - \theta\pi_{J\backslash s} + \frac{\lambda}{\#I_g}\sum_{i\in I_g}\rho_i Emax_{iJ\backslash s}\right\} + \sum_{m'\in M\backslash n}\exp\left\{Z_{m'} - \theta\pi_{J_{m'}} + \frac{\lambda}{\#I_g}\sum_{i\in I_g}\rho_i Emax_{iJ_{m'}}\right\}}.$$

Taking the derivatives of (A9) and (A10) with respect to $\pi_{J\backslash s}$ and π_J, respectively, and applying symmetry, yields the system of first order conditions

$$\sum_g \#I_g\Lambda_{gn}(\pi_{J\backslash s}, \pi_J) - \theta\Lambda_{gn}(\pi_{J\backslash s}, \pi_J)(1-\Lambda_{gn}(\pi_{J\backslash s}, \pi_J))\left(\#I_g(\pi_{J\backslash s} - p_z) - \sum_{i\in I_g}\rho_i\sum_{j\in J\backslash s}\sigma_{ij}^{J\backslash s}(p_j+\tau)\right) = 0. \tag{A11}$$

$$\sum_g \#I_g\Lambda_g(\pi_J, \pi_{J\backslash s}) - \theta\Lambda_g(\pi_J, \pi_{J\backslash s})(1-\Lambda_g(\pi_J, \pi_{J\backslash s}))\left(\#I_g(\pi_J - p_z) - \sum_{i\in I_g}\rho_i\sum_{j\in J}\sigma_{ij}^J(p_j+\tau)\right) = 0. \tag{A12}$$

where

$$\Lambda_{gn}(\pi_{J\backslash s}, \pi_J) \equiv \frac{\exp\left\{Z - \theta\pi_{J\backslash s} + \frac{\lambda}{\#I_g}\sum_{i\in I_g}\rho_i Emax_{iJ\backslash s}\right\}}{1+\exp\left\{Z - \theta\pi_{J\backslash s} + \frac{\lambda}{\#I_g}\sum_{i\in I_g}\rho_i Emax_{iJ\backslash s}\right\} + (\#M-1)\exp\left\{Z - \theta\pi_J + \frac{\lambda}{\#I_g}\sum_{i\in I_g}\rho_i Emax_{iJ}\right\}}$$

and

$$\Lambda_g(\pi_J, \pi_{J\setminus s}) \equiv \frac{\exp\left\{Z - \theta\pi_J + \frac{\lambda}{\#I_g}\sum_{i\in I_g}\rho_i Emax_{iJ}\right\}}{1 + \exp\left\{Z - \theta\pi_{J\setminus s} + \frac{\lambda}{\#I_g}\sum_{i\in I_g}\rho_i Emax_{iJ\setminus s}\right\} + (\#M - 1)\exp\left\{Z - \theta\pi_J + \frac{\lambda}{\#I_g}\sum_{i\in I_g}\rho_i Emax_{iJ}\right\}}.$$

We solve the system given by (A11) and (A12) for $\pi_{J\setminus s}$ and π_J using Newton's method. We apply a two-component stopping rule based on Judd (1998). First, the Euclidean norm of the vector composed of (A11) and (A12) must be less than a given tolerance. Second, the the Euclidean norm of the vector composed of the updates to $\pi_{J\setminus s}$ and π_J must be less than a given tolerance. The first component verifies that the first order conditions are satisfied, and the second verifies that the sequence of guesses of the optimal premiums has converged.

Convergence Criteria C3. *Let ι index iterations in the Newton search for the optimal premiums $(\pi^*_{J\setminus s}, \pi^*_J)$ given hospital prices. We define the equilibrium as $(\pi^\iota_{J\setminus s}, \pi^\iota_J)$ if:*

(i) $\sqrt{\left(\frac{\partial \Pi_n^{J\setminus s}}{\partial \pi_{J\setminus s}}\right)^2 + \left(\frac{\partial \Pi_m^J}{\partial \pi_J}\right)^2}\Big|_{\pi^\iota_{J\setminus s}, \pi^\iota_J} < 10^{-7}$, *and*

(ii) $\sqrt{(\pi^\iota_{J\setminus s} - \pi^{\iota-1}_{J\setminus s})^2 + (\pi^\iota_J - \pi^{\iota-1}_J)^2} < 10^{-7}\left(1 + \sqrt{(\pi^\iota_{J\setminus s})^2 + (\pi^\iota_J)^2}\right)$.

The exclusion equilibrium profit for the insurer n under the exclusion of system s in the insurer oligopoly case is (A9) evaluated at $(\pi^*_{J\setminus s}, \pi^*_J)$.

We solve for the optimal premiums under a hypothetical exclusion of a given hospital system, for either the monopoly or oligopoly insurer case, for each of the $\#S$ hospital systems in the market.

A6.2 Solving the Hospital-Insurer Bargaining Game

In the subsection, we describe our approach to computing the price equilibrium in the Nash bargaining games between hospitals and insurers conditional on the current guesses of the optimal insurance premiums, $\left\{\pi^*_J, \pi^*_{J\setminus 1}, ..., \pi^*_{J\setminus \#S}\right\}$. Since there are $\#S$ hospital systems in each market, and we assume symmetric competition among insurers, the equilibrium is found maximizing (simultaneously) the joint surplus of $\#S$ Nash bargaining games. We compute the equilibrium by solving the system of $\#S$ equations defined by the derivatives of the $\#S$ Nash objective functions with respect to its own price. We solve for the set of optional prices using Newton's method.

As noted in the paper, we impose the restriction that each hospital system and insurer negotiate a single price that is applied to each hospital within the system. Hence, we described computing the equilibrium at the hospital system-insurer level.

Recall that the expected volume for system s from enrollees of insurer n is computed from three stochastic components: the probability that a consumer's insurance group g will select insurer n, the probability that each consumer in group g will require inpatient care, and the probability that each consumer in group g who does require inpatient care will select a hospital in system s. This expected volume is defined as

$$q_{sn} \equiv \sum_g \Lambda_{gn}(\pi_{J_n}) \sum_{i \in I_g} \rho_i \sum_{j \in J_s} \sigma_{ij}^J.$$

The expected volume is defined analogously across all insurers, and, of course, is equal across all insurers in equilibrium because of the assumption of symmetric competition in the insurance market. Note that hospital prices affect expected hospital volumes only indirectly through the premium in the first term, $\Lambda_{gn}(\pi_{J_n})$. The remaining two components of expected hospital volumes, ρ_i and σ_{ij}^J, are exogenous.

Similarly, the expected volume for system s from another insurer m in the event that s does not reach an agreement with insurer n is defined as

$$q_{s(m \backslash n)} \equiv \sum_g \Lambda_{gm}(\pi_J, \pi_{J \backslash s}) \sum_{i \in I_g} \rho_i \sum_{j \in J_s} \sigma_{ij}^J.$$

Next, we turn to defining the cost of providing inpatient care at the hospital system level. Recall that the exogenous cost terms c_j are drawn at the hospital level. Hence, the marginal cost of inpatient care for system s should be the expected volume weighted mean of $\{c_j\}_{j \in J_s}$. Since a component of expected volume (the insurance choice probability $\Lambda_{gn}(\pi_{J_n})$) is endogenous, the weights used to determine system-level cost c_s should be determined in equilibrium. However, almost none of the variation in expected volume across hospitals is due to $\Lambda_{gn}(\pi_{J_n})$. Rather, almost all of this variation is due to variation in the exogenous components, ρ_i and σ_{ij}^J.[43] Hence, including $\Lambda_{gn}(\pi_{J_n})$ in constructing the volume weights would unnecessarily add to the computational burden. Therefore, we use only the exogenous components of expected volume in constructing the weights. Hence, we define system-level marginal cost as

[43] To test this, we evaluate the correlation at the hospital level between the expected volume for hospital j $\sum_g \Lambda_{gn}(\pi_{J_n}) \sum_{i \in I_g} \rho_i \sigma_{ij}^J$ and the expected volume using only the exogenous components $\sum_g \sum_{i \in I_g} \rho_i \sigma_{ij}^J$. Generating these terms for 1,000 simulated markets and computing the correlation between $\sum_g \Lambda_{gn}(\pi_{J_n}) \sum_{i \in I_g} \rho_i \sigma_{ij}^J$ and $\sum_g \sum_{i \in I_g} \rho_i \sigma_{ij}^J$ in each market, we find that the correlation is never less than 0.999 and greater than 0.9999999 in 659 markets.

$$c_s = \frac{\sum_g \sum_{i \in I_g} \rho_i \sum_{j \in J_s} \sigma_{ij}^J c_j}{\sum_g \sum_{i \in I_g} \rho_i \sum_{j \in J_s} \sigma_{ij}^J},$$

and we treat this cost as fixed throughout the search algorithm.

Given these terms, the Nash bargaining objective function for hospital system s and insurer n is

$$NB_{sn} \equiv \left(q_{sn} \left(p_{sn} - c_s\right) - \sum_{m \in M \backslash n} \left(q_{s(m \backslash n)} - q_{sm}\right) \left(p_{sm} - c_s\right) \right)^{\alpha} \left(\Pi_n^J(p_{sn}) - \Pi_n^{J \backslash s} \right)^{1-\alpha}.$$

Note that we list the dependence of the insurer n's equilibrium payoff $\Pi_n^J(p_{sn})$ on the price paid to system s, but not the disagreement payoff $\Pi_n^{J \backslash s}$. This distinction arises because, under no agreement, no enrollees of n will be treated by s. However, each of these payoffs depends on the prices paid to all other systems, as shown in (A5), (A7), and (A9).

The derivative of the Nash objective function with respect to its own price is

$$\frac{\partial \ln(NB_{sn})}{\partial p_{sn}} = \alpha \frac{q_{sn} + \frac{\partial q_{sn}}{\partial \pi_{J_n}} \frac{\partial \pi_{J_n}}{\partial p_{sn}} \left(p_{sn} - c_s\right)}{q_{sn} \left(p_{sn} - c_s\right) - \sum_{m \in M \backslash n} (q_{s(m \backslash n)} - q_{sm}) \left(p_{sm} - c_s\right)} - (1-\alpha) \frac{q_{kn} + \frac{\partial \Pi_n^J(p_{sn})}{\partial \pi_n} \frac{\partial \pi_{J_n}}{\partial p_{sn}}}{\Pi_n^J(p_{sn}) - \Pi_n^{J \backslash s}}. \tag{A13}$$

The price acts indirectly through the insurance premium in both the insurer and hospital system payoffs. The indirect effect in the insurer payoff $\frac{\partial \Pi_n^J(p_{sn})}{\partial \pi_n} \frac{\partial \pi_n}{\partial p_{sn}}$ equals zero in equilibrium by the Envelope Theorem. This equilibrium condition is enforced by the outer loop of our search algorithm in which we search for the premiums that maximize insurer profits. Hence, we can ignore this term in the inner loop component of our search algorithm. However, the indirect effect of price in the hospital system payoff $\frac{\partial q_{sn}}{\partial \pi_n} \frac{\partial \pi_n}{\partial p_{sn}} \left(p_{sn} - c_s\right)$ must be accounted for. This term captures the reduction in hospital system profits from a small increase in price because of the reduction in expected volume through the decline in insurance quantity demanded.[44]

The first term in this indirect effect, which measures the reduction in expected volume due to a premium increase, is

$$\frac{\partial q_{sn}}{\partial \pi_{J_n}} \equiv \frac{\partial \sum_g \Lambda_{gn}(\pi_{J_n}) \sum_{i \in I_g} \rho_i \sum_{j \in J_s} \sigma_{ij}^J}{\partial \pi_{J_n}} = -\theta \sum_g \Lambda_{gn}(\pi_{J_n}) \left(1 - \Lambda_{gn}(\pi_{J_n})\right) \sum_{i \in I_g} \rho_i \sum_{j \in J_s} \sigma_{ij}^J$$

[44] Hence, a price increase reduces the joint surplus that is to be shared between the hospital and the insurer. Because of this term, hospitals always capture less than α percent of the joint surplus.

The second term in this indirect effect, which measures the effect of a small price increase on the equilibrium premium, is evaluated by applying the Implicit Function Theorem to the insurer's first order condition (A6). Hence,

$$\frac{\partial \pi_{J_n}}{\partial p_{sn}} = -\frac{\theta \sum_g \Lambda_{gn}(\pi_{J_n})\left(1-\Lambda_{gn}(\pi_{J_n})\right)\sum_{i\in I_g}\rho_i \sum_{j\in J_s}\sigma_{ij}^J}{-\theta \sum_g \Lambda_{gn}(\pi_{J_n})\left(1-\Lambda_{gn}(\pi_{J_n})\right)\left[2\#I_g + \theta\left(2\Lambda_{gn}(\pi_{J_n})-1\right)\left(\#I_g(\pi_{J_n}-p_z) - \sum_{j\in J_n}(p_{jn}+\tau)\sum_{i\in I_g}\rho_i\sigma_{ij}^J\right)\right]}.$$

Applying symmetry to (A13), and plugging in the expressions for the indirect effects of price, we have the following first order condition for the bargaining problem between a given insurer and hospital system s.

$$\frac{\partial \ln(NB_s)}{\partial p_s} = \alpha \frac{q_s + \frac{\partial q_s}{\partial \pi_J}\frac{\partial \pi_J}{\partial p_s}(p_s - c_s)}{\left(\#M\left(q_s - q_{s\backslash}\right) + q_{s\backslash}\right)(p_s - c_s)} - (1-\alpha)\frac{q_s}{\Pi^J(p_s) - \Pi^{J\backslash s}}, \tag{A14}$$

where $q_{s\backslash}$ denotes the expected volume for system s from each of the competing insurers if the given insurer and system s fail to reach an agreement, and

$$\frac{\partial q_s}{\partial \pi_J}\frac{\partial \pi_J}{\partial p_s} = -\frac{\theta\left(\sum_g \Lambda_g(\pi_{J_n})\left(1-\Lambda_g(\pi_J)\right)\sum_{i\in I_g}\rho_i\sum_{j\in J_s}\sigma_{ij}^J\right)^2}{\sum_g \Lambda_g(\pi_J)\left(1-\Lambda_g(\pi_J)\right)\left[2\#I_g + \theta\left(2\Lambda_g(\pi_J)-1\right)\left(\#I_g(\pi_J - p_z) - \sum_{j\in J}(p_j+\tau)\sum_{i\in I_g}\rho_i\sigma_{ij}^J\right)\right]}.$$

$\Pi^J(p_s)$ denotes the insurer's expected profit if it reaches an agreement with system s. This is defined (A5) and evaluated at the premium π_J^* as defined in (C1). Finally, $\Pi^{J\backslash s}$ denotes the insurer's expected profit if it fails to reach an agreement with system s while each of the other insurers (if any exist) do. This is defined for the monopoly and oligopoly insurer case in (A7) and (A9), respectively, and evaluated at the premium $\pi_{J\backslash s}^*$ as defined in (C2) in the monopoly insurer case or at $(\pi_{J\backslash s}^*, \pi_J^*)$ as defined in (C3) in the oligopoly insurer case.

We solve the system of equations defined by the vector of first order conditions across each of the $\#S$ Nash bargaining problems using Newton's method, applying the following convergence criteria.

Convergence Criteria C4. *Let ι index iterations in the Newton search for the optimal hospital prices $\{p_s^*\}_{s\in S}$ given insurance premiums. We define the equilibrium prices as $\{p_s^\iota\}_{s\in S}$ if:*

(i) $\sqrt{\sum_s\left(\frac{\partial \ln(NB_s)}{\partial p_s}\right)^2}\Big|_{\{p_s^\iota\}_{s\in S}} < 10^{-10}$, *and*

(ii) $\sqrt{\sum_s \left(p_s^\iota - p_s^{\iota-1}\right)^2} < 10^{-7}\left(1 + \sqrt{\sum_s \left(p_s^\iota\right)^2}\right)$.

We alternate the outer loop search (solving the insurer Bertrand games for optimal premiums given hospital prices) and the inner loop search (solving the Nash bargaining games for optimal hospital prices given premiums) until the update in optimal premiums converges across outer loop iterations. This defines our global convergence criteria to compute the equilibrium in any simulated market.

Convergence Criteria C5. *Let $\iota\iota$ index iterations in the outer loop search for optimal premiums given hospital prices $\{p_s^*\}_{s\in S}$. We define the equilibrium as a set of insurance premiums $\left\{\pi_J^{\iota\iota*}, \pi_{J\backslash 1}^{\iota\iota*}, ..., \pi_{J\backslash \#S}^{\iota\iota*}\right\}$ and hospital prices $\{p_s^{\iota\iota*}\}_{s\in S}$ if:*

(i) $\left\{\pi_J^{\iota\iota*}, \pi_{J\backslash 1}^{\iota\iota*}, ..., \pi_{J\backslash \#S}^{\iota\iota*}\right\}$ *satisfies either (C1) and (C2) or (C1) and (C3) given* $\{p_s^{\iota\iota-1*}\}_{s\in S}$*,*

(ii) $\{p_s^{\iota\iota*}\}_{s\in S}$ *satisfies (C4) given* $\left\{\pi_J^{\iota\iota*}, \pi_{J\backslash 1}^{\iota\iota*}, ..., \pi_{J\backslash \#S}^{\iota\iota*}\right\}$*, and*

(iii) $\sqrt{\left(\pi_J^{\iota\iota*} - \pi_J^{\iota\iota-1*}\right)^2 + \sum_s \left(\pi_{J\backslash s}^{\iota\iota*} - \pi_{J\backslash s}^{\iota\iota-1*}\right)^2} < 10^{-7}$.

To summarize the algorithm, we start with an initial guess of hospital system prices $\{p_s^0\}_{s\in S}$. For example, the initial guess for a given hospital system's price could be a small amount above its marginal cost. Given $\{p_s^0\}_{s\in S}$, $\left\{\pi_J^{1*}, \pi_{J\backslash 1}^{1*}, ..., \pi_{J\backslash \#S}^{1*}\right\}$ then satisfy either (C1) and (C2) in the monopoly insurer case or (C1) and (C3) in the oligopoly insurer case. $\{p_s^{1*}\}_{s\in S}$ then satisfy (C4) given $\left\{\pi_J^{1*}, \pi_{J\backslash 1}^{1*}, ..., \pi_{J\backslash \#S}^{1*}\right\}$. We repeat this process until step *(iii)* of (C5) is satisfied.

In each simulated market, we solve the equilibrium for the baseline ownership structure and for each pairwise combination of mergers between hospital systems. We repeat this for each of our 9,000 simulated markets.

A6.3 Uniqueness of the Equilibrium

We do not have a proof regarding the uniqueness of the equilibrium of our theoretical model. However, we test for the possibility of multiple equilibria by testing whether the search algorithm converges at different price vectors given different starting values. We simulate 200 hospital markets. For each market m, we solve for the price equilibrium as described in the previous section 50 times. For each replication r, we set the starting value in the search algorithm for the price of hospital system s in market m, p_{msr}, as a random draw from

$$p^o_{msr} \sim U[c_s + 1, 40],$$

where c_s denotes the marginal cost of hospital system s. Given that the expected value of c_s is 8, this constitutes a broad range of possible starting values for each hospital system price in our search algorithm.

After solving for the price equilibrium for each of the 200 markets 50 times, we take the min and max (within each market) of the set of equilibrium insurance premium $\{\pi^*_{mr}\}$ and the set of each equilibrium hospital system price $\{p^*_{msr}\}$ across replications r. With the min and max of the premium and each hospital system price, we evaluate the distance of a vector consisting of the differences between these max and min values within each market. Finally, we evaluate the max of these distances across markets. That is, we evaluate

$$\max_m \left\{ \left[\left(\max_r\{\pi^*_{mr}\} - \min_r\{\pi^*_{mr}\} \right)^2 + \sum_{s=1}^{\#S_m} \left(\max_r\{p^*_{msr}\} - \min_r\{p^*_{msr}\} \right)^2 \right]^{\frac{1}{2}} \right\}, \tag{A15}$$

where $\#S_m$ denotes the number of hospital systems in market m. The value of A15 in this exercise is approximately 2.6E-6. Based on this value and the broad range of starting values, we conclude that it is likely that the equilibrium in our theoretical model is unique.

References

Ashenfelter, O. and Hosken, D. (2010). The Effect of Mergers on Consumer Prices: Evidence from Five Mergers on the Enforcment Margin. *The Antitrust Law Journal*, 53(3):417–466.

Balan, D. J. (2017). Competitive Effects of Hospital Mergers. Working Paper.

Balan, D. J. and Brand, K. (2014). Bargaining in Hospital Merger Models. Working Paper.

Brand, K. (2013). Price Equilibrium in Empirical Models of Hospital Competition. Working Paper.

Brand, K. and Garmon, C. (2014). Hospital Merger Simulation. AHLA Member Briefing.

Capps, C., Dranove, D., and Satterthwaite, M. (2003). Competition and Market Power in Option Demand Markets. *RAND Journal of Economics*, 34(4):737–763.

Collard-Wexler, A., Gowrisankaran, G., and Lee, R. S. (2017). Nash-in-Nash Bargaining: A Microfoundation for Applied Work. Forthcoming in Journal of Political Economy.

Dafny, L., Ho, K., and Lee, R. S. (2017). The Price Effects of Cross-Market Hospital Mergers. NBER Working Paper 22106.

Farrell, J., Balan, D. J., Brand, K., and Wendling, B. W. (2011). Economics at the FTC: Hospital Mergers, Authorized Generic Drugs, and Consumer Credit Markets. *Review of Industrial Organization*, 39(4):271–296.

Fournier, G. and Gai, Y. (2007). What does Willingness-to-Pay reveal about hospital market power in merger cases? iHEA 2007 6th World Congress.

Garmon, C. (2017). The Accuracy of Hospital Merger Screening Methods. *RAND Journal of Economics*, 48(4):1068–1102.

Gaynor, M., Ho, K., and J.Town, R. (2015). The Industrial Organization of Health-Care Markets. *Journal of Economic Literature*, 53(2):235–284.

Gaynor, M. and Town, R. (2012). Competition in Health Care Markets. In Pauly, M., McGuire, T., and Barros, P., editors, *Handbook of Health Economics, Volume 2*, pages 499–637. Elsevier.

Gowrisankaran, G., Nevo, A., and Town, R. J. (2015). Mergers When Prices are Negotiated: Evidence from the Hospital Industry. *American Economic Review*, 105(1):172–203.

Haas-Wilson, D. and Garmon, C. (2009). Two Hospital Mergers on Chicago's North Shore: A Retrospective Study. FTC Bureau of Economics Working Paper No. 294.

Ho, K. and Lee, R. S. (2017). Insurer Competition in Health Care Markets. *Econometrica*, 85(2):379–417.

Judd, K. (1998). Numerical Methods in Economics. MIT Press.

Katz, M. L. (2011). Insurance, Consumer Choice, and the Equilibrium Price and Quality of Hospital Care. *The B. E. Journal of Theoretical Economics*, 2(5).

Lewis, M. S. and Pflum, K. E. (2015). Diagnosing Hospital System Bargaining Power in Managed Care Networks. *American Economic Journal: Economic Policy*, 7(1):243–274.

Lewis, M. S. and Pflum, K. E. (2017). Hospital Systems and Bargaining Power: Evidence from Out-of-Market Acquisitions. *RAND Journal of Economics*, 48(3):579–610.

May, S. and Noether, M. (2014). Predicting the Price Effects of Hospital Mergers. CRA Insights: Healthcare.

Miller, N. H., Remer, M., Ryan, C., and Sheu, G. (2016). Pass-Through and the Prediction of Merger Price Effects. *Journal of Industrial Economics*, 64(4):683–709.

Miller, N. H., Remer, M., Ryan, C., and Sheu, G. (2017). Upward Pricing Pressure as a Predictor of Merger Price Effects. *International Journal of Industrial Organization*, 52:216–247.

Peters, C. T. (2006). Evaluating the Performance of Merger Simulations: Evidence from the U.S. Airline Industry. *Journal of Law and Economics*, 47(3):627–649.

Peters, C. T. (2014). Bargaining Power and the Effects of Joint Negotiation: The Recapture Effect. Working Paper.

Ramanarayanan, S. (2014). Diversion Analysis as Applied to Hospital Mergers: A Primer. *NERA Economic Consulting*.

Health Care Cost Institute (2015). 2014 Health Care Cost and Utilization Report.

Town, R. and Vistnes, G. (2001). Hospital Competition in HMO Networks. *Journal of Health Economics*, 20(5):733–752.

Vistnes, G. and Sarafidis, Y. (2013). Cross-Market Hospital Mergers: A Holistic Approach. *Antitrust Law Journal*, 79(1):253–293.

Weinberg, M. C. (2011). More Evidence on the Performance of Merger Simualtions. *American Economic Review: Papers and Proceedings*, 101(3):151–155.

Weinberg, M. C. and Hosken, D. (2013). Evidence on the Accuracy of Merger Simulations. *Review of Economics and Statistics*, 95(5):1584–1600.