The Role of Competition and Patient Travel in Hospital Profits: Why Health Insurers Should Subsidize Patient Travel

Joseph S. Durgin
Bowdoin College, jdurgin@bowdoin.edu

Follow this and additional works at: https://digitalcommons.bowdoin.edu/honorsprojects
Part of the Econometrics Commons, Economic Theory Commons, Health Economics Commons, and the Industrial Organization Commons

Recommended Citation
https://digitalcommons.bowdoin.edu/honorsprojects/1

This Open Access Thesis is brought to you for free and open access by the Student Scholarship and Creative Work at Bowdoin Digital Commons. It has been accepted for inclusion in Honors Projects by an authorized administrator of Bowdoin Digital Commons. For more information, please contact mmcderm2@bowdoin.edu.
The Role of Competition and Patient Travel in Hospital Profits: Why Health Insurers
Should Subsidize Patient Travel

An Honors Paper for the Department of Economics

By Joseph Stanley Durgin

Bowdoin College, 2013

©2013 Joseph Stanley Durgin
The Role of Competition and Patient Travel in Hospital Profits: Why Health Insurers Should Subsidize Patient Travel

Abstract

This paper explores the effects of patient travel distance on hospital profit margins, with consideration to the effects of travel subsidies on hospital pricing. We develop a model in which hospital agglomeration leads to a negative relationship between profit margins and patient travel distance, challenging the standard IO theory that profit margins are higher for firms with greater distances of customer travel. Using data on patient visits and hospital finances from the California Office of Statewide Health Planning and Development (OSHPD), we test our theory and confirm that a hospital tends to have less pricing power if it draws patients from beyond its local cluster. We then consider how our results might justify the subsidizing of patient travel by insurers and government payers. Lastly, we present an argument for why the ubiquitous Hirschman-Herfindahl index of market concentration can be robust to owner and system-level hospital cooperation.

I. Introduction

During the last thirty years, health insurers have clearly demonstrated that influencing patient choice can be a profitable business model. While virtually non-existent in the late 1970s, the Managed Care Organization (MCO) is now the dominant form of employer-provided health insurance, with a 72 percent market share in 2012 (Kongstvedt, 2012: 10). The MCOs achieve cost savings by directing their large pools of patients toward lower-priced ‘in-network’ hospitals, encouraging providers to compete for membership in the network. By fostering this competition,
the MCOs have effectively reduced hospital prices and operating costs (Dranove 1993; Bamezai et al. 1999).

Despite the success of the MCO model on a large scale, its performance in local markets depends on hospital concentration. When a hospital has few local substitutes, an MCO may be forced to send its patients there, regardless of price, and gains little leverage from network contracting. Thus, with MCOs as the dominant means of private medical payership, it is little surprise that hospital prices are closely linked with measures of market concentration. In general, when a market has a large number of hospitals, the MCOs are more successful and medical prices are lower (Zwanziger et al. 1994; Bamezai et al. 1999).

While the existing research has noted that MCOs are more effective in less concentrated markets, it offers little prescription for how an insurer might encourage hospitals in a given market to be more competitive. That is the main focus of this paper. We consider how an insurer might use its information on patient origins to encourage hospital competition by subsidizing patient travel costs. To develop a theory of how patient travel affects hospital profits, we use the ‘circular city’ framework of Salop (1979). We first assume that travel costs take a simple linear form, but later in section II.A we relax this assumption and show the construction of a profitable subsidy for a general class of travel cost functions. The key intuition for the subsidy is that, by modifying patient travel costs, an insurer can induce a prisoner’s dilemma in hospital pricing. When this occurs, each hospital will want to reduce its price, given the prices of its neighbors, yet if we assume no cooperation, every hospital lowers its price and is less profitable in the end. This result suggests that MCOs could use their influence over patient choice to encourage hospital price competition in ways other than selective contracting.
While the model we develop may have implications for other industries, it is especially applicable to the market for healthcare, in which payers tend to be very organized. The modern health insurer is effectively a collective bargainer on the behalf of its patients, and—as demonstrated by the success of MCOs—enjoys substantial market power compared to buyers in other industries. The state and federal government payers, moreover, have even greater market influence, and could benefit from the strategies described herein if their objectives include lower hospital prices.

During our discussion of travel cost subsidies, we also consider the spatial distribution of hospitals, with a particular focus on hospital clustering. When hospitals benefit from agglomeration economies, they may choose to co-locate in a compact area, and this clustering reduces the initial cost of the subsidy and has implications for its effectiveness. As we show in section II.A, clustering can also lead to a negative relationship between patient travel distance and hospital profit margins. This result offers a caveat to the common story in industrial economics that firms with greater customer travel distance have higher profits.

Finally, a secondary goal of this paper is to examine the predominant measure of market concentration, the Hirschman-Herfindahl index (HHI), to test whether it is robust to owner-level interactions. Whether as a response to MCOs, or due to new economies of scale, the U.S. hospital industry has seen a surge of mergers and system expansions in the last two decades.\footnote{A hospital merger occurs when two hospitals come under the same ownership, operating under a common license and keeping shared financial records. A health system, on the other hand, is a collection of hospitals with the same owner but separate facilities, licenses, and financial statements (Dranove 2003: 984).}

According to the American Hospital Association, the number of community hospitals decreased from 5,384 to 4,985 between 1990 and 2010, and the number belonging to systems increased from 2,524 to 2,941 in the last decade. To calculate HHI, one must decide whether to use firm-
or owner-level market shares, and with the recent consolidation, it might seem this choice has substantial implications. However, we argue theoretically and empirically that in practical applications, the owner- and hospital-level HHI give similar results in regression models, with the former having a slightly greater effect on hospital pricing.

In section II we discuss the theory of how patient travel affects hospital profits, show the construction of a profitable patient travel subsidy, and examine the effect of hospital consolidation on HHI. Section III covers our data sources, construction of key variables, and the specification of our regression model. The empirical results appear in Section IV, and section V concludes.

II. Theory

A. Patient Travel and Hospital Profits

The existing theoretical literature has much to say about the effects of customer travel on the profits of firms. An early and influential analysis was Hotelling’s ‘linear city’ model, in which customers are evenly distributed along a line segment, and two firms compete by choosing a location and price (Hotelling 1929). The moral of the linear city is very believable—namely, that firms prefer to have distance between them (D’Aspremont et al. 1979) because it functions as a means of product differentiation.

The linear city can be adapted to hold $n > 2$ firms if we transform it into a circular city (Salop 1979), as we show in Figure 3:
The takeaway from the circular city is similar to that of the linear model. If firms are allowed to choose their locations, and if transportation costs have a quadratic form, the firms will place themselves to achieve maximum separation (Economides 1984). The intuition for this result is that when a firm is close to a group of customers, it can earn a profit by charging for the convenience.

However, if we adjust the model once more, placing the firms in clusters as shown in Figure 4, we get a different result. As described by D’Aspremont et al., the nearness of firms creates discontinuities in their profit functions, and this can lead to surprising results about the pricing strategies and existence of Nash-Cournot equilibria. In the rest of this section we formalize these results in the context of hospital markets, and show how insurers and government payers can profit by subsidizing patient travel.
To start, we assume that hospitals are placed exogenously in clusters. In reality, however, hospitals may choose to form clusters due to agglomeration economies: firms prefer to co-locate in areas with a large pooling of skilled labor, concentration of specialized suppliers, and greater interaction and knowledge spillover between workforces. When a hospital is located near a large amount of healthcare activity, these agglomeration economies provide significant cost-savings, particularly in the provision of inpatient services (Cohen & Paul 2008; Bates & Santerre 2005). Later in this section we consider how accounting for endogenous clustering might affect our results, but our initial analysis assumes exogenous placement. Throughout our work in this section we consider a cluster as a group of three hospitals, which is suitable because our model of space is effectively one-dimensional.

With an arrangement like in Figure 4, a hospital inside a cluster has a strong incentive to reduce its price below its neighbors’ prices, especially when the distance between clusters is large. Intuitively, a larger distance between clusters means that an interior hospital would gain more patient volume by capturing its neighbors’ market shares. In the following analysis, we show how this threat of price reductions places an upper bound on the equilibrium prices, which leads to an unusual negative relationship between patient travel and hospital profits.
So suppose we have \( n \) hospitals in a circular city, and each hospital \( H_i \) faces constant marginal costs \( MC_i = C \) for a patient visit. Patients in the city are evenly distributed, with density one, along the edge of the circle, remain on the circle while traveling, and always buy a hospital visit (i.e. no patient is priced out of the market). They face a linear cost of travel \( F(t) \) for a trip of distance \( t \):

\[
F(t) = bt
\]  

(2.1)

We consider what happens when the hospitals reside in clusters of three, as shown in Figure 4. The distance between two adjacent hospitals in a cluster is assumed to be uniform and denoted by \( d_1 \), while the distance between clusters is \( d_2 \). In general terms, if we hold \( d_1 \) constant while increasing \( d_2 \), we observe a greater degree of clustering in the placement of hospitals (see Figure 5 for a concrete picture of this). We want our model to explain how this increase in clustering would affect the pricing strategies of hospitals. For the reasons discussed above, we hypothesize that the interior hospitals will have a greater incentive to capture the market share between clusters, and will be more aggressive in their price competition. To explain this behavior theoretically we need to examine how patients choose a hospital and the implications for hospital pricing.

For any patient in the city, the total cost of visiting a hospital \( i \) is the cost of travelling to that hospital plus the price \( p_i \) of its services. In Figure 5, we project the circular city onto a number line and show the total cost curves for five hospitals—three in a cluster and two adjacent to the cluster—as functions of patient location. Each hospital charges the same price to its patients, but patients observe different travel costs based on their locations. For example, if a patient is located at a distance \( x \) from hospital \( H_2 \), the total cost curve for \( H_2 \) would have the
value $p_2 + bx$ at his location, which is the amount that he would pay *in total* to travel to $H_2$ and purchase a medical visit.

**Figure 5: A Cluster in the City**

![Figure 5: A Cluster in the City](image)

Each patient chooses the hospital that has the lowest cost curve at his location. As the patients are distributed with density one, the total patient volume for hospital $H_i$ is the length of horizontal space over which $H_i$ has the lowest cost curve. The hospitals compete for patient volume by adjusting their prices and shifting their cost curves up or down. To maximize profit, each hospital must weigh the benefits of raising its price against the cost of serving a smaller number of patients. Using the Nash-Cournot criterion, the system will be in equilibrium when every hospital has chosen a price which maximizes its profits, while taking as given the prices of its competitors.

To analyze the prices at a Nash-Cournot equilibrium, we first suppose that each hospital begins with a price $p_1 = C + F(d_1) + \varepsilon_1$ if it resides on the outer edge of a cluster, and $p_2 = C + F(d_1) + \varepsilon_2$ if it resides inside a cluster. By expressing the prices in this way, we can show that every hospital has a gross profit margin greater than the travel cost $bd_1$ simply by showing that $\varepsilon_i$ is positive. This will have implications for the existence of a profitable patient travel...
subsidy, as we explain below. Furthermore, it seems reasonable to assume that hospitals like $H_1$ and $H_3$, which are on the edge of a cluster, should have the same price $p_1$ at equilibrium, because they have the same products and face symmetric positioning with respect to competitors. By the same reasoning, hospitals like $H_2$, which are located within clusters, should have the same price $p_2$ at equilibrium. The approach of assuming symmetry in pricing is a common step in the literature (Tirole 1988: 283).

The profit function of an exterior hospital—take, for example, $H_1$—is the product of its gross profit margin $p_1 - C$ and total patient volume. To find $H_1$’s patient volume we compute the length of the interval over which it has the lowest total cost function. Between hospitals $H_1$ and $H_2$ the intersection of cost curves happens where a patient is indifferent between the two hospitals, i.e. at a distance $x$ where $p_1 + bx = p_2 + b(d_1 - x)$. After substituting for $p_1$ and $p_2$ this condition reduces to the following:

$$C + bd_1 + \varepsilon_1 + bx = C + bd_1 + \varepsilon_2 + b(d_1 - x)$$

or,

$$x = \frac{bd_1 + \varepsilon_2 - \varepsilon_1}{2b}$$

Likewise, between $H_1$ and $H_n$ the intersection of cost curves happens at a distance $x'$ from $H_1$, where

$$C + bd_1 + \varepsilon_2 + bx' = C + bd_1 + \varepsilon_2 + b(d_2 - x')$$

or,
\[ x' = \frac{bd_2}{2b} \tag{2.5} \]

So the total patient volume for the exterior hospital \( H_1 \) is

\[ x + x' = \frac{b(d_1 + d_2) + \varepsilon_2 - \varepsilon_1}{2b} \tag{2.6} \]

which makes the profit \( \pi_1 \) of \( H_1 \) equal to \( x + x'(p_i - c) \), or

\[ \pi_1 = \frac{b(d_1 + d_2) + \varepsilon_2 - \varepsilon_1}{2b} (bd_1 + \varepsilon_1) \tag{2.7} \]

By the same process, an \textit{interior} hospital will have a gross profit margin of \( bd_1 + \varepsilon_2 \), patient volume of \( \frac{bd_1 + \varepsilon_1 - \varepsilon_2}{b} \), and profit \( \pi_2 \) as expressed in Eq. 2.8. Note that, as one might expect, when adjacent hospitals have the same price (i.e. \( \varepsilon_1 = \varepsilon_2 \)), they split the market between them into equal parts.

\[ \pi_2 = \frac{bd_1 + \varepsilon_1 - \varepsilon_2}{b} (bd_1 + \varepsilon_2) \tag{2.8} \]

In order to have a Nash equilibrium, the prices typically must meet the following profit-maximizing conditions:

\[ \frac{\partial \pi_1}{\partial \varepsilon_1} = \frac{bd_2 + \varepsilon_2 - 2\varepsilon_1}{2b} = 0 \tag{2.9} \]

and

\[ \frac{\partial \pi_2}{\partial \varepsilon_2} = \frac{\varepsilon_1 - 2\varepsilon_2}{2b} = 0 \tag{2.10} \]
However, the unique values \( \varepsilon_1^* = \frac{2bd_2}{3} \) and \( \varepsilon_2^* = \frac{bd_2}{3} \) that satisfy these conditions do not necessarily yield the final equilibrium prices. This is because the profit functions of the hospitals are not everywhere differentiable (as in d’Aspremont et al. 1979). Whenever \( \varepsilon_1 \) is positive, hospital \( H_2 \) can potentially earn a profit by dropping its price to \( C + \varepsilon_1 \), thereby capturing the entire patient volume of \( H_1 \) and \( H_3 \) as shown in Figure 6. The profit function of \( H_2 \) is generally discontinuous when the cost curves intersect exactly at the exterior hospitals, because \( H_2 \) can see a large jump in its profit by dropping its price marginally and taking the entire market of its cluster. By comparing the profits of \( H_2 \) before and after the price drop, we can derive a ‘clustering constraint’—i.e., a sufficient condition for when \( \varepsilon_1^* \) and \( \varepsilon_2^* \) do not yield the equilibrium prices.

Figure 6: Hospital \( H_2 \) Reduces Its Price

Before \( H_2 \) drops its price to \( C + \varepsilon_1 \), it has profit \( \pi_2 \) as expressed in Eq. 2.7. After the price decrease, it will have profit \( \pi_2' \):

\[
\pi'_2 = (2d_1 + d_2)(\varepsilon_1)
\]

(2.11)
To calculate this new profit, we find the interior hospital’s new patient volume and multiply it by the new profit margin of $\varepsilon_1$. The new patient volume is clearly $2d_1 + d_2$ because $H_2$ captures the entirety of its cluster ($2d_1$) while also getting half of the market on either side of the cluster ($d_2$). We can also compute patient volume more formally as we do in Eq. 2.2, and the result is the same.

The interior hospital will drop its price, attempting to capture the entire market of its cluster, if and only if $\pi'_2 \geq \pi_2$:

$$\frac{bd_1 + \varepsilon_1 - \varepsilon_2}{b}(bd_1 + \varepsilon_2) \leq (2d_1 + d_2)(\varepsilon_1)$$

(2.12)

or,

$$\frac{b^2d_1^2 - \varepsilon_2^2}{bd_1 + bd_2 - \varepsilon_2} \leq \varepsilon_1$$

(2.13)

By substituting $\varepsilon_1^*$ for $\varepsilon_1$ and $\varepsilon_2^*$ for $\varepsilon_2$ in Eq. 2.13, we can show that, when the prices satisfy the traditional first-order conditions (Eq. 2.9 and 2.10), any interior hospital will try to capture its local market if the distances $d_1$ and $d_2$ are such that

$$d_2 \geq \frac{3d_1(\sqrt{6} - 1)}{5}$$

(2.14)

Intuitively, if the clusters are compact enough, the interior hospitals are more aggressive in their price competition, since they have more to gain by capturing the markets outside of their clusters. When Eq. 2.14 is satisfied, then the Nash prices $p_1 = C + bd_1 + \varepsilon_1^*$ and $p_2 = C + bd_1 + \varepsilon_2^*$ are not the equilibrium prices. Note that the constraint in 2.14 is actually quite loose—it is satisfied, for instance, if $d_1 = d_2$, i.e. when the hospitals are evenly distributed. In other
words, even when clustering is low in intensity, the threat of price reductions by interior hospitals begins to affect the steady state. Our expectation is that, when the degree of clustering increases, this effect will become more substantial.

Of course, this raises the question: what are the equilibrium prices when the hospitals are clustered? It turns out that when Eq. 2.14 holds, the prices of the interior and the exterior hospitals must be lower than the Nash equilibrium prices satisfying Eq. 2.9 and 2.10. To show this, we start by making several observations about the profit maximizing prices for each hospital. Importantly, we do not assume that there is one unique equilibrium, but rather we derive conditions that apply to all possible equilibria. This approach is helpful because there happen to be several steady states, and our conclusions apply to all of them.

First, the minimum value for both $\varepsilon_1$ and $\varepsilon_2$ is $-bd_1$, since any smaller value would have the hospitals pricing below marginal cost. However, any set of prices where $\varepsilon_1 \leq 0$ and $\varepsilon_2 \leq 0$ cannot yield the equilibrium, because if $\varepsilon_1$ and $\varepsilon_2$ are negative we have $\frac{\partial \pi_1}{\partial \varepsilon_1} > 0$, which would make every exterior hospital increase its price. Second, we can tell from the first order conditions in Eq. 2.9 and 2.10 that the equilibrium prices must be such that $\varepsilon_1 \geq 0$ and $\varepsilon_2 \geq 0$, for if $\varepsilon_1 \geq 0$ and $\varepsilon_2 \leq 0$, then $\frac{\partial \pi_2}{\partial \varepsilon_2} > 0$, and if $\varepsilon_1 \leq 0$ and $\varepsilon_2 \geq 0$, then $\frac{\partial \pi_1}{\partial \varepsilon_1} > 0$, so any hospital that has a negative $\varepsilon_i$ would earn a profit by charging a higher price.

Finally, the presence of clustering effectively places a ceiling over $\varepsilon_1$ and $\varepsilon_2$, which binds if the above inequality is satisfied. When the system is in equilibrium, every hospital has chosen a price that maximizes its profit given the prices of its competitors. Hence, at any equilibrium we have $\frac{\partial \pi_1}{\partial \varepsilon_1} \geq 0$ and $\frac{\partial \pi_2}{\partial \varepsilon_2} \geq 0$, because if either derivative were negative, the exterior or interior
hospitals, respectively, would earn a profit by charging lower prices. These inequalities can hold in a non-strict form (i.e. $\frac{\partial \pi_i}{\partial \varepsilon_1} > 0$) if and only if the clustering ceiling prevents the hospitals from increasing $\varepsilon_i$.

In short, there are three possibilities for system equilibria when the clustering inequality is true. Either the ceiling binds for $\varepsilon_1$ alone ($\frac{\partial \pi_1}{\partial \varepsilon_1} > 0$ and $\frac{\partial \pi_2}{\partial \varepsilon_2} = 0$), for $\varepsilon_2$ alone ($\frac{\partial \pi_1}{\partial \varepsilon_1} = 0$ and $\frac{\partial \pi_2}{\partial \varepsilon_2} > 0$), or for both together ($\frac{\partial \pi_1}{\partial \varepsilon_1} > 0$ and $\frac{\partial \pi_2}{\partial \varepsilon_2} > 0$). Notably, whenever one profit derivative is strictly positive ($\frac{\partial \pi_i}{\partial \varepsilon_i} > 0$), it is not a traditional Nash equilibrium. But it is still a stable outcome, because hospital $i$ finds it profitable to keep raising its price until it hits the ceiling, but recognizes that a higher price would lead to zero profits. In fact, it is possible for a given system to have up to three such equilibria, with the final outcome depending on $d_1, d_2$ and the beginning values of $\varepsilon_1$ and $\varepsilon_2$. However, this potential for multiple equilibria does not prevent us from drawing specific conclusions about the relationship between patient travel distance and hospital profits. In particular, we can derive upper bounds for the values of $\varepsilon_1$ and $\varepsilon_2$, and use these to show that a negative relationship exists between $d_2$ and hospital prices. To start, recall from Eq. 2.13 that, for a given $\varepsilon_1$ and $\varepsilon_2$, the interior hospitals will drop their prices to $C + \varepsilon_1$ and capture their neighbor’s market shares if and only if

$$\frac{b^2d_1^2 - \varepsilon_2^2}{bd_1 + bd_2 - \varepsilon_2} \leq \varepsilon_1$$  \hspace{1cm} (2.15)
If this occurs, the exterior hospitals have zero output and profit. Since the exterior hospitals would never be satisfied with such an outcome, any equilibrium must be characterized by

$$
\varepsilon_1 < \frac{b^2 d_1^2 - \varepsilon_2^2}{bd_1 + bd_2 - \varepsilon_2}
$$ (2.16)

But we know from the discussion above that $\frac{\partial \pi_1}{\partial \varepsilon_1} \geq 0$ at any equilibrium, so $0 \leq \varepsilon_2 \leq \frac{bd_2}{3}$ and we have

$$
\varepsilon_1 < \frac{b^2 d_1^2 - \varepsilon_2^2}{bd_1 + bd_2 - \varepsilon_2} < \frac{b^2 d_1^2}{bd_1 + \varepsilon_2} < \frac{b^2 d_1^2}{bd_1 + \frac{2bd_2}{3}}
$$ (2.17)

And we also know that $\frac{\partial \pi_2}{\partial \varepsilon_2} \geq 0$ at equilibrium, so $\varepsilon_2 \leq \frac{\varepsilon_1}{2}$. Hence,

$$
\varepsilon_2 < \frac{b^2 d_1^2}{2bd_1 + \frac{4bd_2}{3}}
$$ (2.18)

Notice that these upper bounds for $\varepsilon_1$ and $\varepsilon_2$ in Eq. 2.17 and 2.18 are actually quite small in dollar terms. In our data from the OSHPD, the average distance between a hospital and its closest neighbor is 5.01 miles, so in practical applications the value of $bd_1$ is likely small compared to the price of a patient visit. Both of the upper bounds are less than $bd_1$—and they can be substantially less when $d_2$ is large compared to $d_1$.

Moreover, when average patient travel distance increases through a rise in $d_2$, the upper bounds in Eq. 2.17 and 2.18 decrease. This squeezes $\varepsilon_1$ and $\varepsilon_2$ toward zero, leading to smaller
hospital gross profit margins. Hence, both interior and exterior hospitals can face smaller profit margins when the average travel distance of their patients increases. Intuitively, the higher $d_2$ makes it more profitable for each interior hospital to try to capture its neighbors’ patient volume (see Eq. 2.13), and as a result the exterior hospitals choose lower, more cautious prices to protect themselves from this threat.

In addition to the negative relationship between patient travel distance and hospital profit margins, another interesting feature of this model is that both $\varepsilon_1$ and $\varepsilon_2$ are strictly positive. By our initial equations for the prices, the positive $\varepsilon_1$ and $\varepsilon_2$ implies that gross profit margins in this system are always at least $bd_1$. Hence, if an insurer or government payer were to subsidize patient travel between any two adjacent hospitals in a cluster, the cost of medical services would fall by more than the amount of the subsidy. The insurer would pay at most $bd_1$ per patient visit, which would cover the expense of travelling between two adjacent hospitals. Patients would then view each hospital as being identical to its nearest neighbor, so the hospitals would engage in Bertrand price competition and price at marginal cost.\(^2\) The entire gross profit margins of $bd_1 + \varepsilon_i$ would dissipate, and the institutional payers would save more on lower medical prices than they spent on issuing the patient travel subsidies.

This model has a number of assumptions which may draw objection. In particular, we assume that travel costs have a very simple linear form, hospitals are placed exogenously in clusters, and every patient buys hospital services (i.e. the market is ‘covered’—no patient is priced out). However, while we do make some strong claims for concreteness and tractability, it turns out that we can show that a profitable patient travel subsidy exists in the context of

---

\(^2\) Interestingly, Hotelling used travel costs to explain why Bertrand competition was not commonly observed ‘in the wild.’ He called this the ‘stability of competition’ (Hotelling 1929). By his reasoning, it makes sense that we could shift the outcome closer to Bertrand competition by allaying travel costs.
agglomeration economies and a very general class of travel cost functions. In particular, as long as interior hospitals have larger gross profit margins than exterior hospitals, and as long as patient travel costs $F(t)$ are monotone-increasing, concave, and such that $F(0) = 0$, we can derive a profitable subsidy. We now briefly cover this expanded model. First, we discuss the implications of agglomeration economies for the cost structure of hospitals. Then, we examine the conditions under which a hospital will decide to reduce its price, which will inform our conditions on $F(t)$ and construction of a profitable subsidy.

Suppose that the production of medical services benefits from agglomeration economies, like the pooling of skilled labor etc. described above. Then, if interior and exterior hospitals have marginal costs of $C_2$ and $C_1$, respectively, we should have $C_2 < C_1$, because interior hospitals are situated near a greater amount of production. These cost benefits for interior hospitals must be sufficient to make their choice of an interior location at least as profitable as an exterior location. Hence, since the markets inside the clusters are smaller than those without, each interior hospital should have a larger gross margin on each unit of service than its exterior neighbors. In short, agglomeration economies imply that interior hospitals have lower marginal costs than exterior hospitals. Moreover, in any model of endogenous clustering, we should expect interior hospitals to have higher gross profit margins. Our revised, expanded model should be robust to both of these variations on the original model.

Now we turn to travel costs and their implications for the construction of a profitable subsidy. If travel costs have a general form $F(t)$, then every interior hospital will produce an amount $2x$ where

$$ C_2 + F(d_1) + F(x) + \epsilon_2 = C_1 + F(d_1) + F(d_1 - x) + \epsilon_1 \tag{2.19} $$
From this equality, we can observe that if an interior hospital chooses to increase its price by \(d\varepsilon_2\), its production quantity will change by \(2dx\), where

\[
\frac{dx}{d\varepsilon_2} = \frac{1}{F'(d_1 - x) - F'(x)} \quad (2.20)
\]

The profits for an interior hospital are equal to its gross margin \(F(d_1) + \varepsilon_2\) multiplied by its total patient volume \(2x.\)

\[
\pi_2 = 2x(F(d_1) + \varepsilon_2) \quad (2.21)
\]

Likewise, the profits for an exterior hospital are the product of its patient volume \(\frac{d_2}{2} + (d_1 - x)\) and its gross margin \(F(d_1) + \varepsilon_1\):

\[
\pi_1 = \left(\frac{d_2}{2} + (d_1 - x)\right)(F(d_1) + \varepsilon_1) \quad (2.22)
\]

By differentiating these profit functions with respect to \(\varepsilon_2\) and \(\varepsilon_1\), respectively, we find the following first order derivatives:

\[
\frac{d\pi_2}{d\varepsilon_2} = 2(F(d_1) + \varepsilon_2)\frac{dx}{d\varepsilon_2} + 2x = \frac{2(F(d_1) + \varepsilon_2)}{F'(d_1 - x) - F'(x)} + 2x \quad (2.23)
\]

and

\[
\frac{d\pi_1}{d\varepsilon_1} = \frac{(F(d_1) + \varepsilon_1)}{F'(d_1 - x) - F'(x)} + \frac{d_2}{2} + (d_1 - x) \quad (2.24)
\]

\[\text{Note that the } dx \text{ and } x \text{ terms are doubled to reflect the symmetric markets around the interior hospital.}\]
We are interested in finding the conditions such that \( \frac{d\pi_1}{dx_1} < 0 \) and \( \frac{d\pi_2}{dx_2} < 0 \), i.e. under which it is profitable on the margin for each hospital to lower its price. Finding these conditions will allow us to construct a subsidy that encourages hospital price reductions. From the expression in Eq. 2.23, we have that \( \frac{d\pi_2}{dx_2} < 0 \) if and only if

\[
\frac{(F(d_1) + \varepsilon_2)}{-x} < (F'(d_1 - x) - F'(x)) \tag{2.25}
\]

But if \( d_2 > d_1 \), we have that \( 2d_2 > x > 0 \), so Eq. 2.25 holds if

\[
\frac{(F(d_1) + \varepsilon_2)}{-2d_2} \leq (F'(d_1 - x) - F'(x)) \tag{2.26}
\]

Now, recall our conditions on \( F(t) \). Namely, we have that \( F(0) = 0 \), \( F'(x) > 0 \) for all \( x \), and \( F''(x) \leq 0 \). In addition, we also have that \( F(d_1) + \varepsilon_2 \geq 0 \), since gross margins must be non-negative. Hence, there exists some non-negative \( \gamma^* \) such that \( \gamma^* \leq F(d_1) \) and \( \gamma^* \leq \frac{(F(d_1) + \varepsilon_2)}{2} \), and such that \( F(s) = \gamma^* \) for some \( s \in (0, d_1) \). In our analysis, this value \( \gamma^* \) will be the maximum cost of the subsidy per patient. Therefore, since \( 0 \leq \gamma^* \leq F(d_1) + \varepsilon_2 \), the condition in Eq. 2.26 is satisfied if

\[
\frac{\gamma^*}{-2d_2} \leq (F'(d_1 - x) - F'(x)) \tag{2.27}
\]

or,

\[
\frac{s\gamma^*}{-2s^{2d_2}} \leq (F'(d_1 - x) - F'(x)) \tag{2.28}
\]
By the mean value theorem, there exists some $\xi_2 \in (0, s)$ such that $F' (\xi_2) = \frac{y^* - F(0)}{s}$. But since $F(0) = 0$, we have $F'(\xi_2) = \frac{y^*}{s}$. So we can rewrite the price-reduction condition in Eq. 2.28 as

\[
\frac{sF'(\xi_2)}{-2d_2} \leq (F'(d_1 - x) - F'(x))
\] (2.29)

However, we assumed that $F'(x) > 0$ and $F''(x) \leq 0$. Hence, we have $0 \leq F'(d_1) \leq F'(\xi_2)$, which implies that Eq. 2.29 is satisfied if

\[
\frac{sF'(d_1)}{-2d_2} \leq (F'(d_1 - x) - F'(x))
\] (2.30)

And by a similar argument, we can show that the price-reduction condition for exterior hospitals is of the form

\[
\frac{(F(d_1) + \varepsilon_1)}{-d_2 - (d_1 - x)} \leq (F'(d_1 - x) - F'(x))
\] (2.31)

which if $d_2 > d_1$ is satisfied when

\[
\frac{rF'(d_1)}{-2d_2} \leq (F'(d_1 - x) - F'(x))
\] (2.32)

for some $r \in (0, d_1)$.

To recap, when Eq. 2.30 and 2.32 are satisfied, we have $\frac{d\pi_1}{d\varepsilon_1} < 0$ and $\frac{d\pi_2}{d\varepsilon_2} < 0$, and each hospital wants to reduce its price. An important thing to notice about Eq. 2.30 and 2.32 is that both sides of the price-reduction inequalities are negative. When the inequalities do not hold, the
reason is either that $F'(d_1 - x)$ is too negative or $F'(x)$ is too positive. Hence, the idea of our subsidy is to twist the slopes of $F'(d_1 - x)$ and $F'(x)$ in such a way as to make the inequalities in 2.30 and 2.32 bind. This makes $\frac{d\pi_1}{dx} < 0$ and $\frac{d\pi_2}{dx} < 0$, effectively inducing a prisoner’s dilemma in the hospitals’ price competition. Given the prices of their competitors, each interior and exterior hospital wants to decrease its price. But when two adjacent hospitals *do* reduce their prices, neither has really profited, yet the insurer or government payer benefits from the lower prices. The precise construction of the subsidy is illustrated in Figure 7 and described in detail below.

**Figure 7: Construction of the Subsidy**

In Figure 7 the hospital $H_2$ is inside a cluster, while $H_1$ is on the edge of a cluster. The original intersection of their two cost curves is labeled $c$. To construct the subsidy, we start by choosing a point $a$ to the right of $c$, such that the vertical line segment $\overline{ab}$ has length $ab \leq \gamma^*$. We choose $e$ to be directly below $c$ and level with the point $b$, i.e. such that the slope of $\overline{eb}$ is
zero. The last significant point that we choose is $d$, which is between $c$ and $e$ and placed where the slope $m_{\overline{d}b} = \max \left( -\frac{rF'(d_1)}{zd_2}, -\frac{sF'(d_1)}{zd_2} \right)$. The point $d$ must exist because $F(t)$ is concave, which makes $m_{\overline{d}b}$ less negative than any tangent slope of $F(t)$ on the interval $(0, d_1)$, and in particular, any $F'(x)$ for $x \in (e, b)$.

With these key points so defined, we can proceed with the construction. Concretely, the subsidy is an ex-post payment that reimburses each patient for his travel expenses, altering the observed total cost functions. For the exterior hospital, the subsidy takes $H_1$’s cost curve between the origin and $a$ and shifts it down to the new curve $\overline{fe} \cup \overline{eb}$. A patient between $H_1$ and $f$ pays no travel costs. For the interior hospital, the subsidy shifts the cost curve between $c$ and $a$ down to $\overline{db}$. We show the new, subsidized cost curves in Figure 8.

**Figure 8: Subsidized Total Cost Curves**

By our construction of the key points, the subsidy costs at most $ab \leq y^*$ for each patient visit. Before any hospital has reacted to the subsidy, a patient located exactly at point $b$ receives this maximum payment $ab$, and patients at other locations receive a smaller amount or no
subsidy at all. Moreover, as Figure 8 shows, the subsidy has the immediate effect that patients in the region $eB$ find that hospital $H_1$ is suddenly cheaper to visit than hospital $H_2$. Therefore, hospital $H_2$ loses the middle-ground market share of $eB$. The key intuition for the subsidy is that $H_2$ will react to this loss by attempting to take back some of the now-contested region, reducing its price and capturing some of $H_1$’s market in the process. Using our price-reduction conditions (Eq. 2.30 and 2.32), we now argue that the subsidy will encourage a round of price reductions that more than makes up for its expense.

First, note that when $H_2$ is considering whether to lower its price at the new intersection $b$, it observes the following:

$$F'(d_1 - x) - F'(x) = 0 + m_{eB}$$  \hspace{1cm} (2.33)

But recall that we chose $m_{eB}$ so that

$$m_{eB} = \max\left(-\frac{sF'(d_1)}{2d_2}, -\frac{rF'(d_1)}{2d_2}\right)$$  \hspace{1cm} (2.34)

Hence, combining Eq. 2.33 and 2.34,

$$F'(d_1 - x) - F'(x) \geq -\frac{sF'(d_1)}{2d_2}$$  \hspace{1cm} (2.35)

So the price-reduction condition in Eq. 2.30 is satisfied, implying that $\frac{d\pi_2}{d\varepsilon_2} < 0$. Thus, the interior hospital wants to reduce its price.\(^4\) But notice that as soon as $H_2$ does reduce its price,

\(^{4}\) Note that it is possible to construct a degenerate case where the right-hand derivative $\left(\frac{d\pi_2}{d\varepsilon_2}\right) > 0$. However, this does not affect the main result. Due to the continuity of $F'(x)$ it is possible to constrain $\gamma^*$ sufficiently so that $\left(\frac{d\pi_2}{d\varepsilon_2}\right) < 0$, in which case the interior hospital will unambiguously decrease its price.
taking some market share back from $H_1$, it immediately become profitable for $H_1$ to reduce its own price. To see this, note that when $H_1$ is considering a price reduction, it observes $F'(d_1 - x)$ and $F'(x)$ as in Eq. 2.33. Hence, since we chose $m_{dB}$ to satisfy the expression in 2.34, we have the following:

$$F'(d_1 - x) - F'(x) \geq \frac{-rF'(d_1)}{2d_2}$$ \hspace{1cm} (2.36)

Therefore, after $H_2$ lowers its price in response to the original shock of losing patient volume, $H_1$ will reduce its own price. As each hospital lower its price, its markup becomes smaller, making the left hand sides of Eq. 2.26 and 2.31 less negative. This tâtonnement of price reductions ends when the exterior hospital $H_1$ has reduced its original gross margin by at least half, after which the condition in Eq. 2.27 is no longer guaranteed to bind. Hence, the new price of $H_1$ is no greater than $C_1 + \frac{F(d_1) + \varepsilon_1}{2}$. Moreover, throughout this process, every price decrease by $H_1$ was accompanied by at least an equivalent reduction by $H_2$, so the interior hospital $H_2$ must have lowered its margins even more than $\frac{F(d_4) + \varepsilon_3}{2}$. Hence, both hospitals reduce their prices by more than $\gamma^*$, and hence more than the maximum amount $ab < \gamma^*$ of the subsidy. The final pricing configuration appears in Figure 9:
Given the construction so far, we have shown that both hospitals reduce their prices by more than the cost of the subsidy per hospital visit. However, to verify that the subsidy is truly profitable for insurers, we also need to show that it does not shift patient volume toward the higher-priced hospital. To see that our subsidy fits this criteria, note that when the price reductions finally cease, the configuration should be like that in Figure 9, where the intersection of cost curves is at point \(d = e\) on \(H_1\)’s curve. The interior hospital \(H_2\) will reduce its price to this point because Eq. 2.35 is satisfied whenever the intersection of cost curves occurs on the segment \(eb\). Due to its higher original gross profit margins, the interior hospital has enough pricing flexibility to lower its price to this point. Once the interior hospital has lowered its price until \(d = e\), the interior hospital will not want to reduce its price any further.\(^5\) Therefore, the distribution of patient volume between hospitals does not change from its original, pre-subsidy

\(^5\) The interior hospital will not reduce its price when \(d = e\) for two reasons: 1) the marginal cost of doing so is strictly negative, since the hospital gains no market share, and 2) if it did lower its price past the discontinuity, the intersection \(c = e\) would mirror the original intersection of the cost curves. We assumed that both hospitals had no desire to reduce their prices at the original intersection, and after the round of subsidy-induced price reductions this will still be true, because reducing the gross margins of a hospital increases its \(\frac{d\pi_i}{d\ell_i}\) at every possible intersection (see Eq. 2.23 and 2.24).
state, and we need not be concerned with shifting patients to the higher-priced hospital. The subsidy we have constructed is therefore profitable.

Thus, in the context of hospital clustering, and for a very general class of function $F(t)$, there exists a profitable patient travel subsidy. From our discussion, it may seem that hospital agglomeration is invoked only sparingly, and may not be important to the main result. However, we required the agglomeration economies to justify our claim that the interior hospitals have higher gross margins than their exterior neighbors. This claim ensures that the subsidy does not risk shifting patient volume toward a higher-priced hospital. In addition, clustering has important implications for the real-world application of the subsidy, because it places a cap on the gross margins of hospitals and reduces the cost of traveling between adjacent hospitals. This has the effect of reducing the cost of the subsidy, which is important because its beneficial effects would occur after some delay.

Throughout our construction of the subsidy, the concavity of $F(t)$ was critically important. The assumption that $F''(x) \leq 0$ ensures that the point $d$ exists, and allows us to find an upper bound for the left-hand side of the price-reduction conditions (Eq. 2.25 and 2.31). In the next section, we present a short empirical argument for why the cost of travel function is concave in the real world.

B. The Marginal Costs of Patient Travel

During the construction of the subsidy in the preceding section, we assumed that travel costs were concave. With circular city-type models, however, it is often more common to use convex travel costs, because they make it easier to show the existence of a unique Nash-Cournot
equilibrium (Tirole 1988: 280). 6 But even though convex travel costs are more common in spatial pricing models, they are not necessarily realistic. That travel costs are concave, like those in our model, is frequently assumed in research on transportation network engineering (Florian 1986; Thomas & Griffin 1996; Yan & Luo 1999), and several economic studies have also used concave cost functions (Stahl 1982; Frutos, Hamoudi, & Jarque, 2001; Hamoudi & Moral 2005).7 Due to the lack of consensus in the literature, in this section we provide empirical support that real-world travel costs take a concave form.

First, suppose we have n patients, and each has the following marginal benefits and non-travel costs of a hospital visit:

\[ MB_j(Q_j) = a - bQ_j \]  

and

\[ MC_{jNT}(Q_j) = c + dQ_j \]  

We wish to examine how travel distance affects the perceived cost of care. Our approach in this analysis is to show that concave travel costs lead to a convex plot of discharges vs. patient travel distance. Then we plot the data from our sample and show that it supports our claim that travel costs are concave.

---

6 In a 2008 paper, Agoudas and Hamoudi noted the absence of non-convex travel costs in spatial pricing theory. They wrote, “Surprisingly, the literature on product differentiation has not focused on this feature, that is, the fact that transportation costs are concave in distance. There, it is assumed that transportation costs are convex and, as a result, demands for firms are connected” (Agoudas & Hamoudi 2008: 93).

7 Hamoudi and Moral point out that the cost of a plane flight is generally concave with respect to distance.
We use the familiar notation of $F(t)$ for patient travel costs, where $t$ is travel distance. For concreteness, we let $F(t)$ have a quadratic form with respect to $t$, but we make no assumptions about the parameters of this function.

$$F(t) = e + ft + gt^2$$  \hspace{1cm} (2.39)

If we want to adjust $MC_i^{NT}$ to account for travel costs, we can simply add $F(t)$ to the non-travel costs. The result is the total marginal costs for patient $j$ of visiting the hospital for a unit of service:

$$MC_j(Q_j) = c + dQ_j + e + ft + gt^2$$  \hspace{1cm} (2.40)

The aggregate marginal benefit and cost functions are now easily found by horizontally summing the individual functions Eq. 2.37 and 2.40 across all patients:

$$MB(Q_{total}) = a - \frac{bQ_{total}}{n}$$  \hspace{1cm} (2.41)

$$MC(Q_{total}) = c + e + ft + gt^2 + \frac{dQ_{total}}{n}$$  \hspace{1cm} (2.42)

So individual patients will decide to consume patient visits until $MB = MC$, at which point the total medical usage is

$$Q_{total} = \frac{n(a - c - F(t))}{b + d}$$  \hspace{1cm} (2.43)
Thus, the patient discharges vs. travel distance plot will be concave if $F(t)$ is convex ($g>0$), convex if $F(t)$ is concave ($g<0$), and linear otherwise ($g=0$). These relationships are illustrated in Figure 10.

![Figure 10. Patient Discharges by Travel Distance](image)

While we used a quadratic form to show this relationship explicitly, it holds for more general functions too. We can see from the expression for $Q_{total}$ that as long as $F(t)$ is concave, the plot of patient discharges vs. travel distance should be convex.

The actual patient discharges vs. travel distance plot is illustrated in Figure 11. The plot is clearly convex, supporting our assumption that $F(t)$ is concave. In our sample the number of patients who traveled between ten and twenty miles is roughly 20.2% of the number who traveled between zero and ten. But the number who traveled between 150 and 160 miles is 91.5% of that which traveled between 140 and 150 miles. While one could argue that the plot’s shape is due to population density effects, the graph is still convex when we restrict the sample to patients from a small area. This supports our hypothesis that patients view travel as having diminishing marginal costs.
C. Measuring Hospital Competition

The degree of price competition between hospitals is important to our theoretical results because it affects hospital margins and, therefore, the construction and cost of patient travel subsidies. In both regulatory and academic contexts, the predominant measure of intra-industry competition is the Hirschman-Herfindahl index, which is found for a given market by summing the squared market shares of the competing firms. Of course, the precise value of the index depends to a large degree on where we draw the bounds of the market—be it at the city, county, state, country or other level—and how we define the firms. The flavors of HHI that appear in the literature on hospital pricing and competition can be thought to fall into two categories: hospital-level and owner-level. In this section we describe the two types, the common rationales for choosing between them, and then we present an argument for why that choice may not be so consequential—i.e. for why the two types of HHI may be effectively interchangeable in regression models.
The hospital-level approach to computing HHI is widely used in the literature (Dranove 1993; Link 1995; Bamezai 1999; Krishnan 2001). It involves treating every hospital in the market as a separate firm in the calculation of market shares. The owner-level measure, on the other hand, treats each hospital owner as a firm. This distinction is important because hospitals are increasingly consolidating into health systems and private corporations, and these large owners tend to negotiate collectively on behalf of their hospitals, even when the hospital facilities and licenses remain separate (Kongstvedt, 2012: 11). One rational for using the owner-level HHI is that it is more economically relevant—it provides a clearer picture of hospital pricing power because it accounts for the fact that co-owned hospitals pool their strength in negotiations. But as we show below, the relationship between owner- and hospital-level HHI is likely to be of a special linear form that makes the measures statistically similar.

Before we directly compare owner- and hospital-level HHIs, we first describe the usual specification for HHI in regression equations. The most common approach is to include HHI on the right-hand side of the regression model with a simple linear form. The theory behind this form is perhaps most commonly rooted in Cournot competition, where each seller has market power and maximizes its profit \( \pi_i = P(Q)q_i \). The first order condition of this profit maximization problem is as follows:

\[
0 = P(Q) + P'(Q)q_i - MC_i(q_i)
\] (2.44)

We can rearrange this to get markups—price minus marginal cost all over price—on the left hand side:
\[
P(P(Q) - MC_i(q_i)) = \frac{-P'(Q)q_i}{P(Q)}
\] (2.45)

Multiplying the right hand side by \(1 = \frac{q}{Q}\) and letting \(s_i = \frac{q_i}{Q}\) be the market share of hospital \(i\), we get:

\[
\frac{P(Q) - MC_i(q_i)}{P(Q)} = \frac{-QP'(Q)}{P(Q)} s_i
\] (2.46)

But \(-\frac{QP'(Q)}{P(Q)}\) is just the inverse price-elasticity of demand, \(\varepsilon\), so we have:

\[
\frac{P(Q) - MC_i(q_i)}{P(Q)} = \frac{s_i}{\varepsilon}
\] (2.47)

Now multiply both sides by \(s_i\) and sum over all hospitals to get

\[
\sum_{i=1}^{n} \frac{P(Q) - MC_i(q_i)}{P(Q)} s_i = \frac{\sum_{i=1}^{n} s_i^2}{\varepsilon} = \frac{HHI}{\varepsilon}
\] (2.48)

The average markup in the hospital market is, therefore, a linear function of HHI. This Cournot competition model is the prevailing rationale for the linear specification of HHI in regression equations.

Now we consider the relationship between owner- and hospital-level HHI. First, note that if the hospital-level HHI is \(H_h = \sum_{i=1}^{n} s_i^2\), and a merger occurs between hospitals \(s_j\) and \(s_k\), the new HHI should be

\[
H'_h = H_h - s_j^2 - s_k^2 + (s_j + s_k)^2 = H_h + 2s_js_k
\] (2.49)
We can think of going from the hospital- to the owner-level HHI as simply recognizing the ‘mergers’ of the co-owned hospitals. If we define a ownership indicator function \( \theta_{ij} \)

\[
\theta_{ij} = \begin{cases} 
0 & \text{if } i \text{ and } j \text{ are co-owned} \\
1 & \text{otherwise}
\end{cases}
\] (2.50)

then the owner-level \( HH_{o} \) can be expressed as

\[
H_{o} = H_{h} + \sum_{i} \sum_{j \neq i} \theta_{ij} s_{i}s_{j}
\] (2.51)

We let \( \varphi \) denote the expected value of \( \theta_{ij} \) when we don’t know a priori if \( i \) and \( j \) are co-owned. The value \( \varphi \) can be thought of as the probability that any two hospitals in the sample are co-owned. This gives us the following:

\[
E(H_{o}) = E(H_{h}) + \varphi \sum_{i} s_{i} \sum_{j \neq i} s_{j}
\] (2.52)

Using the fact that \( \sum_{j \neq i} s_{j} = 1 - s_{i} \), and that the market shares are known, we can find

\[
E(H_{o}) = (1 - \varphi)H_{h} + \varphi
\] (2.53)

or

\[
H_{o} = (1 - \varphi)H_{h} + \varphi + \mu_{i}
\] (2.54)

for some zero-mean random variable \( \mu_{i} \).
When we have some dependent variable $y$ that has a linear relationship with HHI (e.g. markups), and that is independent of $\mu_t$, the estimated coefficient of $HHI_o$ in the regression model should be

$$\hat{\beta}_o = \frac{\text{Cov}(y, HHI_o)}{\text{Var}(HHI_o)}$$

$$= \frac{(1 - \varphi)\text{Cov}(y, HHI_h)}{(1 - \varphi)^2\text{Var}(HHI_h)}$$

$$= \frac{\hat{\beta}_h}{(1 - \varphi)}$$

Therefore, if the market has a large diversity of owners, and the value of $\varphi$ is close to zero, the estimated coefficients on $HHI_o$ and $HHI_h$ will not be significantly different. In short, the decision to use an owner- or hospital-level HHI would not have a material effect on the regression results. To see whether $\varphi$ is close to zero in practice, we estimate the real-world relationship between the owner- and hospital-level variables in the next section. Our results support the hypothesis that the choice between the two measures has little effect on regression estimations.

III. Data & Key Variables

To test the theories described above, we build upon a standard empirical model relating hospital pricing power and concentration. We have two objectives for the model: 1) to measure how patient travel distance affects hospital pricing power, and 2) to determine whether the hospital-level and owner-level Hirschman-Herfindahl indexes are interchangeable in our regression estimations. To measure prices and profits we require detailed hospital financial
statements, and to estimate the effects of travel distance we need a record of patient discharges and origins. Our source for both the financial and patient data is California’s Office of Statewide Health Planning and Development (OSHPD), which collects and publishes data on revenues, costs and volume for California’s more than 450 hospitals, in addition to maintaining a database on the state’s more than 4 million annual patient discharges. We perform our analysis using data from 2008, the most recent year for which patient origin and discharge data is currently available.

We discuss the theoretical form of our model and the construction of our key variables in the next several sub-sections.

A. Measuring Hospital Pricing Power

In other studies on hospital pricing, authors have used dependent variables including gross profit margins, markups, and of course, prices themselves (Dranove 1993: 185). For our regression model we use markups, which we define for a hospital $i$ as the price that insurers pay for the hospital’s services, minus the marginal costs of those services, all divided by price. The use of markups is appealing because it removes scalar effects on price and cost, such as changes in the general price level. This has the benefit of making our estimated coefficients more comparable with the results of studies using data from other time periods. In addition, as we briefly showed in the preceding section, markups have a nice theoretical relationship with the Hirschman-Herfindahl index under Cournot competition.

We calculate markups using the financial data from the OSHPD. For each hospital, the OSHPD reports revenues and units of service broken down by payer (e.g. Medicare, Medicaid, or private insurer) and type of procedure. The revenue statements also report total capitation
premiums paid by Managed Care Organizations (MCOs), along with total revenue deductions from contractual pricing discounts. While some studies have attempted to estimate prices for every kind of procedure (Krishnan 2001), we find this impracticable for our data because the OSHPD aggregates the capitation premiums and contractual deductions across procedure groups, and it would require ad hoc methods to redistribute them. Instead, we calculate the price for each hospital by taking gross revenue from MCOs, subtracting contractual deductions, adding capitation premiums, and dividing by total units of service. The result is effectively the average price paid by an MCO for a unit of medical service—either an inpatient day or outpatient visit. We focus on the prices that private insurers pay, rather than payers like Medicare, because the public-sector payers tend to pay rates that are based on a cost-estimating formula and only loosely related to hospital market power.

We use this method to estimate prices for each hospital, even though every hospital in California is required by law to publish a ‘chargemaster,’ or list of fees for its services. We do this because hospitals usually charge reduced prices to commercial insurers and government payers, so the chargemaster poorly reflects what most patients actually pay. An early study on shifts in hospital competition by Dranove (1993) confirmed that list prices are poorly related with hospital competition and other common explanatory variables. Our price estimations should better reflect what hospitals actually charge, and hence have a stronger relationship with hospital competition and patient travel.

In addition to prices, we also need a measure of marginal costs in order to calculate markups. However, like most financial statements, the OSHPD’s data provides total costs instead of marginal costs. One remedy is to divide total costs by units of service to find average costs, and then use these average costs as a proxy for marginal costs. However, average costs will
overstate marginal costs for all but the largest hospitals, since there are positive economies of scale in hospital operations. The markups that we calculate using average costs will therefore underestimate the true markups. To correct for this inaccuracy, we follow a suggestion by Dranove (1993) and include PPE/Sales on the right hand side of our regression model. By PPE we refer to the standard Plant, Property and Equipment account, net accumulated depreciation. After dividing PPE by sales we get a measure of average fixed costs. This new variable should explain the error caused by using average costs, allowing us to accurately measure the effects of our key variables.

Now that we have calculated markups, we are interested in expressing them as a function of HHI, patient travel distance, and our other independent variables. We can think of each of our explanatory variables as being either hospital-level or market-level, a common expository approach in the literature (Dranove 1993; Krishnan 2001: 219). For example, patient travel is a hospital-level variable, while HHI is a market-level variable—the difference being that two direct competitors in the same market should have the same HHI. The markup of hospital $i$ is then expressed in the following way:

$$M_i = \alpha + \beta^T K_i + \gamma^T H_i + \mu_i$$

(3.1)

where $M_i$ the markup, $\alpha$ is some constant, $K_i$ and $H_i$ are the column-vectors of market and hospital-level characteristics for hospital $i$, and $\beta^T$ and $\gamma^T$ are the row-vectors of coefficients for the market- and hospital-level variables respectively.

In addition to patient travel distance, our other hospital-level variables are as follows:

$Disc$ The hospital’s total patient discharges, measured in thousands of patients.
*Gov* Dummy variable equal to one if the hospital is owned by a city, county, or state, and equal to zero otherwise.

*NetNonOp* The hospital’s net non-operating revenue divided by total revenue. Non-operating revenue can come from the hospital’s investment earnings, research grants, taxing authority, or sales of property. To find net non-operating revenue, we take non-operating revenue and subtract non-operating expenses, like losses on assets. After dividing by total revenue, the result, NetNonOp, effectively measures the hospital’s ability to price below marginal cost—i.e., to have negative markups.

*NP* Dummy variable equal to one if the hospital is owned by a church or other non-profit organization, and equal to zero otherwise.

*PPE/Sales* A measure of fixed cost absorption found by dividing plant, property and equipment (net depreciation) by gross operating revenues.

*Spec* An index that measures hospital specialization, calculated by summing the squared revenue shares of the hospital’s procedure groups. This is like a Herfindahl index for revenue concentration (Zwanziger et al., 1996).

While *PPE/Sales* does control for economies of scale as described above, we still include total patient discharges as an explanatory variable because high patient volume may benefit hospitals for reasons other than fixed cost absorption. The largest hospitals may have stronger brands, better bargaining power with suppliers, or more experienced workforces. Hence, by controlling for discharges we should get better estimates for the effects of competition and patient travel.
The variables \textit{Gov} and \textit{NP} control for hospital ownership. In general, we expect that government and non-profit hospitals will place a greater importance on total patients treated, providing educational services, and other objectives beside profit maximization. When such hospitals have market power, they will still probably raise their prices to some degree, because as Melnick et al. point out, non-profits have an incentive to seek resources to spend on their charitable goals (Melnick et al., 1999). Nonetheless, non-profit and government hospitals will likely have smaller markups than their for-profit counterparts due to their desire to keep medical prices affordable. Hence, we expect \textit{Gov} and \textit{NP} to have negative effects on markups. In addition to their charitable goals, the non-profit hospitals may also benefit from subsidies, favorable tax treatment, and affiliation with wealthy parent organizations. Rather than leave such effects to the dummy variables, we control for them more directly by including net non-operating income (\textit{NetNonOp}) on the right-hand side of our model. Like \textit{Gov} and \textit{NP}, the \textit{NetNonOp} variable should have a negative effect on markups, because it effectively measures the hospital’s ability to price below marginal cost.

Our last hospital-level variable is \textit{Spec}. In our sample, the hospitals with the greatest average patient travel distance also tend to be highly specialized—most commonly in the fields of pediatric care, psychiatric treatment or substance abuse rehabilitation. The correlation between our specialization index \textit{Spec} and patient travel variables is $\rho = .51$. Moreover, specialized hospitals tend to command premium prices because they are relatively scarce and the services they provide often have small elasticities of demand. Hence, it is essential to control for specialization in order to accurately measure the effect of patient travel on hospital markups. Our construction of the specialization index is similar to the “Service Mix HHI” defined by Zwanziger et al. (1996). For each hospital, we divide the revenue from each procedure (e.g.
echocardiography, pediatric intensive care) by the hospital’s total revenue, and then we sum the squares of these proportions across all procedure groups. As noted by Zwanziger et al., the properties of this measure are very similar to those of a Herfindahl index.

For our market-level variables, on the other hand, we include HHI, private insurance penetration, and the average age and income for the area of the hospital’s patients:

**HHI** A measure of market concentration equal to the sum of squared market shares of hospitals (or hospital owners) in the market. One of our empirical goals is to demonstrate that the owner- and hospital-level HHIs are interchangeable in regression models. To this end, we run our model using both the owner- and hospital-level measures and report the estimated coefficients in the results section.

**Age** The average age of patients visiting the hospital.

**Income** The average annual salary in thousands of dollars in the hospital’s market.

**InsurePen** The insurance penetration of the hospital’s market, i.e. the proportion of patients covered by private insurance in the hospital’s market.

The computation of these market-level variables depends to a large degree on how we define our hospital markets. There is no standard way to construct markets in the existing literature, so we discuss the alternatives and describe our method below. Then we detail the calculation of the market level variables.

### B. Defining Hospital Markets

In industrial economics, there is no gold standard for how to define a market, and this is particularly evident in hospital research. To construct hospital markets, some authors use
geopolitical units like cities, counties, and ZIP codes (Dranove 1993; Dranove 2002; Lynk 1995), while others use distance-based measures (Patel 1994). However, both the geopolitical and distance-based markets are less than ideal, because they rely on abstract regions that have little bearing on actual patient choices. For our research, we instead use a ‘patient-flow’ approach, which defines markets based on actual patient locations and discharges (Bamezai et al. 1999; Krishnan 2001). By looking at patient origins, we can see the regions from which each hospital draws its patients, and observe whether any two hospitals are truly competitors. This method should make our variables like HHI more economically relevant. We now describe in detail the construction of our hospital markets and associated variables.

To construct our markets with the patient-flow approach, we use an OSHPD database on patient discharges in California in 2008. For more than four million discharges, the OSHPD reports the patient’s home ZIP code, method of payment, and the hospital at which the patient received service (OSHPD 2008). While our data comes from a different source, the basic procedure of our construction is very similar to that of Bamezai et al. (1999). We define a hospital’s market to be the collection of ZIP codes from which the hospital draws patients. Of course, not every ZIP code contributes equally to the hospital’s total volume. Therefore, when we construct hospital-level variables, we give greater weight to the ZIP codes that are responsible for a larger proportion of the patients. For example, to find the HHI for a given hospital, we calculate a separate HHI for each ZIP code in the hospital’s market, and then perform a weighted average on these HHIs, with the weights given by the proportion of patients that come from each respective code.

To calculate HHI for each ZIP code, we sum the squared market shares of hospitals in the code. With this approach, one technical consideration is that we might count two hospitals as
competitors even if they are specialized in wildly different fields. For example, if a psychiatric hospital and a cardiac hospital serve a given ZIP code, where each accounts for half the patient discharges, we might assign the ZIP code an HHI of one half. However, this would not be an accurate picture of the hospital market in the area, because the two hospitals draw from different pools of patients and are not direct competitors. Therefore, to control for such specialization, we compute procedure-specific HHIs for each of eleven diagnoses groups: infections, neoplasms, endocrine system, psychoses & neuroses, circulatory, respiratory, digestive, genitourinary, pregnancies & neonatal care, musculoskeletal disorders, and injuries and complications. We calculate eleven diagnosis-specific HHIs for each hospital, and then we average them together, using as weights the proportion of the hospital’s patients with the respective diagnosis. With this construction, a specialized hospital will have an HHI that reflects the true competitive environment for its particular services.

The diagnoses-specific method is used to construct both owner- and hospital-level HHIs. For the owner-level version, the only change is that we combine the market shares of co-owned hospitals when calculating the HHI for each ZIP code. To test our theory that owner and hospital-level HHI are practically interchangeable, we report our regression results using both owner- and hospital-level HHI in section IV.

Two other market level variables, $AGE$ and $InsurePen$, are calculated with the same data and general method. The OSHPD discharge data reports age ranges, which we use to calculate the average age of each hospital’s patients. To find insurance penetration, we simply take the proportion of patients who paid with private insurance in each ZIP code, and average this across the ZIP codes in the market of each hospital.
We also control for the average income in each hospital’s market. The income data comes from the 2008 County Business Patterns: ZIP Code Business Statistics Survey of the U.S. Census Bureau. The survey reports the total number of paid employees for every ZIP code, along with the total paid wages in thousands of dollars. To find the average income for each ZIP code, we simply divide the total wages by the number of employees. While this measure only includes cash compensation and omits employee benefits, it should give us a good picture of patient buying power. We expect that hospitals whose patients have higher average incomes will command larger price premiums, due to a greater demand for higher margin, specialized services.

C. Patient Travel

To test if patient travel distance has a negative effect on hospital pricing power, we use two variables: average patient travel distance ($\bar{d}$), and distance from the hospital to its nearest neighbor ($d_1$). We use these two variables to form an estimate of $d_2/d_1$, as in section II.A, which we include as a variable in our regression model. This approach avoids the technical difficulties of assigning hospitals to clusters and calculating $d_2/d_1$ directly.

The OSHPD database reports the physical address of each hospital and the home ZIP code of each patient. We use the addresses and ZIP codes to estimate the latitude and longitude of each hospital and patient origin, and we compute travel distance with a great circle distance formula. By averaging across all discharges, we find the average patient travel distance $\bar{d}_j$ for each hospital. We use the same approach to calculate the distance between each pair of hospitals, and we denote the distance from each hospital $j$ to its closest neighbor by $d_{1,j}$. 
In our model in II.A, the average patient travel distance for each hospital is a linear combination of the variables \( d_1 \) and \( d_2 \):

\[
\bar{d}_j = \omega_{1,j}d_{1,j} + \omega_{2,j}d_{2,j}
\]  
(3.2)

Hence, if we divide \( \bar{d}_j \) by \( d_{1,j} \) for each hospital, the result is a linear function of \( d_{2,j}/d_{1,j} \) with positive parameters \( \omega_{1,j} \) and \( \omega_{2,j} \). We use \( \hat{d} \) to denote the result:

\[
\hat{d} = \frac{\bar{d}_i}{d_{1,j}} = \omega_{1,j} + \omega_{2,j} \frac{d_{2,j}}{d_{1,j}}
\]  
(3.3)

By including \( \hat{d} \) in our regression model, we can test the hypothesis that \( d_2 \) has a negative effect on hospital prices when it is large compared to \( d_1 \). Between hospitals, the parameters \( \omega_{1,j} \) and \( \omega_{2,j} \) can vary, but this approach amounts to using a proxy for \( d_2/d_1 \) if we assume that the unobserved \( \omega_{1,j} \) is independent of \( \omega_{2,j} \) and the explanatory variables. More intuitively, this new variable simply measures the tendency of each hospital to draw patients from beyond the distance to its closest neighbor. This should allow us to test the theory that hospitals have less pricing power when they are close together and draw patients from a wide area. We report the estimated coefficients of our model in the next section.

**IV. Results**

Table 1 shows descriptive statistics for our regression variables. The dependent variable, hospital markup, is slightly negative on average, but this is explained by net non-operating revenue and PPE/sales as discussed in section III.A.
Table 1
Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Markup</td>
<td>-0.01</td>
<td>0.14</td>
</tr>
<tr>
<td>Distance Ratio</td>
<td>19.17</td>
<td>44.19</td>
</tr>
<tr>
<td>Distance Interior</td>
<td>5.01</td>
<td>7.66</td>
</tr>
<tr>
<td>HHI hospital-level</td>
<td>0.29</td>
<td>0.14</td>
</tr>
<tr>
<td>HHI owner-level</td>
<td>0.32</td>
<td>0.15</td>
</tr>
<tr>
<td>Discharges</td>
<td>10.49</td>
<td>8.84</td>
</tr>
<tr>
<td>Government</td>
<td>0.17</td>
<td>0.38</td>
</tr>
<tr>
<td>Non Profit</td>
<td>0.49</td>
<td>0.50</td>
</tr>
<tr>
<td>Net Non-Operating Revenue</td>
<td>0.54</td>
<td>2.19</td>
</tr>
<tr>
<td>Plant, Property &amp; Equipment / Sales</td>
<td>0.34</td>
<td>0.23</td>
</tr>
<tr>
<td>Specialization</td>
<td>0.12</td>
<td>0.10</td>
</tr>
<tr>
<td>Age</td>
<td>45.80</td>
<td>12.83</td>
</tr>
<tr>
<td>Income</td>
<td>36.72</td>
<td>6.17</td>
</tr>
<tr>
<td>Insurance Penetration</td>
<td>0.31</td>
<td>0.17</td>
</tr>
</tbody>
</table>

Source: Author's computation from 2008 OSHPD and U.S. Census.

The owner-level Herfindahl index is always larger than the hospital-level HHI, and its mean is therefore greater. Furthermore, when we regress the owner-level on the hospital-level HHI, we get a strong linear relationship as shown in Figure 12. The estimated slope of this line is close to one, supporting our theory in section II.C that the HHI variables should give similar results in regression models.
Figure 12: Owner vs. Hospital-Level HHI

The estimates for our main model appear in Table 2. The original specification uses owner-level HHI and the full set of control variables, and its coefficients appear in the first column. We report variations on this specification in columns (2), (3), and (4). In column (2), we use a hospital-level HHI, and in (3) and (4) we omit the HHI variable altogether, and (3) includes income while (4) does not.
Table 2
Regression Results for Hospital Markups

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance Ratio</td>
<td>-.00029*</td>
<td>-.00028*</td>
<td>-.00026*</td>
<td>-.00029*</td>
</tr>
<tr>
<td></td>
<td>(-1.83)</td>
<td>(-1.79)</td>
<td>(-1.66)</td>
<td>(-1.81)</td>
</tr>
<tr>
<td>Distance Interior</td>
<td>-.00036</td>
<td>-.00019</td>
<td>.0015</td>
<td>.0022**</td>
</tr>
<tr>
<td></td>
<td>(-.31)</td>
<td>(-.15)</td>
<td>(1.46)</td>
<td>(2.33)</td>
</tr>
<tr>
<td>HHI hospital-level</td>
<td>. . .</td>
<td>.144**</td>
<td>. . .</td>
<td>. . .</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.45)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HHI owner-level</td>
<td>.162***</td>
<td>. . .</td>
<td>. . .</td>
<td>. . .</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.93)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discharges</td>
<td>.002</td>
<td>.002</td>
<td>.002*</td>
<td>.002*</td>
</tr>
<tr>
<td></td>
<td>(1.57)</td>
<td>(1.61)</td>
<td>(1.89)</td>
<td>(1.91)</td>
</tr>
<tr>
<td>Government</td>
<td>-.039**</td>
<td>-.036*</td>
<td>-.032</td>
<td>-.030</td>
</tr>
<tr>
<td></td>
<td>(-2.02)</td>
<td>(-1.88)</td>
<td>(-1.64)</td>
<td>(-1.52)</td>
</tr>
<tr>
<td>Non Profit</td>
<td>-.034**</td>
<td>-.032**</td>
<td>-.031**</td>
<td>-.029**</td>
</tr>
<tr>
<td></td>
<td>(-2.35)</td>
<td>(-2.24)</td>
<td>(-2.12)</td>
<td>(-1.98)</td>
</tr>
<tr>
<td>Net Non-Operating Revenue</td>
<td>-.031***</td>
<td>-.031***</td>
<td>-.032***</td>
<td>-.032***</td>
</tr>
<tr>
<td></td>
<td>(-8.57)</td>
<td>(-8.60)</td>
<td>(-8.80)</td>
<td>(-8.70)</td>
</tr>
<tr>
<td>Plant, Property &amp; Equipment / Sales</td>
<td>-.065**</td>
<td>-.065**</td>
<td>-.065**</td>
<td>-.069**</td>
</tr>
<tr>
<td></td>
<td>(-2.21)</td>
<td>(-2.21)</td>
<td>(-2.17)</td>
<td>(-2.30)</td>
</tr>
<tr>
<td>Specialization</td>
<td>.26***</td>
<td>.25***</td>
<td>.23***</td>
<td>.25***</td>
</tr>
<tr>
<td></td>
<td>(3.27)</td>
<td>(3.19)</td>
<td>(2.96)</td>
<td>(3.17)</td>
</tr>
<tr>
<td>Age</td>
<td>.0017***</td>
<td>.0016***</td>
<td>.0016***</td>
<td>.0014***</td>
</tr>
<tr>
<td></td>
<td>(3.07)</td>
<td>(3.07)</td>
<td>(2.92)</td>
<td>(2.64)</td>
</tr>
<tr>
<td>Income</td>
<td>-.002*</td>
<td>-.002*</td>
<td>-.003**</td>
<td>. . .</td>
</tr>
<tr>
<td></td>
<td>(-1.91)</td>
<td>(-1.96)</td>
<td>(-2.30)</td>
<td></td>
</tr>
<tr>
<td>Insurance Penetration</td>
<td>.20***</td>
<td>.20***</td>
<td>.21***</td>
<td>.18***</td>
</tr>
<tr>
<td></td>
<td>(4.40)</td>
<td>(4.47)</td>
<td>(4.69)</td>
<td>(4.15)</td>
</tr>
<tr>
<td>Intercept</td>
<td>-.100*</td>
<td>-.089</td>
<td>-.044</td>
<td>-.133***</td>
</tr>
<tr>
<td></td>
<td>(-1.77)</td>
<td>(-1.58)</td>
<td>(-.826)</td>
<td>(-3.53)</td>
</tr>
<tr>
<td>Adjusted</td>
<td>.355</td>
<td>.349</td>
<td>.338</td>
<td>.329</td>
</tr>
</tbody>
</table>


*** Significant at 1% level.
** Significant at 5% level.
* Significant at 10% level.
The variables HHI and $d_1$ both represent concentration, so it is little surprise that the coefficient of $d_1$ is small when we include HHI in (1) and (2). However, after removing HHI in (3) and (4), we see that $d_1$ has a positive effect on hospital pricing power, as predicted in II.A. The estimated coefficient for $d_1$ is significant at the 5% level in Eq. (4). Intuitively, when the distance from a hospital to its nearest neighbor is large, that hospital has a greater buffer for its profit margin, and hence higher markups. The estimated coefficient in (4) suggests that increasing this distance by one mile leads to a .2% increase in markups.

In every specification, the coefficient on the distance ratio is negative and significant at the 10% level. This ratio is a proxy for $d_2/d_1$, so the negative coefficient supports our theory in section II.A that $d_2$ reduces pricing power when it is large compared to $d_1$. When $d_2$ is large for a given hospital, its neighbors will be more aggressive in their price competition, as explained in II.A. Thus, to protect its market share, the hospital needs to be more cautious in its pricing, leading to lower markups. The negative coefficient on $d_2/d_1$ also supports our assumption of non-convex patient travel costs. When travel costs are non-convex, each hospital has more to gain by dropping its ‘total cost’ curve, a fact we that used to derive the negative relationship between $d_2$ and hospital prices.

The estimated distance coefficients are not significant at, say, 1%, but this imprecision has several explanations. First, the variable $d_1$ is essentially a one-dimensional measure. It tells us the distance to the nearest hospital, but a more relevant variable might instead consider the several closest surrounding hospitals. Second, the negative relationship between $d_2$ and $d_1$ only occurs when the hospitals are sufficiently clustered (see Eq. 2.14). If a large number of hospitals in our sample are not clustered, the negative relationship between $d_2/d_1$ and hospital pricing power is weakened. When we restrict our sample to the hospitals which have a neighbor within
five miles, the $d_2/d_1$ coefficient is more negative and becomes significant in model (4) at the 5% level. Finally, the OSHPD provides only ZIP codes rather than full patient addresses, so when we calculate travel distance we must estimate the true patient origins. While this measurement error should be uncorrelated with our control variables and should not lead to biased coefficients, it does add imprecision to our estimates.

We also note that the coefficients for the HHI variables agree with our discussion in section II.C. The owner-level HHI has a slightly greater effect than the hospital-level HHI. This is consistent with the predicted relationship in Eq. 2.55, and suggests that the parameter $\varphi$ is close to zero. Hence, while the owner-level HHI may be more economically relevant, our result suggests that the hospital-level HHI still captures much of the useful information. Further consolidation in the hospital industry could change this picture by increasing $\varphi$, but for now, at least, it appears that estimations with hospital- and owner-level HHIs are similar in magnitude and roughly comparable.

Finally, it is worth mentioning that controlling for specialization is crucial to our empirical results. As shown in Table 2, the effect of specialization on hospital pricing power is positive and highly significant. Moreover, specialized hospitals tend to draw patients from farther afield, while often being clustered in cities or around universities. Hence, if we remove specialization from our model, the coefficients on the distance ratio are biased upward and become insignificant. By including specialization we control for this effect and get a more accurate test of our theory. The results confirm that when a hospital draws patients from beyond its local cluster, it tends to have smaller price markups.
V. Discussion

In standard I/O theory, travel distance acts as a shield for profits, because customers are willing to pay higher prices to firms that are nearby. Both theoretically and empirically, we have argued that the interior distance of a cluster does indeed offer such protection—i.e., it is better for a hospital when its closest neighbors are far away. However, our regression model suggests that the exterior distance of a cluster has the opposite effect. When a hospital draws patients from beyond the distance to its nearest neighbor, our model shows that its markups are smaller. Thus, the relationship between average patient travel distance and hospital prices can actually be negative.

This empirical result has two implications. First, in the context of our theory, a negative distance-price relationship only occurs when travel costs are non-convex and hospitals are sufficiently clustered. When we constructed the subsidy, we used non-convex travel costs, so our finding of the negative relationship offers additional support for that assumption. The presence of hospital clustering also made the subsidy less expensive, since the maximum subsidy was strictly less than the cost of traveling between the hospital and its nearest neighbor. Hence, our data supports the theoretical conditions for the subsidy and suggests that its real world cost may not be too onerous.

Second, the negative relationship between travel distance and prices means that the usual I/O narrative requires a caveat. Instead of referring to ‘customer travel distance’ generally, it may be more informative to consider travel distance in two parts: interior, like the $d_1$ in our theory, and exterior, like the $d_2$. When travel costs are non-convex and firms have agglomeration economies, the $d_2$ part of customer travel can actually have a negative effect on firm pricing.
power. An interesting subject of future research would be to take this relationship into account in a more complete model of endogenous firm location.

For our theory of the subsidy, we used a highly stylized setting and various assumptions which may seem unrealistic. However, the intuition of the subsidy has practical appeal. In short, by influencing travel costs, the insurer makes it easier for hospitals to gain patients through price reductions. When each hospital wants to reduce its price, given the prices of its competitors, the result is a prisoner’s dilemma in which every hospital lowers its price and becomes less profitable in the end. Moreover, many of the assumptions of the subsidy are easy to check on a hospital-by-hospital basis, particularly with the information available to insurers. For example, an insurer can estimate hospital gross margins by using billing records to find prices and financial statements to estimate costs. The subsidy is then designed so that its immediate impact is to reduce the market share of the hospital with the higher gross margin, as explained in section II.A. In short, despite our simplified model, the intuition of the subsidy is quite simple: insurers can make markets more competitive by making patients more mobile. As a substantial part of hospital price premiums is due to the local competitive environment, travel subsidies have great potential for promoting lower medical prices.

The relationship between the two HHI variables, derived in Eq. 2.55 and confirmed by our regression model, suggests that the owner-level HHI does have a stronger effect on pricing power, but only by a small amount. If a researcher has data on hospital ownership, it is better to use the owner-level HHI. But if such data is not available, the hospital-level HHI can still capture much of the useful information and yield estimates that are roughly comparable with those from the owner-level approach. As further consolidation takes place, however, it will become more
important to use the owner-level variable. The effects of consolidation are captured by the parameter $\phi$ in our theoretical model.

Finally, we note that the travel subsidies may allow MCOs to obtain price reductions in markets where traditional methods have failed. A travel subsidy does not entirely rely on the existing level of competition, but rather encourages new competition by creating contested regions. When a hospital is unswayed by network contracting, it may still respond to a loss of patient volume by cutting its price, initiating the round of reductions described in our theory. Moreover, since insurers have good information on hospital prices, the subsidy could be further refined to direct patients to the lower-priced hospital. The subsequent savings would defray part of the initial cost, and give the subsidy an immediate positive impact. When patients have poor information, the subsidy has even greater potential to be profitable, because it allows insurers to signal to patients which hospital has the better value.

References


