Hospital Choices, Hospital Prices, and Financial Incentives to Physicians[†]

By KATE HO AND ARIEL PAKES*

We estimate an insurer-specific preference function which rationalizes hospital referrals for privately insured births in California. The function is additively separable in: a hospital price paid by the insurer, the distance traveled, and plan- and severity-specific hospital fixed effects (capturing hospital quality). We use an inequality estimator that allows for errors in price and detailed hospital-severity interactions and obtain markedly different results than those from a logit. The estimates indicate that insurers with more capitated physicians are more responsive to price. Capitated plans send patients further to utilize similar quality, lower-priced hospitals; but the cost-quality trade-off does not vary with capitation rates. (JEL G22, H51, I11, I13, I18, J44)

The Patient Protection and Affordable Care Act of 2010 includes provisions to establish accountable care organizations (ACOs) in the Medicare program. These are groups of health care providers who offer services to a large number of patients and are eligible to share in any cost savings they achieve for the Medicare program.[1] Private-sector ACOs are forming in parallel to this initiative, often with very similar structures and payment arrangements. The proportion of savings which ACOs keep is linked to their performance on quality standards. The goal of the initiative is to reduce health care costs in both Medicare and private insurance and to improve coordination of care. The tying of cost savings to quality standards is designed to mitigate the concern that an increase in incentives to control costs might lead to a decrease in the quality of care.

The health policy literature has noted that similar cost control incentives are currently utilized by some health maintenance organizations (HMOs) in California and elsewhere.[2] In this paper we investigate the impact of these incentives on the cost and quality of care provided to patients, focusing in particular on the choice of hospital for privately insured patients giving birth. We use hospital discharge data for managed care enrollees from California in 2003 to estimate a hospital referral model

which posits that referrals are a plan-specific function of the prices insurers pay to the hospitals, distance from home to hospital, and a severity-specific measure of "hospital quality." The results are used first to study the relationship between referrals to hospitals, hospital prices, and differences in the cost-control incentives used by different insurers. We then refine our estimates in a way that allows us to analyze how the implicit trade-offs between price, our severity-specific measure of hospital quality, and distance differ with these incentives.

The question of whether patients, enrolled in insurers which give physician groups an incentive to control hospital costs, are admitted to low-priced hospitals is important for several reasons. Hospital costs make up more than 30 percent of national health care spending so our findings are potentially important for the overall impact of ACOs on costs. Moreover, any analysis of the effect of hospital mergers on hospital prices will require an assumption regarding the effect of price increases on referrals. Assuming away this effect will likely result in an overestimate of the price increases that will result from a merger. Similarly a merger analysis which does not take into account price effects will likely overestimate the merged entity's incentives to invest in new high-cost technologies and underestimate its incentives to invest in cost-reducing technologies.

The process by which a patient chooses a hospital involves multiple players. Decisions are made by referring physicians in consultation with their patients. Insurers often attempt to influence physician choices through direct financial incentives. In California in particular, they often remunerate the physician group through fixed (capitated) monthly payments per patient to cover the cost of patient services.[3] The most common type of capitation involves payments covering only services provided by the physician group but when this is the case "shared risk arrangements" almost always apply, under which a target is set for total (including hospital) spending and cost savings or overruns relative to the target are shared between the physician group and the insurer. Less commonly, global capitation payments cover the cost of all services received by patients, including inpatient hospital stays. Both types of capitation contracts generate incentives for physician groups to refer their patients to low-cost hospitals. Physician groups often pass the financial incentives on to their member physicians, for example through profit-sharing arrangements and bonuses based on costs per enrollee. They may also make physicians' promotion on the pay scale contingent (formally or informally) on their management of costs. Thus, if we assume that higher-quality hospitals negotiate higher prices, physicians in groups receiving capitation payments face a trade-off between incentives to reduce costs and other aspects of the choice such as quality of care and convenience. Moreover the extent to which patients and/or doctors care about these factors is likely to vary with the severity of illness.

The incentives for ACOs and for the California insurers in our data are similar in that the provider group either bears the financial risk for hospital payments or benefits from hospital savings relative to a benchmark.[4] Also in both cases the incentives

---

[3] Several previous papers describe the contractual arrangements between insurers and physicians in California, including Rosenthal et al. (2001, 2002) and Grumbach et al. (1998a,b).

[4] The Affordable Care Act also introduces other provider payment mechanisms that are designed to control costs, such as the replacement of fee-for-service payments with "bundled" payments where providers receive a fixed fee for

are based on the costs incurred by the group (rather than by individual physicians), with no formal guidance on how these incentives are passed down to individual physicians or patients. Approximately 430 ACOs had been set up by January 2013 (Muhlestein 2013). The details of their payment arrangements are evolving over time, but they all involve some form of shared savings when costs are less than a benchmark.[5] Their structure also varies: in 2013, 189 were integrated with a hospital system, while most of the remainder were sponsored by physician groups which contracted with hospitals outside the organization. We come back to some of the likely implications of our results for these distinctions below.

Our analysis uses hospital discharge data to estimate an insurer-specific choice function which rationalizes hospital referrals for women giving birth in California. Unfortunately our dataset does not identify the physician referring each patient to her hospital; we therefore cannot observe physician behavior directly. However we do observe each patient's insurance carrier and the extent to which each carrier uses capitation payments. In 2003, 73 percent of payments made to primary physicians by the six largest carriers in our data were capitation payments; the proportions varied substantially across carriers from 97 percent for Pacificare to 38 percent for Blue Cross (more detail on the data underlying these numbers is given below). We ask whether the observed referrals for patients whose insurers have different capitation rates indicate different trade-offs between price, quality, and patient convenience factors. We view the allocation of a patient to a hospital as the outcome of a multi-stage process that includes, for example, the patient's choice of obstetrician and the obstetrician's allocation to affiliated hospitals as well as the obstetrician's choice of hospital for the particular patient. We do not model the protocol that leads to these choices; rather, we estimate the relative weights on different factors that emanate from that protocol and the extent to which these weights are related to capitation rates.

The analysis builds on the previous literature on hospital demand. Previous papers consider the factors affecting patients' hospital choices in some detail but almost exclusively make the simplifying assumption that the hospital is chosen without regard for the price paid by the insurer. To include the price variable one has to address three problems. First, the observed price is a "list price" for the relevant hospital discharge. Prices actually paid by the insurer are discounted versions of these list prices. To address this issue we import data on the average hospital discount from hospital financial reports. The true discount could vary across insurers and we treat that in two complementary ways: we allow for errors in our price variable and present results which use additional data to estimate the variation in discounts across insurers. Second the expected price that generates hospital choices is inherently unobservable. We assume that expectations are on average correct, and construct a price variable which is the average realized price for patients admitted with the same diagnosis and similar comorbidities at that hospital. Those predictions will have estimation error,

---

an episode of care rather than for each individual service. As noted by a referee, the efficacy of these arrangements is also likely to be affected by the extent of capitation, but these relationships are beyond the scope of this paper.

[5] See the Department of Health and Human Services (2012). Initial arrangements for Medicare ACOs were largely based on shared savings with no requirement to share in any losses relative to the benchmark, but as time passes increasingly they involve physician groups bearing at least some financial risk.

but the estimation procedure we develop averages those errors out.[6] The third problem relates to price endogeneity: the expected price for a patient with a particular diagnosis is likely to be correlated with the unobserved hospital quality for that diagnosis. We control for these unobservables by developing an estimation procedure which allows for hospital fixed effects that vary freely with severity of diagnosis.

We begin with a standard logit model for hospital choice. This does not allow for errors in the price variable and has a limited ability to allow for hospital fixed effects that vary with severity of diagnosis. We expect the omission of hospital severity interactions to bias the price coefficient upwards, and measurement error in the price measure to bias the price coefficient toward zero. When we pool all delivery and birth discharges we obtain a positive and significant coefficient on price. However, when we narrow the sample to more homogeneous diagnosis groups the price coefficient becomes negative. We then allow the price coefficient to vary by insurance carrier and find that the carriers with the highest proportion of payments to physicians made through capitation contracts have price coefficients that are significantly more negative than those of other carriers with a higher proportion of fee-for-service contracts. Since neither endogeneity nor errors in variables are fully addressed by this technique, we doubt that the estimates obtained here measure accurately the true responses to price.

So we develop a methodology which addresses these problems. It is based on revealed preference; we assume that the hospital chosen for each patient is preferred to any of the other hospitals in her choice set. The choice function is assumed to be additively separable in the price paid by the insurer, distance, and an insurer- and severity-specific hospital quality. We identify pairs of patients who have the same severity and are enrollees in the same insurer but who chose different hospitals. By defining the alternative of each patient as the chosen hospital of the other and summing the two patients' inequalities, we difference out the severity-specific hospital quality terms from the utility equation. By averaging the resulting inequalities over patients and hospitals we eliminate the effects of errors in price measurement. The result is a relatively straightforward estimator of bounds on the (normalized) price coefficient.[7]

The estimates indicate that the price coefficients are far more negative than in the logit analysis and are ordered with respect to the plans' capitation rates. That is, the price paid by insurers to hospitals does impact referrals, and the price response is more elastic for insurers whose physician groups are more highly capitated. We show that these results are robust to a number of perturbations to the specification used in the estimation. We then use the price coefficients to back out bounds on the plan-, hospital-, and severity-specific quality terms and find them to be highly correlated across plans. We therefore add structure and estimate a model where the quality terms for the different plans are affine transforms of one another. This allows us to represent preferences as a linear function of price, quality, and distance which differs across plans only in the coefficients of these variables. We can then examine how the trade-offs between

---

[6] Our baseline results assume no measurement error in distance. Later we consider a specification which is robust to errors generated by the distance between the centroids of our geographic areas and true locations, and find no major difference in results.

[7] The analysis is similar in spirit to previous papers that match treatment to control groups based on observable data and assume that unobserved information does not affect response to treatment. The propensity score literature (see Rosenbaum and Ruben 1983), and difference-in-differences analyses more generally, fall into this category. A difference is that we base our estimates on averages of the moment inequalities implied by our model, as in Pakes et al. (2011) and Pakes (2010).

price, quality, and distance vary with capitation rates. Though in absolute value the price coefficient varies directly with the capitation rate, the ratio between the price and quality coefficients, or trade-off between price and quality, is evaluated in much the same way across plans. In contrast the trade-off between distance and price is evaluated differently. That is, highly capitated, more price-sensitive plans tend to send their patients longer distances to obtain the same quality service at a lower price (but do not trade off costs against quality differently). To the extent that "hospital quality" is related to health outcomes, we can examine the relationship of quality to capitation rates directly by importing data on measures of health outcomes for the mothers and infants into our data. Conditional on patient severity, there is no evidence that capitation rates are related to any of our health outcome measures.

The remainder of the paper is structured as follows. In Section I we discuss the relevant previous literature. Section II describes important features of the market and Section III describes the data. Section IV sets out the full model we wish to analyze; Sections V and VI summarize the restrictions required for the logit and inequalities methods and set out their results; Section VII considers the trade-off between price, quality, and distance; and Section VIII concludes.

I. Previous Literature

Two sets of previous papers are relevant for our analysis. The first, summarized by Glied (2000), considers HMO gatekeeping and cost controls. Glied's summary suggests that HMOs have lower inpatient admissions and costs than other insurers; however these papers do not analyze the relationship between hospital price and referrals.[8] There are a few more recent studies that consider similar questions. For example, Cutler, McClellan, and Newhouse (2000) compare the treatment of heart disease in HMOs and traditional insurance plans and find that HMOs have 30 percent to 40 percent lower expenditures. Virtually all the difference comes from lower unit prices rather than differences in actual treatments. However they do not investigate whether price reductions are due to lower negotiated prices within a hospital or due to referring patients to cheaper hospitals (the focus of our study). Gaynor, Rebitzer, and Taylor (2004) look in more detail at how HMOs achieve cost savings. They analyze physician responses to group-based financial incentive contracts within a single HMO. They find that spending on medical utilization, particularly for outpatient services, increases with the size of the physician group receiving group-based incentives. That is, spending is negatively correlated with the power of incentives to limit these expenditures.[9]

There are also some papers evaluating recent initiatives which implement cost-control incentives like those that may be used for accountable care organizations. For example the Alternative Quality Contract (AQC) was adopted by Blue Cross Blue Shield of Massachusetts in 2009. It introduced physician incentives similar to those created by the global capitation contracts in our data. Physician groups entered into five-year global budget contracts under which they received a budget per enrolled patient and were accountable for costs of all services provided to those

---

[8] More recent reviews by Chandra, Cutler, and Song (2012) and McClellan (2011) come to similar conclusions.
[9] Other recent papers considering the responsiveness of health care providers to financial incentives include Ketcham, Léger, and Lucarelli (2012), Limbrock (2011), and Bajari et al. (2012).

patients, including inpatient care. Song et al. (2011) use a difference-in-differences analysis (where the first difference is across time and the second is between intervention and control groups) to analyze the impact of this initiative. They find that, in the first year, it was associated with reduced growth in spending on outpatient services and improved quality of care and that most of the savings came from referring patients to lower-priced providers. We also conducted a preliminary analysis of the response of the hospital referrals of physicians to capitation payments in Ho and Pakes (2011). There we regressed a severity-adjusted price measure on the proportion of the insurer's payments to primary physicians that were capitated and market fixed effects.[10] The coefficient on capitation was negative and statistically significant; consistent with the hypothesis that insurer capitation payments influence physician referrals. However, simple regressions like these cannot provide more than suggestive evidence since they do not account for the trade-offs made between price, other hospital characteristics, and convenience factors in the hospital choice equation.[11]

The second relevant literature estimates discrete choice models of hospital demand: see Gaynor and Vogt (2000) for a survey.[12] Almost all of these papers exclude the price paid by the insurer to the hospital from the utility equation. One exception is Gaynor and Vogt (2003), which uses assumptions to define a single price index for each hospital that is included in the utility equation. However, that paper assumes away interactions between patient and hospital characteristics in determining procedures and therefore prices. It also does not consider the impact of physician incentives on the price coefficient.

II. Background on the Market

The analysis in this paper focuses on enrollees of health maintenance organizations (HMOs).[13] The seven largest HMOs had 87 percent of the California HMO market at the end of 2002. Our analysis focuses on six of these seven: we exclude Kaiser Permanente because the prices paid by this vertically integrated insurer to its hospitals are not observed in our data.[14]

Each HMO contracts with a network of providers (physicians and hospitals); enrollees seek care within that network. Each pregnant woman chooses an obstetrician from within the network and is referred to one of the small number of network hospitals with which the obstetrician is affiliated. Our analysis controls for any variation in networks across insurers by conditioning on the observed provider

---

[10] We adjusted for severity by constructing the following price ratio measure: $p_i^{ratio} = \frac{p_i}{\overline{p}_{s_i}}$ where p_i was the hospital price for patient i and $\overline{p}_{s_i}$ was the average of that variable for same-severity patients across all hospitals in the sample.

[11] Duggan (2000) considers hospital referrals for Medicaid patients. He finds that private hospitals in California responded to new financial incentives to treat Medicaid patients by cream-skimming the most profitable Medicaid patients from publicly owned hospitals. The reallocation was especially pronounced for pregnant women.

[12] Examples include Luft et al. (1990), Burns and Wholey (1992), Town and Vistnes (2001), Capps, Dranove, and Satterthwaite (2003), and Ho (2006)—all of which either omit price entirely or include only the list price.

[13] The 2003 California medical care market is described in detail in Baumgarten (2004). Several previous papers describe the contractual arrangements between health plans and physicians in California, including Rosenthal et al. (2001) and Grumbach et al. (1998a,b).

[14] The insurers we do consider are Blue Cross, Blue Shield, Health Net, Pacificare, Aetna, and Cigna. Blue Cross of California is independent of other Blue plans, including Blue Shield of California.

network when estimating the choice equation.[15] While HMOs could, in theory, further influence hospital referrals for their enrollees by requiring patient cost-sharing for within-network hospitals, in practice this is not usually the case. There are two exceptions. One is the use of out-of-pocket payment schemes that require the consumer to share in hospital costs. Most out-of-pocket payments are fees (or copays) which would not affect hospital choice since they are fixed across hospitals and much smaller than hospital costs. However a minority of patients pay a co-insurance rate (i.e., a fixed percentage of hospital costs). Our data does not contain information on co-insurance rates, but we know from industry reports that they were used by only a small fraction of enrollees in 2003, and to the extent they were used, they were used disproportionately by the low capitation insurers in our dataset, and therefore do not explain our findings on the differences in price sensitivities across insurers.[16] The second exception is the possible existence of tiered networks. These are arrangements where insurers group hospitals into categories based on costs and patients pay more for services from higher-cost hospitals. They were first introduced in California in 2002–2003. We do not observe the details of any tiered networks and so cannot condition on them. However the available evidence indicates that only Blue Shield had more than 2 percent of its enrollees in a tiered network. This and the fact that Blue Shield is the only not-for-profit insurer in our data help explain the estimates below which, in turn, will induce us to treat Blue Shield differently in the rest of the paper.[17]

HMOs in California do not generally use hospital payment mechanisms that provide incentives either to control costs or improve quality. Most hospitals in California are paid by the insurance carrier on a per service or per diem basis.[18] Payment arrangements for physicians, in contrast, are often structured to generate cost-control incentives. Most HMOs contract on a non-exclusive basis with large physician groups,[19] making capitated (fixed) monthly payments to the group for every enrollee who uses it as his or her primary care clinic. The most common alternative is a fee-for-service payment arrangement. The extent of financial risk passed to the medical group varies across capitated contracts. In around 20 percent of cases the monthly ("global capitation") payment covers all services needed by the physician group's patients including inpatient hospital stays. These physician groups have a clear incentive to refer their patients to lower-cost hospitals. The

---

[15] Melnick and Ketcham (2008) find that 55 percent of hospitals were include in an HMO's network in California in 2003 on average. Ho (2006) finds a higher proportion: 83 percent of hospitals were included on average in a sample of 43 large markets (including 7 in California) in the same year.

[16] The Kaiser Family Foundation Employer Health Benefits Survey 2003 reports that 5 percent of covered workers in an HMO paid a co-insurance rate for hospital admissions while 49 percent paid a deductible or copay; the remaining enrollees paid neither. In contrast 14 percent of PPO enrollees paid a co-insurance rate, 26 percent paid a deductible or copay and the remainder paid neither. Our data contains only HMO enrollees for four out of six insurers. However we note below that for Blue Shield and Blue Cross only we observe admissions from PPO as well as HMO plans, since both are included in the Knox Keene plan for these two insurers. Blue Shield (a not-for-profit) and Blue Cross have the lowest capitation rates in our data. This use of patient cost-sharing is therefore negatively correlated with capitation, so it is likely to bias our estimates towards finding no difference in price coefficients between high- and low-capitation insurers. See Section III and online Appendix 1 for details.

[17] Pacificare, Health Net, and Blue Shield were the earliest adopters of tiered networks (in 2002) and only Blue Shield made the tiered product mandatory for any employers. See Yegian (2003) and Robinson (2003) for details.

[18] Hospital capitation payments existed but were uncommon in 2003. See online Appendix 1 for more details.

[19] There are two types of physician groups: medical groups and Independent Practice Associations (IPAs). See online Appendix 1 for details.

remaining 80 percent of capitation contracts involve payments that cover only the cost of services provided by physicians within the group, perhaps with the addition of ancillary services like outpatient medical tests. The HMO makes separate payments to hospitals for providing secondary care. However physician groups again have incentives to control hospital costs because "shared risk arrangements" almost always apply, under which a spending or utilization target is set and cost savings or overruns relative to the target are shared between the physician group and the HMO.[20] These arrangements are very similar to the "shared savings" arrangements that are often instituted for accountable care organizations.

Our dataset does not distinguish between global and nonglobal capitation arrangements. We investigate the extent to which referrals from physician groups with more highly capitated contracts of any kind are more (or less) sensitive to price. We also do not observe the physician or physician group referring each patient to a hospital. We do see the name of each patient's insurer and the percent of each insurer's primary services and other medical professional services that are capitated. In the analysis below we compare the importance of price in determining the hospital choice for patients enrolled in high-capitation insurers to its importance for those in low-capitation insurers.

Our finding that referrals from high-capitation insurers tend to be to less costly hospitals could be due to individual physicians referring capitated patients to cheaper hospitals and noncapitated patients to others (which is consistent with Melichar 2009), or to highly capitated physician groups using lower-priced hospitals on average. Physicians in more highly capitated physician groups could use lower-priced hospitals either because they respond to incentives they face in those groups, or because inherently cost-conscious doctors gravitate towards high-capitation groups. Our data does not allow us to distinguish between these alternative mechanisms, although it is clear that more highly capitated physician groups have increased incentives to make all their doctors aware of, and responsive to, monetary trade-offs.

III. The Dataset

A. *Data Sources and the Price Measure*

We use five datasets. The first is hospital discharge data covering all patient discharges from hospitals in California in the year 2003 from the state's Office of Statewide Planning and Development (OSHPD). This provides information on each patient's zip code, demographic characteristics, health insurer, the hospital chosen, and patient diagnosis details: both the "principal" diagnosis recorded as the major cause of admission and a list of up to 24 other diagnoses for each patient. We link this to OSHPD hospital financial data, to the OSHPD Birth Cohort file which contains outcome variables for both mother and infant, and to hospital characteristics data from the American Hospital Association for 2003. Finally we have access to the State of California Department of Managed Health Care Annual Financial Reporting Forms for 2003.

---

[20] Rosenthal et al. (2001) note that 85–90 percent of nonglobal capitation revenues were generated from contracts with shared hospital risk. Robinson and Casalino (2001) report similar findings. According to this literature, fee-for-service contracts at the time of our data did not generally involve shared hospital risk arrangements.

These include balance sheets, income statements, and some information on enrollment, utilization, and types of payment to providers for all Knox Keene plans (essentially the same as HMOs) in California. We consider only birth- and delivery-related admissions records and only private Knox Keene enrollees.[21] Our analysis covers only the six largest insurers other than Kaiser Permanente: these make up approximately 90 percent of the non-Kaiser observations in the data. We infer the hospital network of each insurer using the discharge data by assuming that a hospital is in the network if at least three patients are admitted from the particular insurer.[22] Consistent with Kessler and McClellan (2000), we assume that patients consider traveling up to 35 miles to visit a general hospital and up to 100 miles to visit a teaching hospital.[23]

We do not observe the price charged to the insurer by the hospital, and, as a result, have to construct our price variable. Our data does include the list price for every discharge. List prices are a standard set of prices listed by hospitals in each year for all their services. All patients are quoted the same list price for the same service. However, only uninsured patients and some patients using an out-of-network provider are actually asked to pay the list price, and even they are frequently offered a discount by the hospital. Each insurance company has a contract with each provider in its network that defines the price it actually pays for services for its enrollees. We also observe the average negotiated discount at the hospital level, calculated as the total contractual adjustments from private managed-care payors divided by the total charges (the sum of list prices for all inpatient and outpatient episodes) for the relevant hospital-year. Both variables are recorded in the hospital's financial statements.[24]

The price we need to construct is the price that the decision makers expect to pay for a patient entering the hospital in a given condition, defined below by a combination of diagnosis, age, and comorbidity information known at the time the patient is admitted to the hospital. We assume that expected prices are on average correct, and construct a baseline price variable as the average realized list price for a given condition in a particular hospital multiplied by 1 minus the average hospital discount. Since

---

[21] Knox Keene plans are defined as plans that are overseen by the California Department of Managed Health Care (DMHC) and subject to the Knox Keene Act. They are not precisely the same as HMOs: while most insurers designate only their HMO plans as Knox Keene plans, Blue Shield and Blue Cross PPO products were also included in this category in 2003. We cannot distinguish between PPO and HMO enrollees for these two insurers at the individual discharge level. Capitation rates are also reported for the full Knox Keene plan. This likely generates some of the cross-insurer variation in capitation rates in the data. To account for the differences between HMO and PPO plan types we drop hospitals to which very few patients are admitted for these two insurers (thereby removing from the data the hospitals that are most likely to be out-of-network). To the extent that we have not dropped all those PPO patients who pay a co-insurance rate rather than a fixed copay, we bias our estimates against finding a positive relationship between price-sensitivity and capitation rates. This is because Blue Shield and Blue Cross are the lowest capitation insurers in our dataset. See discussion in online Appendix 1 for details.

[22] We check the implied network definitions against hand-collected data (described in detail in Ho 2006) from seven California markets in 2003. The definition is conservative: that is, the networks implied by our methodology contain fewer hospitals than the networks in the hand-collected data and if an implied network contains a particular hospital it is also included in the hand-collected data in the vast majority of cases.

[23] We repeated the analysis using a 20-mile radius for low-risk pregnancies in Los Angeles, since a 35-mile radius could be too large in this urban area. The results were very similar to those reported below. Still we note that our market definitions may mask some remaining variation in the data.

[24] The fee-for-service arrangement is not the only way hospitals are reimbursed. Per diem and DRG or case-based payments are also possible: see online Appendix 1 for details. In principle we have the right average discount regardless of the form of reimbursement since contractual adjustments are defined as the sum of charges for the hospital's private managed care patients less its total revenue for those patients. However since the discount applies to all such patients and we only analyze Knox Keene inpatient labor and delivery admissions, there may be some error in it. In our main analysis we average over patients in different hospitals which, under assumptions to be specified below, addresses the measurement error issue.

the estimation methodology we rely on for our conclusion averages over agents, any remaining expectational error should average out when we sum across severities.[25]

We demonstrate below that there is meaningful variation in this price measure both across patients admitted with different conditions and across hospitals for a given condition. However, it is clearly subject to measurement problems. There is a trade-off between aggregation error, if our groups of similar patients for the expected list price calculation are defined too broadly, and measurement error if they are too narrow implying small sample problems. We return to this issue below. There may also be specification error since we observe the discount at the hospital rather than the hospital-insurer level.[26] We examine the robustness of our results to specification error in the price variable in Section VI and online Appendix 2. There we use additional data on the share of each hospital's total inpatient revenues coming from each insurer to estimate a model of the discount as a function of hospital, insurer, and market characteristics. We then repeat our inequalities analysis using price measures derived from the estimated discount, and find only minor differences in our results.

B. Summary Statistics

Tables 1 and 2 set out summary data on the six insurers included in the analysis; data for Kaiser are also included for comparison. These data give a broader picture of the insurers we consider than can be provided by our specific dataset. Since the effect of capitation payments on the price coefficient will be identified from variation across these six insurers, our goal here is to summarize the differences between them on other relevant dimensions. Table 1 provides enrollment data, showing that of the insurers we consider, Blue Cross and Blue Shield have the largest commercial plan enrollment while Aetna and Cigna have the smallest. Every insurer in our dataset has over 70 percent of its enrollment in commercial plans. The fourth column of Table 1 lists the number of delivery discharges included in our analysis for each plan; the breakdown is approximately proportionate to the commercial enrollment numbers. The first column of Table 2 lists the percent of each Knox Keene insurer's primary services that are capitated.[27] There is considerable dispersion across insurers, from Pacificare with 97 percent capitated payments to Blue Cross with 38 percent. The simple correlation between percent capitation and premiums is negative (around -0.44), consistent with more highly capitated insurers having lower hospital costs and therefore lower prices.[28] The rest of the table demonstrates that

---

[25] Gaynor and Vogt (2003) use a similar methodology, defining price as the observed list price multiplied by 1 minus the average discount.

[26] Specification error may also be generated because the observed discount is an average for both inpatient and outpatient services and for all managed care payors (including Point of Service plans) rather than just for Knox Keene inpatient events. An examination of the robustness of our results to this source of error is given in Section VI and online Appendix 2.

[27] Capitation payments for primary professional services are defined in the Knox Keene insurers' Annual Financial Statements as "capitation costs incurred by the reporting entity to primary care physicians, dentists and other professionals for the delivery of medical services." They include capitation payments to obstetricians. The statements also record capitation payments for other medical professional services, defined as including, for example, services provided by optometrists, nurses, ambulance drivers, and technicians.

[28] The simple correlations from our main dataset (described in detail below) also have the same qualitative features as our findings. The simple correlation between the weighted average distance traveled and insurer percent capitation (where the average is taken across severity groups and the weight is the number of Blue Cross patients in the relevant severity) is 0.57. The correlation between the weighted average price paid to hospitals and the percent

TABLE 1—ENROLLMENT DATA BY INSURER

	2002 enrollment			Birth discharges
	Com. plans	Medicare	Medi-Cal	
Aetna	485,787	37,312	0	6,291
Blue Cross	3,486,358	251,299	1,099,044	25,038
Blue Shield	2,231,350	67,049	0	16,302
Cigna	634,568	0	0	8,097
Health Net	1,665,221	101,317	349,826	16,950
Pacificare	1,543,000	386,076	0	15,479
Kaiser	5,790,348	671,858	104,844	N/A

Notes: Enrollment data on the six insurers in our analysis and on Kaiser Permanente. Source for 2002 enrollment: Baumgarten (2004). 2002 enrollment listed for commercial plans, Medicare plans and Medi-Cal/Healthy Families plans. "Birth discharges" is discharges in our data sample.

TABLE 2—SUMMARY STATISTICS BY INSURER

	Percent Prim Capitn	Tax status	Premium pmpm	Admin expense	Medical loss ratio	Inpatient utilizn		Prescrip drugs
						Discharges	Days	
Aetna	0.91	FP	152.42	19.33	86.2 percent	38.4	139.8	23.15
Blue Cross	0.38	FP	186.86	21.22	78.9 percent	38.4	142.4	20.92
Blue Shield	0.57	NFP	146.33	22.72	83.5 percent	50.3	176.4	20.51
Cigna	0.75	FP	-*	27.07	84.6 percent	39.8	137.1	15.63
Health Net	0.80	FP	184.92	18.60	86.3 percent	39.0	137.8	21.08
Pacificare	0.97	FP	149.92	24.51	88.4 percent	44.5	156.5	20.48
Kaiser		NFP	163.44	5.23	97.7 percent	49.1	158.1	0.44

Notes: Data on the six insurers in our analysis and on Kaiser Permanente. Source for all fields except percent capitation: Baumgarten (2004). "Percent Prim Capitn": percent of 2003 payments to primary providers made on capitated basis (source: State of California Department of Managed Health Care Annual Financial Reporting Forms). "Tax status" indicates for-profit (FP) or not-for-profit (NFP) status. "Premium pmpm": HMO commercial premium revenue per member per month (pmpm). "Admin expense": 2002 pmpm administrative expenses for entire insurer. "Medical loss ratio": 2002 medical and hospital expenses/premium revenues for entire insurer. Inpatient utilization and prescription drug data are for 2002 commercial plan. "Discharges" is 2002 discharges per 1,000 members; "Days" is acute days per 1,000 members; and "Prescrip drugs" is outpatient prescription drug expenses pmpm.*Information for Cigna not included in source data.

insurers with a high percent of capitated payments are not obviously different from other insurers on several observable dimensions. In particular, inpatient utilization and prescription drug costs are not highly correlated with capitation rates. That is to the extent that these variables reflect which types of patients choose which insurers, this suggests that patient selection into insurers is not related to capitation rates.

Blue Shield and Blue Cross, which have the lowest proportion of capitated payments, were historically different from other insurers. They were $501(c)(4)$ tax exempt as social welfare plans, acting as administrators of Medicare and providing coverage to state and federal government employees. By 2003, however, Blue Cross

---

capitation is -0.22. Both correlations are consistent with enrollees in higher-capitation insurers traveling further to access lower-priced hospitals. However patient locations and insurers' hospital networks, both of which are endogenous to the supply side of the market, will affect these correlations. The full model set out below conditions on these aspects of the environment and then provides a quantitative analysis of how the trade-off between price and distance is related to capitation.

and Blue Shield companies were franchisees, independent of the association and each other. They were no longer tax exempt and could be for-profit corporations. In California, Blue Cross was an investor-owned for-profit organization with a lower medical-loss ratio (defined as medical and hospital expenses divided by premium revenues for the whole insurer) and similar inpatient utilization to other insurers in the market. Blue Shield was still quite different from the other insurers we consider. It was the only not-for-profit company we analyzed and had relatively high inpatient utilization figures (although its premiums and medical loss ratio were quite low). As a result Blue Shield's administrators and physicians may have been less receptive to financial incentives than those of other insurers, an issue we return to below.

Table 3 provides summary statistics on the discharges in the dataset. There are 88,157 patients and 195 hospitals.[29] There are 38 hospitals in the average patient's choice set. Twenty-seven percent of discharges are from teaching hospitals. The average price paid (approximated as list price $\times$ (1 − average discount)) is \$4,317 for birth admissions. The average length of stay is 2.5 days. The importance of the distance between the patient's home and her hospital is clear from the raw data. The average distance between a patient and a hospital in her choice set is 24.6 miles; the average distance to the chosen hospital is 6.7 miles. Distance will be an important variable in the utility equation estimated below.

Table 3 also records means for three potential measures of outcomes: death while in hospital, transfer to an acute care setting (at this hospital or a different hospital), and transfer to a skilled nursing facility (again at either this or a different hospital). These are useful inputs to an initial investigation of the patterns in the data although we will not use them in our full model. The average probability of each event is low for delivery admissions: 0.01 percent for death, 0.3 percent for acute care transfer, and 1.5 percent for transfer to a skilled nursing facility.

Table 4 demonstrates that the variation in price and in outcomes across patient ages and comorbidities is intuitive. Here we add both data on infant outcome variables and data that follows both mother and baby over time which enables us to calculate the probability of readmission within a 12-month period.[30] We also aggregate the probabilities of death, acute care transfer, and skilled nursing facility transfer into a single probability of discharge to a location other than home.

Comparing older to younger women who give birth, we find that the delivery episodes of women who are aged over 40 are significantly more expensive and that these older women have higher probabilities of readmission within 12 months and of discharge to "other than home" than younger women. Also the infants of the older woman incur significantly larger hospital expenses and are significantly more likely

---

[29] This is the sample used for the logit analysis. We follow the previous literature by accounting for all hospitals in the choice set for that analysis. Average discount data is missing for some hospitals; we fill it in using regression analysis. We exclude these hospitals from the inequalities analysis because only pairwise comparisons between hospitals on which we have complete price information are required for the inequality estimation procedure. For similar reasons we also exclude the small number of hospitals reporting more than 5 percent capitated revenues. We are left with 64,691 patients and 157 hospitals. Using this smaller sample for the logits has no qualitative effect on the estimates.

[30] The data are taken from the OSHPD Birth Cohort File for 2003. All summary statistics are very similar to those of our main dataset.

TABLE 3—SUMMARY STATISTICS BY DISCHARGE

	Birth only	
	Mean	SD
Number of patients	88,157	
Number of hospitals	195	
Number of insurers	6	
Hospitals per patient choice set	38	
Teaching hospital	0.27	
Distance to all hospitals (miles)	24.6	25.6
Distance to chosen hospital	6.7	10.3
List price	$13,312	$13,213
Discounted price	$4,317	$4,596
Length of stay	2.54	2.39
Died (percent)	0.01	0.004
Acute transfer (percent)	0.3	0.02
Skilled nursing transfer (percent)	1.5	0.04

Notes: Summary statistics for dataset comprising private enrollees of the six largest HMOs excluding Kaiser who are admitted for delivery-related diagnoses. "Discounted price" is list price $\times$ (1-discount). "Died" is the probability of death while in hospital, "Acute transfer" the probability of transfer to an acute care setting (in this or a different hospital), and "Skilled Nursing Transfer" the probability of transfer to a skilled nursing facility (again at this or a different hospital). Standard Deviation for "Died," "Acute transfer," and "Skilled nursing transfer" are calculated under the assumption that the 0/1 variable is binomially distributed.

TABLE 4—PRICES AND OUTCOMES BY PATIENT TYPE

		Mother			Infant		
	Obs.	Price$\times(1-$disc$)$	Readmission	Not home	Price$\times(1-$disc$)$	Readmission	Not home
Overall	73,118	4,291 (4,373)	2.39% (0.06%)	1.62% (0.05%)	2,675 (18,324)	9.42% (0.1%)	6.60% (0.1%)
Age							
< 40	71,074	4,259 (4,329)	2.36% (0.1%)	1.60% (0.1%)	2,627 (18,103)	9.41% (0.1%)	6.50% (0.1%)
> 40	2,044	5,420 (5,571)	3.52% (0.4%)	2.10% (0.3%)	4,337 (24,787)	9.64% (0.6%)	9.88% (0.7%)
Signif diff		0.000	0.000	0.038	0.000	0.365	0.000
Charlson							
0	71,804	4,256 (4,265)	2.33% (0.1%)	1.58% (0.1%)	2,619 (18,101)	9.36% (0.1%)	6.53% (0.1%)
> 0	1,314	6,227 (8,135)	5.78% (0.6%)	3.42% (0.5%)	5,735 (27,745)	12.3% (0.9%)	10.5% (0.9%)
Signif diff		0.000	0.000	0.000	0.000	0.000	0.000

Notes: Data taken from OSHPD Birth Cohort 2003 (a slightly different dataset that includes infant outcome variables). "Readmission" is percent of patients readmitted to hospital within 12 months of birth episode. "Not home" is percent of patients discharged somewhere other than home; this includes transfer to acute care setting, transfer to skilled nursing facility, discharge against medical advice, and death. Standard deviations in parentheses; for "Readmission" and "Not Home" we report standard errors which are calculated assuming that the 0/1 variables are binomially distributed. Charlson scores assign weights to comorbidities (known on admission to hospital) other than principal diagnosis where higher weight indicates higher severity. Value 0–6 are observed in the data. "Signif diff" states significance level at which we cannot reject the hypothesis that the means in the two samples are the same; these are the results of a *t*-test for price $\times$ (1 − discount) and a *z*-test assuming two binomial distributions for "Readmission" and "Not Home."

to be discharged to "other than home." Interestingly, infant readmission probabilities are not significantly different across these two groups.[31]

We use the Charlson score (Charlson et al. 1987) as a measure of patient severity: this assigns integer-valued weights (from 0 to 6) to comorbidities other than

---

[31] Our price sensitivity analysis includes only the expenses charged to the mother. Infant health, and therefore expected expenses, are generally not known when the hospital is chosen and therefore are not relevant for this part of the analysis. However hospital quality for the infant as well as the mother should be an input to the hospital choice. So we consider both infant and maternal outcomes in Section VII where we ask whether health outcomes differ by insurer (conditional on severity).

principal diagnosis where higher weights indicate higher severity. The weights are summed to generate a single integer-valued index. For example, patients with comorbidities indicating that they have diabetes or mild liver disease would receive a Charlson score of 1; those with renal disease or any malignancy would have a Charlson score of 2; those with a metastatic solid tumor or AIDS would have a Charlson score of 6. A patient with both diabetes and renal disease would have a score of 3. The index was developed by physicians and is widely used to measure severity based on diagnoses listed in patient records. Table 4 indicates that women with higher Charlson scores in our data, and their infants, had more costly deliveries and higher probabilities of adverse outcomes than women with lower Charlson scores. All of these differences are significant at $p = 0.05$. Our main analysis will allow the Charlson score, interacted with other severity measures such as age and principal diagnosis, to affect preferences for different hospitals.

IV. The Model

As noted, the hospital chosen is a result of a complex decision process. The woman first chooses an obstetrician, typically with knowledge of which hospitals the obstetrician can admit patients to, and then the obstetrician, in consultation with the patient, chooses among the hospitals where he (or she) has admitting privileges. We assume this process generates an ordering of the hospitals in the insurer's network which is derived from the patient's and doctor's preferences. The patient's preferences are affected by the distance from her home to the hospital, her assessment of the severity of her condition, and by the hospital's characteristics (some of which we, the researchers, may not have measures of). The physician's choice is influenced by the patient's preferences, their assessment of the severity of the patient's condition, and the quality of the hospital services for that severity, and the price charged by the hospital to the insurer. The potentially observable part of the hospital referral function whose maximum determines the hospital (h) that patient i of insurer π is allocated to, is assumed to take the additively separable form

$$(1) \qquad W_{i,\pi,h} = \theta_{p,\pi}\, p(c_i, h, \pi) + g_\pi(q_h(s), s_i) + f_\pi(d(l_i, l_h)),$$

where
- $p(c_i, h, \pi)$ is the price insurer π is expected to pay at hospital h for a patient who enters in condition (which includes diagnosis) c_i;
- s_i is a measure of the severity of the patient's condition;
- $q_h(s)$ is a vector of perceived qualities of hospital h, one for each different severity;
- $g_\pi(q_h(s), s_i)$ is the plan- and severity-specific function which determines the impact of the quality of a hospital for the given severity on the choice of hospitals; and
- l_i is patient i's location, l_h is hospital h's location, $d(\cdot)$ provides the distance between the two locations, and $f_\pi(\cdot)$ is an increasing function of that distance which may differ by plan, for $i = 1, \ldots, n_{\pi,s}$, $h = 1, \ldots, H$, and $\pi = 1, \ldots, \Pi$.

The condition of the patient on admission (our c_i) will determine the expected price associated with that admission at different hospitals and therefore will sometimes be referred to as the price group. The severity groups (s_i) are aggregates of the admission conditions c_i (so there is variance in price conditional on severity). Both these terms are defined based on information in the discharge data. Their precise definitions differ somewhat across estimation strategies and will be defined below.[32] Also when implementing these alternative estimation strategies we will have to add an unobserved (or disturbance) term whose properties will be assumed to differ across those strategies.

In its most general form, which will be our preferred form, the function $g_\pi(\cdot)$ is allowed to differ arbitrarily: across plans, among sickness levels for a given hospital, and across hospitals. This allows particular hospitals to have higher quality for some diagnoses and/or sickness levels than for others. So different aspects of a hospital's characteristics can be differentially important for hospital choice for different severities.[33] It also permits physicians to differ in their intensity of preferences for quality relative to price and distance when considering patients of different sickness levels, and allows the referral process in different insurers to differ in both the quality assessment of hospitals and the weight given to quality relative to price and distance. Finally keep in mind that the ordering implicit in equation (1) is an outcome of a decision-making protocol that involves both the physicians' and the patients' preferences, so the analysis allows for patient (as well as physician) differences across plans.

For some of the specifications we will have to constrain $g_\pi(\cdot)$ to be a parametric function of patient and hospital characteristics. To the extent that the parametric assumption does not capture all the variance in $g_\pi(\cdot)$ the residual variance will create an additional unobservable that may bias the other parameters of interest. In particular if the "unobserved quality" represented by this residual is correlated with price we would expect it to cause a positive bias in the price coefficient. This is the reason that our preferred mode of analysis starts by not constraining $g_\pi(\cdot)$ in any way. Only after estimating the price coefficients for each plan do we come back to asking whether there are constraints that the hospital-severity plan-specific quality estimates seem to satisfy.

V. Logit Analysis

We begin with a multinomial logit model of hospital choice, as it provides a familiar starting point for investigating the patterns in the data. The logit model makes the following assumptions:

$$(2) \qquad\qquad p(c_i, h, \pi) \;=\; \delta_h^o \, lp^o(c_i, h),$$

---

[32] In the main analysis, which is the inequality analysis in Section VI, severity groups (our s_i) are defined by unique combinations of age, principal diagnosis, and other factors observed in the discharge data including the severity rank (defined on a scale from 1 to 3 by obstetrical experts) of the most serious comorbidity in the patient's discharge record. Price groups (our c_i) are severity groups subdivided by the number of most seriously ranked comorbidities. See Section VI for details.

[33] For example, Goldman and Romley (2008) find evidence that amenities such as food quality and staff attentiveness play an important role in hospital demand, and their relative importance in the hospital choice equation may differ by severity of illness.

where $lp^o(c_i, h)$ is the average list price of patients who enter hospital h with condition c_i, and δ_h^o is 1 minus the average discount rate hospital h gives managed care providers,

$$(3) \qquad\qquad g_\pi(q_h(s), s_i) = q_h + \beta z_h x(s_i),$$

where q_h are hospital fixed effects, $x(s_i)$ are functions of the sickness level of the patient, and z_h are hospital characteristics, both of which are specified below, and

$$(4) \qquad\qquad f_\pi(d(l_i, l_h)) = \theta_{d1} d(l_i, l_h) + \theta_{d2} d(l_i, l_h)^2.$$

Adding the disturbance $\varepsilon_{i,\pi,h}$, and substituting into equation (1) the logit model becomes

$$(5)\ \ W^l_{i,\pi,h} = \theta_{p,\pi} p(c_i, h, \pi) + g_\pi(q_h(s), s_i) + \theta_{d1} d(l_i, l_h) + \theta_{d2} d(l_i, l_h)^2 + \varepsilon_{i,\pi,h}.$$

The properties of the model are completed by assuming our composite agent knows $\varepsilon_{i,\pi,h}$ at the time the hospital choice is made, and the vector of disturbances has a distribution, conditional on the other right hand side variables, which is i.i.d. Type 1 extreme value. Notice that since we have not indexed an "outside" option, the logit analysis conditions on women who only consider giving birth at a hospital. We estimate this model using maximum likelihood.

We consider three different assumptions regarding the price coefficient $\theta_{p,\pi}$, in the estimation

$$(6) \qquad\qquad\qquad\qquad (a)\ \theta_{p,\pi} = \theta_p;$$

$$(b)\ \theta_{p,\pi} = \theta_{p,\pi};$$

$$(c)\ \theta_{p,\pi} = \theta_0 + \theta_1 . pcap_\pi ,$$

where $pcap_\pi$ is the insurer's capitation rate.

Equation (2) states that the price is exactly equal to our measure of the expected list price for the patient's diagnosis multiplied by 1 minus the observed average discount: i.e., this specification assumes no measurement or expectational error in price. We define the expected list price to be the average list price for the particular hospital over patients in categories defined by unique combinations of: age (categories 11–19, 20–39, 40–49, and 50–64); principal diagnosis (21 categories for women giving birth including, for example, "normal delivery," "previous Cesarean Section," and "early labor"); Charlson score, and diagnosis generating the Charlson score. Both principal diagnosis and Charlson score are based only on diagnoses known on admission. We are constrained to using these fairly broad definitions of similar patients because we encounter small sample problems when we define narrower groups. To the extent that either the aggregation generates measurement error in our price measure, and/or the small cell sizes generate estimation error in our estimate of expected price, we expect an attenuation bias in the estimated price coefficient.

Equation (3) restricts the $g_\pi(\cdot)$ term in a way consistent with the previous literature: we assume it is determined by a hospital fixed effect plus interactions between hospital characteristics and patient characteristics that are known on admission and expected to be correlated with severity. In the inequalities analysis below we define over 100 patient severity groups and allow these to freely interact with hospital fixed effects. We can not do this in the logit analysis because it would imply estimating almost 20,000 coefficients and a similar number of expected price terms (without error), putting us in a range of values where an incidental parameter problem, similar to that described in Neyman and Scott (1948), would make coefficient estimates very unreliable. So we assume the interaction terms are determined by linear interactions between hospital and consumer diagnostic characteristics. Included in z_h are the number of nurses per bed and indicators for teaching hospitals, for-profit hospitals and hospitals that offer transplant services (a proxy for high-tech hospitals). We also include a measure of the quality of delivery and birth services taken from Ho (2006): hospitals were rated on a scale from 0 to 1, where 0 indicated that no delivery/birth services were provided and a higher rating indicated that a less common (assumed to be higher-tech) service was offered. The patient characteristics in x_i are the expected probabilities of death in hospital and of transfer to acute care setting or skilled nursing facility given the patient's age group, principal diagnosis, and Charlson score.

While these interactions, like those used in the previous literature, are sensible given the constraints imposed by the methodology, we do not expect them to be sufficient to fully address the price endogeneity issue noted above. To the extent they do not there will be an error in the approximation in equation (3) so that equation becomes

$$g_\pi(q_h(s), s_i) = q_h + \beta z_h x(s_i) + e_\pi(q_h(s), s_i).$$

The logits assume $e_\pi(q_h(s), s_i) \equiv 0$. If this is not the case and hospital quality is both regarded as more important for more severely ill women and is positively correlated with hospital price, we expect this error to bias the price coefficient upwards. Note, however, that if we restrict the analysis to a subsample of the data which only contains women with similar severities (i.e., for whom $s_i \approx s^*$), then when estimating the logit on this subsample the effect of severity is almost captured by the hospital fixed effect (it would be captured if severity was exactly the same for all women in the subsample). So a separate analysis for women with similar severities should mitigate this source of omitted variable bias.

The logit analysis assumes that both the distance coefficient and the quality-severity interaction terms do not vary across insurers. We begin the analysis of all the subsamples we consider by assuming a common price coefficient across insurers, and then provide separate results where this coefficient differs across insurers, paying particular attention to the relationship of the price coefficient to the insurer's capitation rate. After presenting the results from the full sample, we provide another set of results which partially controls for the omitted variation in patient-severity hospital-quality interactions by restricting our attention to the least sick patients in our data. For comparison we also estimate separately on the rest of the sample, a group with a more diverse set of severity conditions. As noted we do not have enough data to obtain meaningful results if we allow for hospital-severity

TABLE 5—LOGIT ANALYSIS RESULTS

	All birth patients	Least sick patients			Sickest patients		
Price	0.010** (0.002)	−0.017* (0.009)		0.069** (0.014)	0.012** (0.002)		0.028** (0.006)
Price interactions:							
Percent capitated				−0.127** (0.016)			−0.025** (0.008)
Pacificare			−0.077** (0.012)			−0.006 (0.006)	
Aetna			−0.011 (0.016)			0.021** (0.008)	
Health Net			−0.038** (0.011)			0.007 (0.005)	
Cigna			−0.021 (0.014)			0.004 (0.007)	
Blue Shield			0.018 (0.011)			0.024** (0.004)	
Blue Cross			0.008 (0.011)			0.014** (0.003)	
Distance	−0.215** (0.001)	−0.215** (0.002)	−0.215** (0.002)	−0.215** (0.002)	−0.217** (0.002)	−0.216** (0.002)	−0.216** (0.002)
Distance squared	0.001** (0.000)	0.001** (0.000)	0.001** (0.000)	0.001** (0.000)	0.001** (0.000)	0.001** (0.000)	0.001** (0.000)
$z_h x_i$ controls (15 coeffts)	Y	Y	Y	Y	Y	Y	Y
Hospital FEs (194 coeffts)	Y	Y	Y	Y	Y	Y	Y
Observations	88,157	43,742	43,742	43,742	44,059	44,059	44,059

Notes: Least sick: aged 20–39 with zero Charlson scores and "routine" principal diagnoses and comorbidities (see online Appendix 1). Sickest: all other patients. $z_h x_i$: interactions between hospital characteristics (teaching hospitals, for profit hospitals and hospitals that offer transplants, number of nurses per bed, a variable summarizing quality of delivery services provided) and patient characteristics (probabilities of death while in hospital, transfer to an acute care facility, and transfer to a skilled nursing facility conditional on principal diagnosis, age category and Charlson score). Sample sizes for sickest and least sick populations do not sum to total observations because a small number of hospitals with very few admissions were dropped when estimating on each subsample; patients choosing these hospitals were also dropped.

**Significant at the 5 percent level.
*Significant at the 10 percent level.

specific fixed effects, but we can push the subsampling a bit further by looking at the results from subsamples that have only the largest six severity groups in both less-sick and sicker subsamples of the dataset.

A. *Logit Results*

A summary of the results is reported in Table 5. The price coefficients, price interaction terms, and distance coefficients are reported, together with the sample size, for each specification. In each case the distance coefficient is negative and highly significant, with a magnitude that is consistent with estimates from the previous literature.[34]

---

[34] See, for example, Gaynor and Vogt (2003) and Ho (2006).

The price coefficient from the full sample of delivery/birth discharges is positive and significant with a t-value of approximately 5. Recall that we would expect a positive bias in that coefficient if high-priced hospitals were high-quality hospitals, and the severity-hospital interactions we included were not sufficient to control for hospital quality conditional on the severities that determined hospital choice. To see if this might be the source of a problem, the next column provides the results from the same specification when we restrict the sample to the least sick women. We identified these women with the help of obstetrical experts at Columbia Presbyterian hospital. They are the subset of patients who are aged 20–39, have a Charlson score of 0, and whose principal diagnosis and comorbidities were defined by the obstetricians to be "routine." Our sample contains 43,742 of these patients. Since they have similar severities we would expect there to be less variance across patients in the importance of hospital quality differences, and this should mitigate the omitted variable bias.

When we use the sample of less-sick women, the price coefficient becomes negative (magnitude -0.017) and marginally significant (standard error 0.009). The same specification on the subsample with the sickest patients, the group of patients where we think the quality severity interactions are likely to be both more variant and more important in determining hospital choice, yields a positive price coefficient again (of 0.012), and this time with a t-value that increases to 6 (despite the fact that this subsample is only about one-half the size of the full sample). We conclude that we need a better way to control for hospital quality/patient severity interactions.

Next we look for interactions between price and insurer fixed effects. Insurers in the table are sorted by declining proportion of capitated payments to primary physicians. When we use the sample with the least sick patients, Blue Cross and Blue Shield, the plans which have the lowest proportions of capitated, payments have small, positive, and insignificant price coefficients. All four of the remaining insurers have price coefficients less than zero. The negative price coefficients are significant for Pacificare and Health Net, two of the three carriers that favor capitation the most (97 percent of payments for Pacificare and 80 percent for Health Net). The remaining carriers, Aetna and Cigna, have relatively small sample sizes (6,291 and 8,097 birth discharges respectively, compared to 15,479 for Pacificare and 16,950 for Health Net), which helps explain the larger standard errors on their price coefficients. When we remove the price-insurer interactions and instead include an interaction between price and the percent capitation in the insurer, the price coefficient is positive and the interaction term negative with almost twice the magnitude of the price coefficient. Both are highly significant; the t-value of the capitation interaction is 7.7.

When we do the same exercise with the subsample of sicker patients, the price-insurer interaction term is still negative for Pacificare, although insignificant at $p = 0.05$ and smaller in magnitude than for the healthier population. All other insurers' price coefficients are positive and three out of five are statistically significant, again pointing to the need for a better way to control for hospital quality/patient severity interactions. The third specification, including a price-percent capitation interaction, again generates a positive price coefficient and a negative interaction term (implying that insurers that favor capitated payments generate physician referrals that are more price-based than those of other physicians). However, the magnitudes are much smaller than for the least sick sample and the implied overall price coefficient is positive even for insurers with 100 percent capitated payments to primary physicians.

Online Appendix 3 takes the logit analysis one step further. We now restrict our attention to the six largest severities in either the less-sick or the sicker subsample of patients. The incidental parameters problem prevents us going further than this: indeed we have to use only the four largest markets, in which the number of patients per hospital-severity is relatively large, to obtain these estimates.[35] The first column for each subsample (sick and less-sick) sets out the estimates in a specification identical to that in column 1 of Table 5 except that we remove the squared distance term, since this makes the results more comparable to the inequalities analysis, and include only the four largest markets. As in Table 5 the price coefficient for the less-sick sample is negative and insignificant while that for the sicker sample is positive and very significant. We also report results that include an interaction between price and the percent capitation of the insurer. As before the interaction term is negative and significant in both samples, but for the less-sick sample it is much larger than the price coefficient (price coefficient 0.10, price-capitation interaction -0.17) while for the sicker sample the two coefficients are essentially the same magnitude (0.03 and -0.03 respectively).

The next set of results for each subsample drops all but the six largest severities. This has very little impact on the less-sick sample (we lose only 10 percent of observations) but a marked effect on the sick sample (we lose 65 percent of the original sample). Accordingly, and because the less-sick sample is already fairly homogeneous in terms of severity, restricting to the six largest less-sick severities has very little effect on the estimates. However in the sicker sample the price coefficient moves from positive and significant to negative (although still insignificant). When we go further with these subsamples and allow the price coefficient to differ across insurers by including an interaction between price and percent capitation, the price-capitation interaction for the sicker sample triples in value when we include only the six largest severities, from -0.034 (0.010) to -0.092 (0.038), and is now quite a bit larger than the price coefficient which is 0.059 (0.033).

There were too few patients per hospital-severity pair for us to restrict the logit sample to a smaller number of severities.[36] However we did use a related procedure to investigate whether measurement error in price was likely to also pose a problem for logit estimators. In particular we repeated the analysis of the six largest severities using only the hospital-severity pairs with more than 30 observations. Since the variance of the measurement error should be inversely proportional to the number of patients, dropping the small hospital-severity pairs ought to mean dropping those pairs with the most problematic price measures. Restricting the subsamples for the six largest severities in this way excluded about 15 percent of the less-sick subsample, but about one-half of the sicker subsample. For both subsamples the uninteracted price coefficient changed very little, but once we allowed insurers to differ in their price response by including a price-percent capitation interaction, in both subsamples the price coefficient became more positive and the price-capitation

[35] The four largest markets are Los Angeles, Orange, San Diego, and the Bay Area. We use the same severity definitions as in the inequalities analysis in Section VI.

[36] More formally for consistency in this specification we require the number of observations for each hospital to grow large and this is not a reasonable approximation for samples based on a single severity group. In addition, since there have to be multiple price groups in each analysis, use of the separate logit fixed effect analysis for each severity group would require us to use more detailed price groups than in the logits reported here and therefore accentuate the price measurement problem.

interaction more negative and both were highly significant. Most noticeably the price-capitation interaction for the sicker population tripled in value once again and the price effects in the sick and less-sick samples became similar: the less-sick sample had a price coefficient 0.162 (0.028), and a price-capitation interaction −0.250 (0.025), while the sicker sample had a price coefficient 0.183 (0.073), and a price-capitation interaction −0.264 (0.053). By now, however, we had dropped over 65 percent of our original sample and still had reason to believe we had not fully eliminated either the omitted variables or the errors in variables bias.

Before leaving the logit analysis we go back to the full sample of the least sick patients and look for the implications of the logit results on the relative magnitudes of the distance and price effects. Consider first the distance coefficient. We calculate the impact of a one-mile increase in distance for hospital h, holding all else fixed, on the probability that a particular patient i visits that hospital. We then take the average over patients and a weighted average over hospitals. The average effect of the one-mile distance increase is a 13.7 percent reduction in the probability that the hospital is chosen.[37] Next we conduct a similar exercise to evaluate the magnitude of the price effect. Consider Pacificare's referrals for its least sick patients; the insurer with the most negative estimated price coefficient. The implied average effect of a $1,000 increase in a hospital's price, holding all other prices constant, is a 5.2 percent reduction in the probability that the hospital is chosen. So the price increase we would require in order to compensate for a one mile increase in distance would be approximately $2,600. This is more than two-thirds the average price for the less-sick patients (which is $3,380 and has a standard deviation of $1,870.) All the other insurers' price coefficients are considerably less negative, implying a considerably larger price distance trade-off. These numbers, together with the results we obtain when we use subsamples, accentuate our worry that omitted variables and errors in observed variables may be causing important biases in the logit estimators.

VI. Inequalities

As noted, we are worried that the logit analysis does not fully control for variation in quality conditional on severity and that this might cause a positive bias in the price coefficient. In addition, that analysis compels us to use average prices within quite broadly defined patient groups because narrower groups would increase the variance in our estimated price accentuating the impact of measurement error in price. The estimates in Table 5 and online Appendix 3 show that restricting the sample to relatively similar severity groups, and dropping hospital-severity pairs where measurement error is likely to be large, has the expected effects on the estimates. However we cannot fully address either omitted variable bias or measurement error in a multinomial logit framework. We now provide an alternative estimation technique that addresses both issues.[38]

---

[37] The average distance to the chosen hospital for the less-sick patients included in the sample is 6.45 miles; the standard deviation is 10.11 miles. The weighted average probability that a particular hospital is chosen is 2.7 percent, where the weight is the number of discharges.

[38] In addition to the logits reported above we also tried Chamberlain's (1980) conditional likelihood estimator. At least under particular distributional assumptions this addresses the absence of quality controls, though not the measurement error in price. However the conditional likelihood estimator is not suitable for our problem because

The method is based on a revealed preference inequality: it is assumed that the chosen hospital is preferred to feasible alternative hospitals. Consider again the referral function in equation (1) where patients are assigned to detailed severity groups s_i (the same groups used to define samples in online Appendix 3) and price groups c_i. We consider all couples of same-insurer, same-severity patients whose chosen hospitals differ but both of whose choices were feasible for both agents. Within each couple we sum the inequalities obtained from the fact that each patient's choice is preferred to the choice made for the other. Since the severity-hospital interactions from the two inequalities are equal but opposite in sign, when we sum the inequalities the interaction terms difference out. Revealed preference implies that this sum is positive, and this constrains the remaining parameters.

More formally, let $S_\pi(h, h', s)$ be the set of patients from plan π with severity s who chose hospital h but had hospital h' in their choice set. For notational simplicity let $\Delta x(i, h, h') = x_{i,h} - x_{i,h'}$ for $x = W(\cdot)$ or price, and let $\Delta f_\pi(l_i, l_h, l_{h'}) = f_\pi(d(l_i, l_h)) - f_\pi(d(l_i, l_{h'}))$ provide the difference between the distance from patient $i's$ home to hospitals h and h'. Then the observable part of our inequalities is formed by taking couples of patients $i \in S_\pi(h, h', s)$ and $i' \in S_\pi(h', h, s)$ and using equation (1) to form

$$(7) \quad \Delta W(i, h, h') + \Delta W(i', h', h)$$

$$= \theta_{p,\pi}[\Delta p(c_i, h, h') + \Delta p(c_{i'}, h', h)] + [\Delta f_\pi(l_i, l_h, l_{h'}) + \Delta f_\pi(l_{i'}, l_{h'}, l_h)].$$

Revealed preference implies that this equation is expected to be greater than zero when we evaluate it at the true value of θ and the expected prices. We then average over all couples $i \in S_\pi(h, h', s)$ and $i' \in S_\pi(h', h, s)$ for all $h' \neq h$, and this averages out the expectational and measurement errors in price.

Since we have removed the quality/severity interaction terms and no longer need to estimate their coefficients we can define our severity groups at a more detailed level than was possible in the logit analysis. Moreover since we average over all such couples in all severity groups we eliminate the effect of estimation error in price so we can define the price terms in as narrow a set of price groupings as we like. Note, however, that this procedure does rely on the expected price varying within a hospital across patients who have the same severity of illness; otherwise the price terms will be differenced out along with the interaction terms.

A. Severity and Price Groups

Our severity groups are assumed to be defined in sufficient detail that the severity-hospital interactions absorb all unobserved variation, other than price and distance, that affects choices and may be correlated with price. That is though we require price variation across patients in different price groups within a given severity, that variance

---

its computational burden grows as the combinatorial formula for the number of ways the total patients in a severity group can be divided among the hospitals conditional on the given number of patients (in the severity group) and hospitals. The number of patients and hospitals in our study is just too large to make this feasible.

cannot affect choices except through the price variable itself. We now provide details of our severity and price definitions and consider whether these requirements are satisfied. Our definitions are chosen following the advice of the obstetricians we consulted at Columbia Presbyterian Hospital. As one input to the definitions, these experts assessed the list of principal diagnoses and comorbidities in our data, assigning each a rank from 1 to 3. A "1" indicated a routine diagnosis (such as normal birth or immunization of the newborn) and "3" indicated something more serious; see online Appendix 4 for a complete list.

We use narrower definitions for severity and price than were used in the logit analysis. Severity groups are now defined by a unique combination of age, principal diagnosis, Charlson score, diagnosis generating the Charlson score, and the rank of the most serious comorbidity, other than principal diagnosis, that is listed in the discharge record.[39] Our expected list price, on the other hand, is now defined as the average list price for the particular hospital across women with the same severity (as just defined) who also have the same number of most seriously ranked comorbidities.[40] That list price is interacted with 1 minus the average hospital discount to calculate our baseline price variable.[41] These definitions generate many more groups than those used in the logit analysis. For example, for the first insurer in our data, there are 63 populated groups defining prices using the logit-based categories (recall that the logits did not allow severity groups to interact with the hospital fixed effects). There are 106 severities and 272 price groups under the more detailed definitions.

The obstetrical experts we interviewed advised us that these detailed price groups, conditional on severity, were unlikely to be important in terms of hospital choice. The price groupings are more detailed than those used for severity only in that they break out patients by the number of comorbidities of the highest rank as well as the identity of that rank. The number of comorbidities of a given rank, conditional on severity, was considered unlikely to affect the hospital choice. While a physician might refer a pregnant woman with a normal delivery but a comorbidity of rank 2 (such as a viral infection or a thyroid disorder) to a different hospital from a similar patient with only rank-1 comorbidities, this would be a hospital well-equipped to deal with high-risk pregnancies rather than the specific comorbidity, and the presence of two rather than one rank-2 comorbidities would not affect the referral decision. In contrast, our experts agreed that the number of comorbidities of a particular rank would be likely to affect the tests performed and drugs prescribed and therefore the price.

The price variation used in the inequality analysis is a difference in price differences. More precisely it is a difference in expected price differences, where an

---

[39] For example, a woman aged 25 with a normal delivery, Charlson score of 1 caused by diabetes, and a maximum rank of 2 would be in a different severity group from a woman with the same age, principal diagnosis, and maximum rank but whose Charlson score of 1 was caused by mild liver disease. A woman aged 25 with a normal delivery, Charlson score of 0, and maximum rank of 1 would have a different severity group from a similar woman whose maximum rank was 2 (but where that comorbidity was not severe enough to trigger a Charlson score above 0).

[40] For example if two women have the same age and principal diagnosis and a zero Charlson score but one has a migraine (a rank 1 comorbidity) and one has a viral infection (rank 2), the women have different severities and different prices. If neither woman had a migraine but one had a viral infection and the other had a viral infection and also a thyroid disorder (both rank 2 comorbidities), they would have the same severity but different prices.

[41] Provided expectations are unbiased the average of actual prices will converge to the average of expected prices, so we could have used actual and not a measure of expected prices in our inequalities. However the expected price variable we do use has the advantage that it uses the information from all same-plan same-price group patients in comparisons across couples of hospitals, not just those for whom there was a feasible switch.

TABLE 6—PRICE VARIANCE ACROSS AGGREGATED PRICE AND SEVERITY GROUPS

Number diags of max rank	Pats	Price ($)	SD
Panel A. Max rank 1			
1	23,029	3,431 (15)	1,612
2	11,757	4,145 (28)	2,180
3	4,077	4,682 (60)	2,356
4	1,179	5,505 (149)	2,590
5	331	6,189 (254)	3,123
≥ 5	95	7,663 (936)	4,896
Total	40,468	3,857 (15)	
Panel B. Max rank 2			
1	13,128	4,968 (42)	2,476
2	4,196	6,019 (88)	2,785
3	1,274	7,428 (212)	3,609
4	380	8,602 (462)	5,283
5	110	10,186 (1,002)	6,084
≥ 5	55	13,365 (1,596)	8,880
Total	19,143	5,488 (40)	
Panel C. Max rank 3			
1	1,273	7,448 (356)	4,256
2	64	11,536 (2,337)	20,370
3	8	12,733 (4,009)	11,338
4	1	25,573 (—)	—
5	0	—	
≥ 5	0	—	
Total	1,346	7,687 (13,065)	

Notes: Distribution of patients from inequalities sample who have a Charlson score of 0 across comorbidity ranks. "Pats" shows the number of patients in each "Max rank" group and each "number diags of max rank" group. Here "Max rank *j*" means the maximum rank of a comorbidity for this patient, as defined by obstetrical experts at Columbia Presbyterian Hospital, is *j*. "Number diags of max rank" groups patients according to the number of comorbidities in their discharge record with the relevant max rank. Patients in different rows of a particular column of the table within a panel will have different price groups. "Price ($)" is the average observed price × (1-discount) for patients in this group; standard errors in parentheses. SD is the cross-hospital standard deviation of the mean observed price × (1-discount) in this hospital for the patients in this group.

expected price is measured by average prices associated with a given admission condition (defined by a combination of diagnosis, age, and comorbidities) at a hospital. Take patient i who goes to hospital h and i' who goes to hospital h'. Assume they both have access to each other's chosen hospital and both have the same severity and plan. Then when we add the inequalities for the two of them the severity-specific hospital quality terms difference out. Patient i who goes to h faced an expected price difference of $p(c_i, h) - p(c_i, h')$ while patient i' who goes to h' faced $p(c_{i'}, h') - p(c_{i'}, h)$. These are the only price comparisons across hospitals that are used in the analysis and they are comparisons for exactly the same admission price group. The estimates of the price coefficient come from finding out just how much further the patient with the highest price differences is willing to travel than the patient with the lowest price difference.[42]

---

[42] The intuition of the method is somewhat similar to including fixed effects at the severity-hospital level in a linear regression. However the regression approach would not average out the error in the price measure and therefore would have an errors-in-variables problem.

Table 6 provides summary statistics on the price differences used in the analysis. The data used in the table are from patients who have a Charlson score of 0 and a given maximum rank but different numbers of most seriously ranked comorbidities. 63,450 out of 64,691 patients in our inequalities sample have a Charlson score of 0.[43] Cells in the table correspond to a particular maximum rank and number of diagnoses of that rank, but average over 4 age groups and 21 diagnoses which are disaggregated in the actual analysis.[44] The difference across hospitals for a patient who is admitted in a given price group is quantified in the columns labelled SD which give the standard deviation for that price group across hospitals (averaged over age and diagnosis). There clearly are differences across hospitals in a given price group and these differences increase as we move down the rows of the table within a panel (i.e., as we increase the number of comorbidities of maximum rank). The second difference is across price groups within a severity group (defined by diagnosis, age, and the maximum rank of the comorbidities). This is illustrated by the differences across rows in the columns labelled Price ($) within each panel of Table 6 (again averaged over age and diagnosis). Clearly the mean prices are ordered as expected, and the differences are usually highly significant (the bracketed numbers provide the standard error of the mean).

Finally we note that when we go to the actual disaggregated groups (disaggregated further by age and diagnosis), an analysis of variance indicates that moving from severity to price groups explains an additional 12 percent of the variance in price (from 50 percent to 62 percent of the total variance), ensuring there is meaningful variance in price after we fully condition on our severity groups. In addition to requiring that there is variance in price conditional on severity we are also assuming that hospital "quality" does not differ across price groups within a severity. There may be many dimensions of quality but expected health outcomes are clearly among them. Online Appendix 5 provides the analog of Table 6 when the four mother and infant outcome measures used in Table 4 are substituted for the "Price ($)" columns in Table 6.[45] There we see that as we move from max rank 1 to 2 and then max rank 2 to 3 in a given row (i.e., conditional on a given number of max rank comorbidities), 19 out of 20 of the differences are of expected sign (the single difference with the wrong sign is not significant) and most are statistically significant (17 out of 19).[46] When we consider differences in outcomes across entries that correspond to different numbers of comorbidities conditional on the same maximum rank (illustrating the differences across price groups within a severity), 6 out of 28 are decreases when we expect increases (though none are significant decreases), and of the remaining 22 only 8 are significant. Though the latter differences (those across rows within a given column) are small relative to those across columns, they are large enough to induce us to check for differences in outcome measures when we condition on the more detailed severity groups actually used in the analysis.

[43] We exclude a few patients who have no comorbidities known on admission besides the principal diagnosis and therefore have zero diagnoses of maximum rank.

[44] We use aggregated groups because for many of the hospitals there are no patients in some of our actual severity groups and these hospital severity pairs would not be used in the inequality analysis which follows.

[45] As for Table 4 we use a slightly different dataset that includes infant outcome variables. We exclude a small number of patients with no comorbidities known on admission other than the principal diagnosis.

[46] These numbers only consider differences where both entries are based on more than 400 patients. We thank Jesse Shapiro for suggesting this table.

When we further disaggregate by diagnosis and age (i.e., to the severity groups we actually use) and perform χ^2 tests for the outcome differences across price groups conditional on severity groups and hospitals (to correspond to our severity specific hospital effects) we find no significant differences.

B. *Inequality Analysis*

We work with each plan's data separately, so we omit the plan (our π) index from the notation below, with the understanding that all coefficients are plan specific. The inequality model makes the following assumptions:

$$(8) \qquad p(c_i, h) = \delta_h^o lp^o(c_i, h) - \epsilon_{p(c_i), h} \equiv p^o(c_i, h,) - \epsilon_{p(c_i), h}.$$

$\delta_h^o lp^o(c_i, h)$ was assumed equal to expected price in the logit analysis, so the difference between this specification and that used in the logit analysis is that the inequality analysis allows for an error in price and the logit analysis did not. Also

$$(9) \qquad g(q_h(s), s_i) = g^o(q_h(s), s_i) - \epsilon_{s_i, h}:$$

i.e., the inequality analysis places no restrictions on the quality severity interactions and allows for classification error in those interactions. The only assumption we will need on the two errors, $(\epsilon_{p(c_i), h}, \epsilon_{s_i, h})$, is that they be mean zero conditional on the patient's chosen plan and hospital and independent of the distance traveled to the hospital (we consider the robustness of our estimates with respect to these assumptions below).

Finally

$$(10) \qquad f(d(l_i, l_h)) = \theta_d d(l_i, l_h),$$

which differs from the specification in the logit analysis in that the squared term in distance has been eliminated because it did not affect any of the parameters of interest and would complicate the algebra below.

Substituting into equation (7) for a same-plan same-severity couple who could have chosen each other's hospital and are in different price groups (an $i \in S(h, h', s, c)$ and $i' \in S(h', h, s = s_i, c \neq c_i)$) our revealed preference inequality becomes

$$(11)\ 0 \leq \Delta W(i, h, h') + \Delta W(i', h', h) = \theta_p[\Delta p^o(c_i, h, h') + \Delta p^o(c_{i'}, h', h)]$$

$$+ \theta_d[\Delta d(l_i, l_h, l_{h'}) + \Delta d(l_{i'} l_{h'}, l_h)]$$

$$- \Delta \epsilon_{i, h, h'} - \Delta \epsilon_{i', h', h},$$

where $\Delta \epsilon_{i, h, h'} \equiv \theta_p \Delta \epsilon_{p(c_i), h, h'} + \Delta \epsilon_{s_{i'}, h', h}$.

Our inequalities for hospital h and insurer π are simply averages of equation (11) across switches between patients who chose hospital h and those who chose another hospital but could have chosen h and had the same severity and insurer,

but a different price group, as the patient who chose h. Formally, let $N_{h,h',s}^s$ be the number of switches between patients $i \in S(h, h', s, c)$ and $i' \in S(h', h, s = s_i, c \neq c_i)$ and for any $x(\cdot)$ define

$$(12) \qquad \Delta \bar{x}(h, h', s) \equiv \frac{1}{N_{h,h',s}^s} \sum_{i \in S(h,h',s)} \sum_{i' \in S(h',h,s,c \neq c_i)} \Delta x(i, h, h').$$

Then averaging equation (11) over switches we get

$$\theta_p(\Delta \bar{p}(h, h', s) + \Delta \bar{p}(h', h, s)) + \theta_d(\Delta \bar{d}(h, h', s) + \Delta \bar{d}(h', h, s))$$

$$- \Delta \bar{\epsilon}(h, h', s) - \Delta \bar{\epsilon}(h', h, s) \geq 0.$$

The moment inequalities we use in estimation are weighted averages of these inequalities where the weights are given by the fraction of comparisons that each contributes, or

$$w(h, h', s) \equiv \frac{N_{h,h',s}^s}{\sum_s \sum_{h'>h} N_{h,h',s}^s}.$$

If we let $\rightarrow_P$ read converges in probability, and note that our assumptions imply

$$\sum_{s,h'>h} w(h, h', s)\left(-\Delta \bar{\epsilon}(h, h', s) - \Delta \bar{\epsilon}(h', h, s)\right) \rightarrow_p 0,$$

then our model implies that at $\theta = \theta_0$

$$(13) \qquad m(h, \theta) \equiv \sum_{s,h'>h} w(h, h', s)[\theta_p(\Delta \bar{p}(h, h', s) + \Delta \bar{p}(h', h, s))$$

$$+ \theta_d(\Delta \bar{d}(h, h', s) + \Delta \bar{d}(h', h, s))] \rightarrow_P \kappa \geq 0.$$

The inequality in equation (13) is in terms of observables and the parameters of interest. Since it is an inequality, if a particular $[\theta_p, \theta_d]$ satisfies (13), then so will $[\kappa \theta_p, \kappa \theta_d]$ for any $\kappa > 0$. This implies that there is a free normalization, so we set $\theta_d = -1$. Thus we will only be able to estimate the ratio $\theta_p / ||\theta_d||$, which at the risk of some notational confusion, we will henceforth simply call θ_p.[47]

---

[47] Though the inequality in equation (13) allows for detailed hospital quality/patient severity interactions and errors in the price variable, it does rule out determinants of choice that are patient-hospital specific and are both: (i) not controlled for by the severity/quality interactions, price, or distance, and (ii) not differenced out by adding the difference in preference for hospital h over h' for one patient to the difference in preference for hospital h' over h for the other. The logit model does not allow for patient severity/hospital quality interactions or errors in the measure of price, but does allow for unobservables that are patient- and hospital-specific provided they are independently and identically distributed across both patients and hospitals and have an extreme value distribution. To the extent that the inequality model does not account for all the patient/hospital-specific determinants of choice there will be a selection problem in the resulting estimates which would be expected to narrow the estimated bounds. If the selection term is important we should expect a (nonrandom) fraction of our inequalities to switch signs, and if selection is important enough we will reject the null that there are values of θ which satisfy all the inequalities. In fact we accept below. Moreover, as we will see, there is no evidence at all of a disproportionate number of negative inequalities.

The inequality in (13) bounds θ_p. We can generate additional inequalities, and therefore bounds, by multiplying each inequality in equation (11) with an "instrument," say z, whose sign is the same for all observations. The additional moments will generate lower bounds if the expected value of $[\Delta\bar{p}(h, h', s) + \Delta\bar{p}(h', h, s)] \times z$ is positive and upper bounds otherwise. For a variable to be an instrument it must be known by the agents when their decisions are made and mean independent of the measurement errors. Our assumption that $(\epsilon_{p(c_i), h}, \epsilon_{s_i, h})$ are independent of the distance travelled implies that we can use as instruments the positive and negative parts of the distance differences, that is: $\Delta d(l_i, l_h, l_{h'})_+ \equiv \max\{d(l_i, l_h, l_{h'}), 0\}$, $\Delta d(l_i, l_h, l_{h'})_- \equiv \min\{d(l_i, l_h, l_{h'}), 0\}$, and analogously, $\Delta d(l_{i'}, l_{h'}, l_h)_+$, $\Delta d(l_{i'}, l_{h'}, l_h)_-$. We label the additional inequalities for instrument z, $m(h, z, \theta)$.

Details.—As noted we conduct the initial analyses separately by insurer. This allows all hospital quality-patient severity interactions as well as the price coefficient to differ by plan. The left-hand side of equation (13) is computed separately for each hospital and instrument. We then weight each of these terms by its estimated standard error. For small hospitals we are concerned that the average error either in the inequality, or more likely in the estimate of its standard error, may not be close to zero. As a result we develop a separate inequality for each hospital that has more than 1,000 patient switches but average over all hospitals with less than 1,000 patient switches. Overall we have between 78 and 285 moments per insurer: one for each combination of an instrument and a major hospital and an additional moment per instrument that includes hospitals with fewer patients.[48]

Our model tells us that at the true θ_0 each of the weighted moments should be positive, so we should reject any value of θ which makes any moment negative. However we have no reason to prefer one value of θ that generates positive moments over another that does also. So we form the "negative" part of the moments, say $m(z, h, \theta)_- = \min(0, m(z, h, \theta)$, weight them by their standard errors, and accept any value of θ which makes them all zero. If $\mathbf{m}(\theta)_-$ is the vector obtained after weighting each moment by its estimated standard error and then arranging them into a vector, and for any vector x, $||x|| \equiv \sum_j x_j^2$, then this is formally identical to accepting any θ which is in

$$(14) \qquad \hat{\Theta} = \arg\min_\theta ||\mathbf{m}(\theta)_-||.$$

$\hat{\Theta}$ could be a set of values all of which make all the moments zero, or it could be a point, indicating that no value of θ satisfies all the moment conditions. If it is a point then we test whether the fact that no value of θ satisfies all the constraints is a result of sampling error or is because the model is misspecified.[49]

---

[48] We exclude from the analysis hospitals that have fewer than 150 switches with all other hospitals in the sample combined (when instruments are included, each pair of hospitals is required to have at least 150 switches whose value of the instrument is nonzero). We also tried estimating the coefficients keeping the smaller-hospital moments separate. The estimated coefficients were almost always smaller in magnitude than our baseline results, consistent with small hospitals introducing measurement error, but in qualitative terms the story did not change.

[49] The test uses the moment shifting technique in Andrews and Soares (2010), and for computational reasons we use a related technique for 95 percent confidence intervals developed in Pakes et al. (2011). Both require development of the variance covariance of the inequalities across moments, and the formula for that matrix is available upon request.

C. *Robustness Tests*

As noted, there are possible issues with price measurement that could affect the analysis. In particular we do not observe the actual discount negotiated by each hospital-insurer pair. We do observe the average negotiated discount for private managed care patients at each hospital, so were each inequality obtained by averaging over all private managed care patients we would have the correct average price. There are at least two possible problems. First our inequalities are insurer-specific, and insurers' contracts with the same hospital can differ in ways that generate differences in average discounts. Second we consider only birth episodes and the average discount for birth episodes may differ from the overall average. In addition we did not allow for errors in the distance measure when in fact we only have information on the zip code in which the patient resides and calculate distances from the centroid of that zip code. We now develop procedures for checking whether our conclusions might be affected by these problems.

Allowing for Insurer-Specific Average Hospital Discounts.—Our data include the average negotiated discount at each hospital, our d_h. Our procedure for testing whether our results are robust to insurer-specific average hospital discounts begins by writing d_h as a share-weighted average of the average insurer discount at each hospital. We substitute into this equation data on insurer-specific revenue shares at each hospital and a parametric model for the insurer-specific average hospital discounts. This generates an estimable model for the needed discounts. The additional data on the insurer-specific share of hospital revenues needed for this analysis comes from the OSHPD hospital discharge and financial records for 2003. The average discounts are modelled as a function of hospital, plan, and market characteristics.

The model for the average insurer-specific hospital discount, our $d_{\pi,h}$, is the sum of two components; one for the average hospital discount and one for the insurer-specific deviation from that average. There are two ways to use the estimates to generate a prediction for $d_{\pi,h}$. First we can use the model's prediction directly; we denote this $\hat{d}^1_{\pi,h}$. Alternatively we can subtract the predicted average discounts of other insurers (appropriately weighted) from the observed d_h to generate a second prediction $\hat{d}^2_{\pi,h}$. The first prediction will not account for errors in predicting the hospital's average discount. The second prediction does account for hospital-specific prediction errors but does not account for the insurer-specific prediction errors of the other insurers. The two predictions are used to define price measures $p^1(\cdot) = (1 - \hat{d}^1_{\pi,h})lp^o(c_i, h)$ and $p^2(\cdot) = (1 - \hat{d}^2_{\pi,h})lp^o(c_i, h)$ which are substituted for our estimates of price in equation (8) in the inequality analysis. Recall that the predicted insurer-specific hospital discount in our baseline model is just the actual average hospital discount. The error in the moment conditions that the use of the different price measures generates is the difference between the predicted and actual average insurer-specific hospital discount times the average price.

We specify a logistic functional form for $d_{\pi,h}$. The explanatory variables that are interacted with plan revenue shares at the hospital include indicators for for-profit hospitals and hospitals that are members of systems (groups of providers that bargain jointly with insurers), indicators for teaching hospitals, and both insurer and

market fixed effects. However we also estimate other specifications that replace the fixed effects with market and insurer characteristics to check that our results are consistent with previous papers analyzing the impact of those characteristics on hospital prices. Details on both the models estimated and their coefficient estimates are provided in online Appendix 2. Though the fits are not extraordinary (with R^2s just under 0.5), the results are intuitive and accord with the prior literature (for example discounts increase with the number of hospitals per population and decrease with the number of insurers per population).[50]

Allowing for Variation in Discounts across Diagnoses.—The discount analysis just described makes the assumption that discounts are fixed across diagnoses within a hospital-insurer pair. In reality discounts may vary across services within a hospital. To ensure that ignoring this variation was not biasing our results we went back to our discount model and allowed the discount for deliveries to differ from that for other diagnoses. We also ran specifications allowing the labor/birth discount to differ across different types of hospitals: for example hospitals with high-tech delivery services could have a particularly good reputation for obstetrics and therefore negotiate low discounts. Augmenting our baseline analysis to allow for a separate discount for delivery episodes generated a significant coefficient which implied that births had a 6 percent higher discount than the average for other diagnoses. However we did not find significant differences in this discount across hospitals. When we substituted the delivery-specific discounts into the inequality analysis the coefficients differed very little from the baseline specification. We also tried estimating a different discount for Cesarean sections but the estimated coefficient was not statistically significant. See online Appendix 2 for further details.

Allowing for Errors in Our Distance Measure.—Our measure of the distance between patients and hospitals is the distance between the centroid of the patient's home zip code and the zip code of the hospital. It therefore contains measurement error.[51] Assuming the error is mean zero implies that it does not affect the properties of our original inequality (equation 13) as that simply averages out the estimation error. However when we take the positive and negative parts of the distance and use them as instruments, those instruments will, in general, contain an error which is

---

[50]Dranove and Satterthwaite (2000) and Gaynor and Vogt (2000) provide good reviews of the literature on discounts. Several other specifications were investigated. For example we replaced the insurer fixed effects with the plan percent capitation. This coefficient was positive and the other coefficient estimates were qualitatively unaffected by this change, foreshadowing our results that accounting for variation in discounts across insurers does not change the major results of our inequalities analysis. We also investigated whether the proportion of the insurer's patients sent to a particular hospital was correlated with the discount by including an interaction of this proportion with insurer fixed effects in the model. This relationship between the "channeling" of patients to a particular provider and the prices negotiated with that provider is analyzed in Sorensen (2003). When we excluded market fixed effects we estimated a significant positive relationship between patient channeling and discounts (a negative relationship between channeling and prices) for just one insurer, Blue Shield. The coefficient became insignificant when we added market fixed effects. We repeated the inequalities analysis for Blue Shield using this discount specification and the results changed very little.

[51]Latitudes and longitudes of the hospital's and patient's zip codes are taken from the US Census Bureau 1999 zip code file. Additional measurement error is caused by the fact that, if the zip code could not be located in the Bureau's internal database, the county internal point was assigned to the zip code. Multiple zip codes therefore have the same recorded latitude and longitude in some cases.

TABLE 7—RESULTS OF INEQUALITIES ANALYSIS

		Percent capitated	(1) $p(\cdot) = (1 - d_h)lp(c_i, h)$		(2) $\tilde{d}(i_h, h, h')_+$	
			$\hat{\theta}$	$[CI_{LB}, CI_{UB}]$	$\hat{\theta}$	$[CI_{LB}, CI_{UB}]$
Pacificare		0.97	−1.50**	[−1.68, −1.34]	−1.42**	[−1.66, −0.94]
Aetna		0.91	−0.92**	[−0.95, −0.86]	−0.73**	[−0.77, −0.66]
Health Net	Baseline	0.80	−0.17**	[−0.27, −0.13]	−0.60**	[−0.71, −0.51]
	Drop $t \leq -2$		−0.78**	[−0.80, −0.44]	−2.16**	[−2.26, −1.13]
Cigna		0.75	−0.35**	[−0.40, −0.33]	−0.66**	[−0.71, −0.58]
Blue Shield		0.57	−0.06	[−0.15, 0.23]	set	[−0.95, 0.65]
Blue Cross	Baseline	0.38	−0.10**	[−0.24, −0.01]	−0.27**	[−0.31, −0.25]
	Drop $t \leq -2$		−0.29**	[−0.31, −0.25]	−0.31**	[−0.34, −0.27]

Notes: Results of inequalities analysis. We include 157 hospitals in total. Estimated coefficient is the ratio of the price coefficient to the distance coefficient in the utility equation, where prices are measured in thousands of US dollars and distance in tens of miles. Price calculated using the observed average hospital discount. Specification includes four distance-based instruments (positive and negative parts of $d(i_h, h) - d(i_h, h')$ for each patient) plus a constant in the instrument set. In column 2 these instruments are replaced with indicators for distance differences being greater than three miles. The rows labeled "Drop $t \leq -2$" report results when we dropped moments whose t-statistic values were less than −2 (2 out of 182 for Health Net; 5 out of 285 for Blue Cross) and repeated the estimation process. "set" for Blue Shield in column 2 indicates a range of values.
    ** Significant at the 5 percent level.
    * Significant at the 10 percent level.

correlated with the error in distance in the original equation, and this can generate biases in our estimate of θ_p.

To ensure that this was not having a major impact on our results we modified our distance instruments $(\Delta d(l_i, l_h, l_{h'})_+, \ \Delta d(l_i, l_h, l_{h'})_-, \ \Delta d(l_{i'}, l_{h'}, l_h)_+, \ \Delta d(l_{i'}, l_{h'}, l_h)_-)$ as follows. Instead of using $\Delta d(l_i, l_h, l_{h'})_+$ we used $\Delta \tilde{d}(l_i, l_h, l_{h'})_+$ where

$$\Delta \tilde{d}(l_i, l_h, l_{h'})_+ \ = \ 1 \quad \text{if} \quad \Delta d(l_i, l_h, l_{h'})_+ \ \geq \ 3, \quad \text{and} \quad \Delta \tilde{d}(l_i, l_h, l_{h'})_+ \ = \ 0 \quad \text{otherwise.}$$

We did the analogous transform to the other distance instruments. This addresses the problem if we assume that the error in the distance is not greater than three miles, so that we know the incremental distance between hospitals is positive (negative) if the observed differences are greater than three miles.

D. *Inequality Results*

Tables 7 and 8 report results from the inequalities analysis. The first column of Table 7 reports our main results. These assume that the price measure obtained by multiplying hospital-specific discounts by the expected list price is correct up to an error which is mean zero conditional on the plan and the choice of hospital. The second column uses the modified distance measure described above to check for the possible impacts of errors in the distance measure. Table 8 reports results using $p^1(\cdot)$ and $p^2(\cdot)$ respectively (and assuming no errors in the distance measure). Their validity requires additional assumptions on the prediction errors generated when forming these variables, but they allow us to investigate whether our results are robust to allowing for plan-specific discounts within hospitals.

There is no value for $\theta_{p, \pi}$ that satisfies all the inequality constraints in any specification except one (Blue Shield in column 2 of Table 7). When this occurs the

TABLE 8—INEQUALITIES ANALYSIS, DIFFERENT PRICE MEASURES

		Percent capitated	(1) $p(\cdot) = (1 - \hat{d}^1_{\pi,h})lp(c_i, h)$		(2) $p(\cdot) = (1 - \hat{d}^2_{\pi,h})lp(c_i, h)$	
			$\hat{\theta}$	$[CI_{LB}, CI_{UB}]$	$\hat{\theta}$	$[CI_{LB}, CI_{UB}]$
Pacificare		0.97	−1.07**	[−1.52, −0.62]	−1.47**	[−1.64, −1.32]
Aetna		0.91	−0.68**	[−0.72, −0.62]	−0.77**	[−0.81, −0.71]
Health Net	Baseline	0.80	−0.11**	[−0.23, −0.07]	−0.20**	[−0.30, −0.17]
	Drop $t \leq -2$		−0.41	[−0.43, 0.94]	−1.87**	[−1.89, −1.33]
Cigna		0.75	−0.35**	[−0.39, −0.33]	−0.32**	[−0.36, −0.30]
Blue Shield		0.57	0.18	[−0.16, 0.79]	0.004	[−0.28, 0.70]
Blue Cross	Baseline	0.38	−0.03	[−0.18, 0.39]	−0.09**	[−0.22, −0.01]
	Drop $t \leq -2$		−0.12**	[−0.14, −0.05]	−0.18**	[−0.21, −0.14]

Notes: Results of inequalities analysis using different price measures. See notes to Table 7 for details. Price measures are calculated using the two estimated hospital-insurer level discounts discussed in Section VI. Rows labeled "Drop $t \leq -2$" report results when we dropped moments whose t-statistic values were less than −2 (2 out of 182 for Health Net; 5 out of 285 for Blue Cross) and repeated the estimation process.

** Significant at the 5 percent level.
* Significant at the 10 percent level.

TABLE 9—SUMMARY OF t-STATISTICS FROM INEQUALITIES ANALYSIS

	Pacificare	Aetna	Health Net	Cigna	Blue Shield	Blue Cross
Summary of t-statistics						
Number positive	152	75	173	93	170	254
Avg value of positive	12.7	22.5	17.1	19.5	19.1	21.5
Number negative	11	3	9	2	4	31
Number $t < -2$	0	0	2	0	0	5

Notes: Summary of estimated t-statistics of the moments used in inequalities analysis. t-statistic = value of the moment at the estimated $\theta_{\pi,p}$ (for specification where $p(\cdot) = (1 - d_h)lp(c_i, h)$). Under the model all moments should be non-negative.

estimation algorithm produces a point estimate: the value of $\theta_{p,\pi}$ that minimizes the sum of squares of the negative part of the (standardized) moments. Given the number of inequalities we have for each of our plans we are not surprised to find point estimates. In finite samples when each moment is evaluated at the true value of the parameter vector it generates a variable which distributes approximately normally. Consequently the greatest of the values from the moments which provide lower bounds has a positive bias. Similarly the least upper bound has a negative bias. So depending on the magnitude of the biases the bounds can easily cross producing a point estimate. The expected magnitude of these biases increases with the number of moments.

There is a standard statistical test for whether sampling errors of this form exist and the results always indicated that we could accept the null that there were values of θ that satisfied all the inequalities. Table 9 makes it clear why we accept the null. It provides estimates of the t-statistics obtained when we evaluate all moments used at the estimated value of $\theta_{p,\pi}$ for each plan. The model predicts that the expectation of all moments are non-negative. Our results indicate that of the 977 moments evaluated only 60, or about 6 percent, are less than zero, and only 7 out of 977, or 0.7 percent of the moments, have t-values less than −2. In 4 of the 6 plans studied none of the moments are significantly negative at the traditional p-value of 0.05. Health Net has 2 out of 182 moments with t-value less than −2 and Blue Cross has
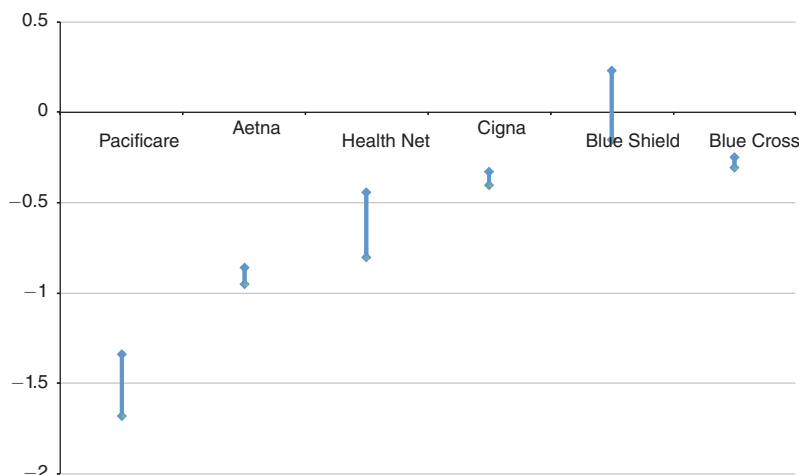
Figure 1. Correlation of Estimated Price Coefficient
with Insurer's Percent Capitation Payments

Notes: Graph to illustrate confidence intervals for insurer price coefficients, reported in Table 7. Estimates are from model where $p(\cdot) = (1 - d_h)lp(c_i, h)$.

5 out of 285 with *t*-values less than -2. For these two insurers we re-estimate $\theta_{p, \pi}$ after dropping the moments with *t*-statistics less than -2. The results are reported in the rows of Tables 7 and 8 labeled "Drop $t < -2$".

In the first column of Table 7 the price coefficients for all insurers other than Blue Shield are negative and statistically significant at $p = 0.05$. That for Blue Shield is small, negative and statistically insignificant. As is traditional for set estimators, we focus on confidence intervals for $\theta_{p, \pi}$. These are illustrated for each insurer in Figure 1. The coefficients for all insurers except Blue Shield are "set ordered" by decreasing percent capitation; that is, the upper bound of the confidence interval for one insurer is below the lower bound for the insurer with the next-highest percent capitation. The picture is less clear for Blue Shield. Its confidence interval is above that for Blue Cross and crosses zero.

The results from substituting $p^1(\cdot)$ and $p^2(\cdot)$ for the $p^o(\cdot)$ in the inequality analysis are provided in Table 8. They are similar to the results in our main specification. The two major differences occur when using $p^2(\cdot)$; then the Health Net coefficient estimated when we drop the two negative moments is larger in absolute value, and the Blue Shield coefficient is positive with a confidence interval that crosses zero (instead of being negative with a confidence interval crossing zero). We conducted a number of other robustness tests that involved the price variable, but none had anything but the expected effect on the results.[52]

Column 2 of Table 7 provides the results using the modified distance instrument that takes account of the possibility of error in our distance measure (it uses the

---

[52] We repeated the inequalities analysis using just the list price (rather than its interaction with the discount). The pattern of results was unchanged in that high-capitation insurers had more negative price coefficients in general than other insurers. However all price coefficients were closer to zero than those in Table 7, consistent with our expectation that measurement error should affect these results.

TABLE 10—MAGNITUDES OF LOGIT AND INEQUALITY RESULTS

	Percent capitated	Logits (less-sick patients) average η_i	Inequalities (all patients) average η_i
Pacificare	0.97	0.33	11.08
Aetna	0.91	0.10	11.47
Health Net	0.80	0.15	6.52
Cigna	0.75	0.10	2.49
Blue Shield	0.57	−0.08	0.51
Blue Cross	0.38	−0.03	3.24

Notes: Estimated cross-patient average value of $\eta_i = \frac{\partial d_i}{\partial p_i} \frac{p_i}{d_i}$ for each insurer implied by logit and inequality analyses. logit model uses less-sick population as defined in notes to Table 5. Inequality model uses price defined using discount δ_h (column 1 of Table 7).

$\tilde{d}(\cdot)$ instruments defined in the prior subsection). The results are very similar to those above. All coefficients but those for Blue Shield, and to a lesser extent Health Net are similar to those in column 1. The confidence interval for Blue Shield indicates that the data are not informative about the Blue Shield price coefficient, and the Health Net coefficient, though always significantly negative, varies in magnitude with whether or not we keep the two inequalities that are significantly negative.

With the possible exception of Blue Shield, these results indicate that the allocation of patients to hospitals responds to the prices the insurer pays those hospitals. Moreover insurers with more capitated payments to physicians have hospital referral processes that place a more negative weight on prices than other insurers. We cannot say much about Blue Shield. The bounds for its coefficient are large and vary quite a bit across the specifications we tried. Recall that Blue Shield is the only not-for-profit insurer in our data, and so might be expected to differ.[53] As a result we disregard Blue Shield in the analysis that follows.

The difference between the inequality results and those from the logit analysis are striking. To get an idea of the importance of this difference Table 10 compares the elasticity of price with respect to distance computed using the logit estimates for the least sick patients (Table 5), to those same elasticities computed using the price coefficient estimated from the inequalities (column 1 of Table 7). That is, we measure on average how much further the consumer would have to drive (in percentage terms) to just offset a one percent price increase.

Consider first the comparison of elasticities derived from the logit price coefficients to those derived from the inequality estimates for the plans where the logits estimated a negative price coefficient. All the elasticities obtained from the inequality estimates are more than an order of magnitude larger than those obtained from the logit estimates, and some are more than two orders of magnitude larger. In addition two of the elasticities obtained from the logit estimates have the wrong sign. Notice also that the inequality estimates indicate that the average elasticity increases by almost a factor of four when we move from the least capitated for-profit insurer (Blue Cross) with a capitation rate of 38 percent to Pacificare whose capitation rate is 97 percent.

---

[53] As noted above, it is also the only insurer that mandated tiered networks by 2003, and the available evidence indicates that it is the only insurer with a noticeable fraction of enrollees on a tiered network in that year. This is yet another reason to think the analysis for Blue Shield should be different.

VII. Cost, Quality, and Distance Trade-Offs

This section is divided into two parts. In the first we derive insurer- and severity-specific estimates of quality differences across hospitals. The estimated qualities are quite similar across insurers. We illustrate this by estimating a model which requires the within-severity ratings of different plans to be linear transforms of each other, and then providing a plot of the restricted on the unrestricted quality estimates. The next subsection examines implications of these results. When the within-severity orderings are linear transforms of each other each plan's preference ordering over hospitals (equation (1)) is a different linear function of price, distance, and a common quality index. This allows us to investigate how the quality-price-distance trade-off varies across insurers. We conclude this subsection by comparing our results on these trade-offs to the implications of the data on health outcomes (the data that underlie Table 4). This reinforces our results and provides some external validity for our quality controls.

A. *Plan- and Severity-Specific Hospital Quality Terms*

The revealed preference inequality in equation (11) implies that any given value of $\theta_{p,\pi}$ generates a set of bounds for differences in the quality terms across hospitals. We evaluate these differences at the estimates of $\theta_{p,\pi}$ given in column 1 of Table 7.[54] Since we can only recover differences in quality and we can only compare hospitals within a market, we estimate quality coefficients that are: plan-, severity- and market-specific. To ease notation we will omit the plan and market indices below.

Recall that $\Delta\bar{x}(h, h', s)$ is the average of $\Delta x(i, h, h')$ among the patients with severity s who chose hospital h but could have chosen h'. Then since every one of those patients chose h over h' revealed preference implies

$$g(q_h, s) \ - \ g(q_{h'}, s) \ \geq \ -\theta_p \Delta\bar{p}(h, h', s) \ + \ \Delta\bar{d}(h, h', s) \ + \ \Delta\bar{\epsilon}_{h, h', s}$$

$$\equiv \ \hat{q}(h, h', s) \ + \ \Delta\bar{\epsilon}_{h, h', s},$$

with $\hat{q}(h, h', s)$ observable. Moreover since the $\bar{\epsilon}_{h, h', s}$ are mean zero conditional on the hospital choice

$$(15) \qquad \hat{q}(h, h', s) \ \sim> \ \mathcal{N}\big(\underline{q}(h, h', s), \sigma^2(h, h', s)/N_{h, h', s}\big),$$

where

$$\underline{q}(h, h', s) \ \leq \ g(q_h, s) \ - \ g(q_{h'}, s),$$

"$\sim>$" reads converges in distribution to, $\mathcal{N}(\cdot, \cdot)$ is the normal distribution, $\sigma^2(h, h', s)$ is the variance of $(-\theta_p \Delta p(c_i, h, h') \ + \ \Delta d(l_i, l_h, l_{h'}))$ across observations in $S(h, h', s)$, and $N_{h, h', s}$ is the cardinality of that set.

---

[54] The correlation of the quality terms across different values within the confidence intervals of $\theta_{p,\pi}$ is nearly one, so it makes little difference which of the values in the confidence intervals reported there we use.

Each couple of hospitals generates two quality bounds of this form: one from the patients who chose h but could have chosen h' and one from those that chose h' and could have chosen h. The former provides a lower bound and the latter an upper bound to the difference in quality between h and h'. So if there are H hospitals in a market, there are $H(H-1)$ estimates of quality bounds for each severity.

Estimating the Quality Bounds.—Recall that we can only bound *differences* in hospital quality. So we set one hospital's quality to be zero (the same hospital for each insurer for a given severity and market). Indexing that hospital by H one can show that the inequalities that relate to hospital h are given by

$$(16) \qquad \bar{q}(h,s) \equiv \min_{h' \neq h} E[-\hat{q}(H,h',s) - \hat{q}(h',h,s)] \geq g(q_h, s)$$

$$\geq \max_{h' \neq h} E[\hat{q}(h',H,s) + \hat{q}(h,h',s)] \equiv \underline{q}(h,s).$$

We stack these inequalities for each hospital, weight each by its estimated standard error, and then find the (set) estimator that minimizes the squared inequality violations.

Recall that we used very detailed severity groups for the estimation of price coefficients so as to ensure we eliminated biases that might be caused by unobserved quality terms. The sample sizes associated with those groups are quite small and to obtain the quality estimates we do not average over severity groups as we did to obtain the price coefficient estimates. Moreover we are not interested in orderings of hospitals at that fine a level of severity. We therefore use the severity classifications given to us by the obstetricians we consulted to aggregate into five "super-severity" groups. These consist of four groups all of whose patients have identical principal diagnosis and comorbidity rankings, and a fifth group which contains all the remaining patients.[55] Finally the actual estimates of the quality bounds depend on the prior estimates of $\theta_{p,\pi}$. The results presented below use the point estimates from the first column of Table 7, but the implied quality bounds varied very little when we considered other points within their respective confidence intervals.

In computing the quality estimates we included moments for patients who went to hospital h but could have chosen hospital h' for a given severity if there were five or more patients who were admitted to hospital h and had hospital h' in their choice set. Over our 12 markets and 5 severity groups, we obtained 1,176 quality estimates.[56] Almost all of these estimates, 1,078 of them, come from our 5 largest markets (Los

---

[55] The super-severity groups are: Group I contains 55 percent of patients who have a rank 1 (routine) principal diagnosis, rank 1 comorbidities, and are young; Group II has 11 percent of patients who have rank 2 principal diagnosis, rank 1 comorbidities, and young; Group III has 15 percent of patients who have rank 1 principal diagnosis, maximum rank 2 comorbidities, and are young; Group IV has 12 percent of patients and they have a rank 2 principal diagnosis, maximum rank 2 comorbidities, and young; and Group V has 6 percent of patients that are not included in the other groups.

[56] 477 of these were sets and 699 were points. We tested whether the points satisfied the appropriate vector of moment inequality constraints (the sets necessarily do). Slightly more than half did. However when we go to switches between individual hospitals for a given severity and plan there is a limited amount of data per moment. So the asymptotic approximation inherent in the moment inequality test statistic is questionable. Moreover, as we show below the actual estimates satisfy most of the properties our priors might associate with them.

Angeles, Orange County, Inland Empire, the Bay Area, and San Diego), so we confine the remainder of the analysis to these 5 markets.

Does the Implied Order Make Logical Sense?—For the ordering across hospitals to make logical sense it must obey transitivity. There are at least two ways we can check this, one of which does not rely on our estimates of the price coefficient and one of which does.

Temporarily ignore estimation error. Then if both $\Delta\bar{p}(h, h', s)$ and $\Delta\bar{d}(h, h', s)$ are positive the perceived quality of h for severity s must be higher than that of h'. This because patients chose hospital h over h' despite the fact that h was both more distant and had higher prices. This fact generates a partial ordering across hospitals that *does not require* either estimates of the price coefficients or estimates of the quality terms. Alternatively we could use our estimate of $\theta_{p,\pi}$ to ask whether the partial order obtained from the sign of $\Sigma(\theta_{p,\pi}\Delta\bar{p}(h, h', s) - \Delta\bar{d}(h, h', s))$ for each pair (h, h') obeys transitivity.

The estimation procedure we use does not guarantee that the order we obtain from using either of these sets of inequalities satisfies the logical condition of transitivity. That is there could be cycles of the form

$$h_1 \succ h_2, \quad h_2 \succ h_3, \quad \text{but} \quad h_3 \succ h_1,$$

or even more simply we could find that

$$h_1 \succ h_2 \quad \text{but} \quad h_2 \succ h_1.$$

To check this we compute all possible cycles from both the *nonparametric* bounds and the bounds that use our estimates of θ_p for each of our five insurers, in each of our five markets for each of our five severities. The nonparametric procedure yields only 543 possible orderings, and none violate transitivity. When we use our estimates of θ_p and the estimation algorithm described above there are 10,526 possible orderings, and of these 1,069, or about 11 percent, actually cycle. However almost all of these are associated with bounds that are estimated imprecisely. Only 3 or 0.03 percent of the possible cycles are significantly negative at the 5 percent level. We take this as evidence that the data generates a hospital ordering that satisfies rationality constraints. We now consider that ordering in more detail.

Similarity of the Implied Orders across Plans.—1,078 estimates is still too many to examine individually, and our primary interest is not in the quality estimates per se but in the implied trade-off between price, quality, and distance. Moreover, the similarity in the estimated rankings of hospitals across insurers within our severity groups is striking, and this implies that some aggregation across plans is warranted. Note that since patients are assigned to a unique insurer, there is no statistical relationship between the moments used for the different insurers (except the relationship due to our using the same price measure across insurers). The similarity in ranks is a result of the similarity in (almost) statistically independent quality estimates generated by the referral processes of the different plans.
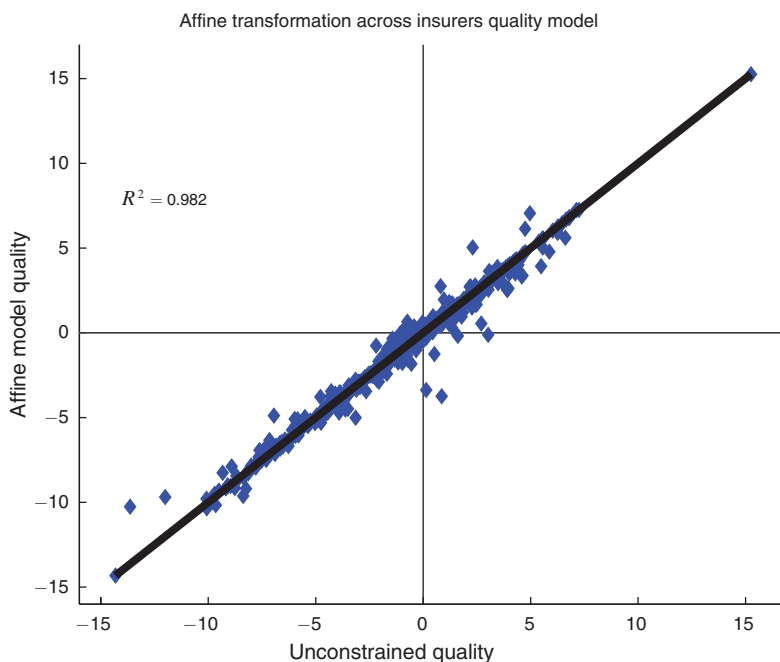
Figure 2. Graph of Constrained against Unconstrained Quality Estimates

Note: Figure 2 plots constrained quality estimates against unconstrained estimates, where unconstrained are $g_\pi(q_{h,s})$ and constrained are defined as $\alpha_{\pi,m,s}^0 + \alpha_{\pi,m,s}q_{h,s}$.

Figure 2 illustrates this similarity. This figure plots the estimates obtained from imposing the constraint that the plan-specific orderings for our five severity groups in our five largest markets are linear transforms of one another. That is, reintroducing the plan (π) and market (m) indices, we substitute

$$(17) \qquad g_\pi(q_h, s) = \alpha_{\pi,m,s}^0 + \alpha_{\pi,m,s}q_{h,s}$$

for the quality terms into equation (16), and re-estimate. When we do this the α^0 coefficients can not be separated from the quality of the reference hospital in each market and, since we can only compare quality estimates across plans, the $\alpha_{\pi,m,s}$ can only be analyzed proportionately to those of a base plan. Consequently in what follows we set the $\alpha_{\pi,m,s}$ coefficient for Blue Cross equal to 1 in each market and severity.

When we impose the constraints in equation (17) we reduce the number of coefficients estimated from 1,078 to 452 coefficients. Figure 2 plots the constrained against the unconstrained estimates.[57] The fitted line captures 98.2 percent of the variance of the unconstrained estimates. We then imposed the further constraint that

[57] When the unconstrained quality estimate was a set the error in the fit of the point was set to zero if the line went through the set, and was set to the distance between the set's bound and the line when it did not. Sixty-two of the 452 constrained coefficients were sets. When the constrained estimate was a set and the line went through the
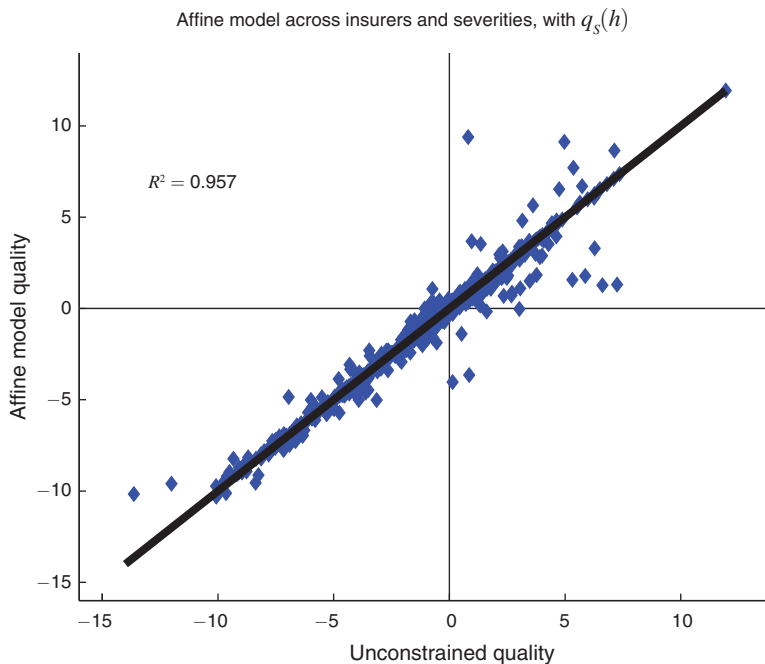
Figure 3. Adding a Constraint to the Quality Estimates

Notes: Figure 3 repeats the exercise in Figure 2 but defines constrained estimates as $\alpha^0_{\pi,m,s} + \alpha_\pi q h, s$. See Section VII for details.

$$(18) \qquad \alpha_{\pi,m,s} = \alpha_\pi.$$

This reduced the number of parameters estimated to 380. Figure 3 plots the constrained against the unconstrained estimates after imposing the additional constraint. The fitted line now captures 95.7 percent of the unconstrained variance. Though the difference in fit between Figures 2 and 3 is noticeable, it is rather small; we lose about three-hundredths of one percent of the fit per additional constraint. Moreover, if we impose the constraint in equation (18) there is a straightforward way to compare the way different plans trade-off costs, quality, and distance.

B. *Trade-Offs*

We now accept the constraints in equations (17) and (18) and substitute the results into the equation which determines hospital choice (equation (1)). To get directly at the price-quality and distance-quality trade-offs we divide the resulting equation by α_π so

$$(19) \qquad W_{i,\pi,h} \propto \left(\frac{\theta_{p,\pi}}{\alpha_\pi}\right) p(c_i, h, \pi) - \left(\frac{1}{\alpha_\pi}\right) d(l_i, l_h) + q_{h,s_i} + \left(\frac{1}{\alpha_\pi}\right) \epsilon_{i,\pi,h}.$$

---

set, we placed the point on the line in the figures, and when the line did not go through the set we chose the closest value to the line from the set.

TABLE 11—TRADE-OFFS AGGREGATED OVER MARKETS AND SEVERITIES

Insurer	P-care	Aetna	HNet	Cigna	BC
Percent cap	0.97	0.91	0.80	0.75	0.38
$\theta_{p,\pi}$	-1.50	-0.92	-0.78	-0.35	-0.29
α_π	5.13	3.12	2.63	1.20	1.00
$\theta_{p,\pi}/\alpha_\pi$	-0.293	-0.295	-0.297	-0.291	-0.290
$1/\alpha_\pi$	0.20	0.32	0.38	0.83	1.00
Upper and lower bounds on CI $\theta_{p,\pi}/\alpha_\pi^*$					
Lower	-0.38	-0.36	-0.35	-0.40	-0.31
Upper	-0.23	-0.23	-0.15	-0.22	-0.25

*Calculated as lower bound (upper bound) θ_p divided by upper bound (lower bound) α_π.

Table 11 provides the plan-specific estimates of the coefficients in equation (19).

The first two rows of the table reproduce the capitation rates and price coefficients from prior tables. As noted the (absolute value of the) price coefficients are ordered by the capitation rates. The third row shows that the quality coefficients are ordered in exactly the same way. As a result the ratio of the price coefficient to the quality coefficient is virtually constant across plans. The fourth row shows that this ratio lies between -0.29 and -0.30 for all five plans. In addition this ratio is estimated quite precisely. If we take upper and lower bounds to that ratio obtained by dividing the upper (lower) limit of the confidence interval for $\theta_{p,\pi}$ by the lower (upper) limit of the confidence interval for $1/\alpha_\pi$, we find that the lower limit only varies between -0.31 and -0.40 while, with the exception of Health Net, the upper bounds vary only between -0.22 and -0.25 (Health Net has an upper bound of -0.15, and as noted earlier our estimates of its values are somewhat sensitive to the precise specification of the price and distance terms).

Table 12 provides the estimates of $\theta_{p,\pi}/\alpha_\pi$ before we impose the constraint in equation (18); i.e., for each market and severity separately. There we see that Table 11 does hide some variance in the estimates of the parameter determining the cost-quality trade-off across markets and severities. However the difference between these numbers and those for $\theta_{p,\pi}/\alpha_\pi$ in Table 11 is largely in the smaller markets and severities. As a result when we impose the constraint in (18), the LA, Bay Area, and (to a lesser extent) Orange County moments for the first three severities dominate, and they do not differ much across either plans or severities.

The ratio of the price to the quality coefficient represents the trade-off between costs and quality. What the estimates are telling us is that the cost-quality trade-off is, as far as we can tell, independent of the capitation rate. This despite the fact that the higher the capitation rate the more sensitive hospital referrals are to price. Apparently though the high capitation plans are willing to send their patients to relatively far-off hospitals to save on hospital costs, they are not willing to sacrifice quality for cost savings. The trade-off between cost, quality, and distance only differs between plans in the trade-off between patient convenience and cost, not between quality and cost.

Of course our "quality" measure is simply whatever is implicit in the referral process: it captures everything that makes the hospital attractive after accounting for price and distance. The results we report above indicate that the quality rankings are similar across insurance plans. Online Appendix 6 shows that there is fairly

TABLE 12—COST-QUALITY TRADE-OFFS BY MARKET AND SEVERITY

Insurer	P-care	Aetna	HNet	Cigna	BC
LA S1	−0.29	−0.29	−0.29	−0.29	−0.29
LA S2	−0.33	−0.31	−0.31	−0.39	−0.29
LA S3	−0.28	N/A	−0.30	−0.29	−0.29
LA S4	−0.31	−0.30	−0.32	−0.34	−0.29
LA S5	−0.29	N/A	N/A	N/A	−0.29
Bay S1	−0.34	−0.32	−0.32	−0.30	−0.29
Bay S2	−0.44	−0.89	−0.48	−0.43	−0.29
Bay S3	−0.37	−0.35	−0.39	−0.33	−0.29
Bay S4	−0.38	−0.31	−0.29	−0.33	−0.29
Bay S5	N/A	N/A	−0.19	−0.35	−0.29
Ora S1	N/A	−0.28	N/A	−0.28	−0.290
Ora S2	N/A	−0.21	N/A	−0.31	−0.290
Ora S3	N/A	−0.24	N/A	−0.30	−0.290
Ora S4	N/A	−0.41	N/A	−0.40	−0.290
Ora S5	N/A	N/A	N/A	−0.69	−0.290
SD S1	−0.42	−0.55	−0.41	−0.26	−0.290
SD S2	−2.07	−0.88	−0.50	−107.9	−0.290
SD S3	−0.40	−0.43	−0.51	−0.31	−0.290
SD S4	−0.28	−0.23	−0.28	−0.73	−0.290
SD S5	−1.31	N/A	−56.5	N/A	−0.290
IE S1	−2.02	−0.68	−0.81	−0.46	−0.290
IE S2	−1.37	N/A	−0.78	N/A	−0.290
IE S3	−1.02	N/A	−0.50	N/A	−0.290
IE S4	−0.52	N/A	−0.67	N/A	−0.290
IE S5	N/A	N/A	N/A	N/A	N/A

Note: LA: Los Angeles, Bay: Bay Area, Ora: Orange County, SD: San Diego, IE: Inland Empire.

substantial variation in our measure of quality. There we convert the quality variable to thousands of US dollars by calculating $\alpha_\pi q_{h,s}/\theta_{p,\pi}$, and then we calculate its value for each patient-hospital pair. We take the standard deviation across hospitals for each patient and then average over patients in each super-severity group and each market. The average variation in this quality measure is of the same order of magnitude as the variation in price within each market and super-severity group. Also, perhaps not surprisingly, the estimates indicate that the variation in both price and the quality variable increases with the severity of illness.[58]

This variance in our quality measure could be attributable to many things: patient preferences for hospital amenities, physician perceptions of clinical quality, and any other factors (other than price and distance) which affect referrals. We have separate work in progress to investigate which observable characteristics are most related to it. We can however make one connection which throws further light on our focus here: the relationship between capitation rates and the trade-off between price, convenience, and quality. Table 4 showed that older and sicker women had significantly higher probabilities of readmission within 12 months and of discharge to "other than home," and their infants were also more likely to be discharged to "other than home." We calculated χ^2 test statistics for differences across plans in the probability

---

[58] Online Appendix 6 summarizes the variation in quality and in price faced by the typical patient across hospitals. The table records, for each super-severity group and each market where quality terms were estimated, the cross-patient average of the standard deviation in expected price (in thousands of US dollars) and in quality (measured in the units described above) across hospitals in the choice set. The variance in both variables increases with severity of illness and the two are comparable in magnitude (for example for the least-sick super-severity the average standard deviation in quality in Los Angeles is $1,136 while the average price standard deviation is $1,278).

of each adverse outcome conditional on each of our five super-severity groups. Not 1 of the 40 test statistics (10 pairwise comparisons of insurers, for 4 outcomes each) was significant at the traditional 5 percent level.

VIII. Conclusions

The results of this paper indicate that the prices paid by insurers to hospitals for obstetric care: (i) affect allocations of patients across the hospitals in the network, and (ii) have an impact which is greater the more highly capitated the insurer. Our second major finding relates to the trade-offs made between price, quality, and patient convenience. We find that, in the more highly capitated insurers, price reductions are achieved by sending patients to relatively far-off hospitals. There is no evidence that quality of care, or health outcomes, suffer as a result of this behavior: the trade-off between quality and price is constant across plans while that between convenience and price is not. Our estimates summarize the preferences generated by a complicated decision-making process which involves both physician and patient choices. The data do not allow us to investigate the extent to which particular mechanisms drive the estimates, so we have to leave that question to future research.

At least in the context of obstetrics, these findings have obvious implications for the impact of the ongoing institutional changes in the health care sector. Most importantly, the use of capitation payments in accountable care organizations is likely to reduce costs and is unlikely to result in a reduction in quality of care. There are also several other possible implications. For example, as noted earlier, currently just under one-half of accountable care organizations include a member hospital. There may be a benefit to such vertical integration through improvements in coordination of care. However, if providers favor within-ACO hospitals, vertical integration is likely to limit the cost reductions that would otherwise result from capitation. This is one of several trade-offs which merit further investigation. Also the capitation incentives used to reimburse ACOs are supposed to condition on quality, and the right measure of quality is unclear. Our methods result in a measure of hospital quality that reflects patient and physician preferences. There is a question worth exploring about whether our quality measure would be helpful in this context.

REFERENCES

Andrews, Donald W. K., and Gustavo Soares. 2010. "Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection." *Econometrica* 78 (1): 119–57.

Bajari, Patrick, Han Hong, Minjung Park, and Robert Town. 2011. "Regression Discontinuity Designs with an Endogenous Forcing Variable and an Application to Contracting in Health Care." National Bureau of Economic Research Working Paper 17643.

Baumgarten, Allan. 2004. "California Health Care Market Report 2004." California Healthcare Foundation. http://www.chcf.org/topics/view.cfm?itemID=114640 (accessed October 8, 2014).

Berwick, Donald M. 2011. "Making Good on ACOs' Promise—The Final Rule for the Medicare Shared Savings Program." *New England Journal of Medicine* 365 (19): 1753–6.

Burns, Lawton R., and Douglas R. Wholey. 1992. "The Impact of Physician Characteristics in Conditional Choice Models for Hospital Care." *Journal of Health Economics* 11 (1): 43–62.

Capps, Cory, David Dranove, and Mark Satterthwaite. 2003. "Competition and Market Power in Option Demand Markets." *RAND Journal of Economics* 34 (4): 737–63.

Chamberlain, Gary. 1980. "Analysis of Covariance with Qualitative Data." *Review of Economic Studies* 47 (1): 225–38.

Chandra, Amitabh, David M. Cutler, and Zirui Song. 2012. "Who Ordered That? The Economics of Treatment Choices in Medical Care." In *Handbook of Health Economics.* Vol. 2, edited by Mark V. Pauly, Thomas G. McGuire, and Pedro Pita Barros, 397–432. Amsterdam: Elsevier B.V.

Charlson, Mary E., Peter Pompei, Kathy L. Ales, and C. Ronald MacKenzie. 1987. "A New Method of Classifying Prognostic Comorbidity in Longitudinal Studies: Development and Validation." *Journal of Chronic Diseases* 40 (5): 373–83.

Cutler, David M., Mark B. McClellan, and Joseph P. Newhouse. 2000. "How Does Managed Care Do It?" *RAND Journal of Economics* 31 (3): 526–48.

Department of Health and Human Services. 2012. "Accountable Care Organizations: What Providers Need to Know." *Medicare Learning Network ICN 907406.*

Dranove, David, and Mark A. Satterthwaite. 2000. "The Industrial Organization of Health Care Markets." In *Handbook of Health Economics.* Vol. 1B, edited by Anthony J. Culyer and Joseph P. Newhouse, 1093–139. Amsterdam: Elsevier B.V.

Duggan, Mark G. 2000. "Hospital Ownership and Public Medical Spending." *Quarterly Journal of Economics* 115 (4): 1343–73.

Gaynor, Martin, James B. Rebitzer, and Lowell J. Taylor. 2004. "Physician Incentives in Health Maintenance Organizations." *Journal of Political Economy* 112 (4): 915–31.

Gaynor, Martin, and William B. Vogt. 2000. "Antitrust and Competition in Health Care Markets." In *Handbook of Health Economics.* Vol. 1B, edited by Anthony J. Culyer and Joseph P. Newhouse, 1405–87. Amsterdam: Elsevier B.V.

Gaynor, Martin, and William B. Vogt. 2003. "Competition among Hospitals." *RAND Journal of Economics* 34 (4): 764–85.

Glied, Sherry. 2000. "Managed Care." In *Handbook of Health Economics.* Vol. 1A, edited by Anthony J. Culyer and Joseph P. Newhouse, 707–53. Amsterdam: Elsevier B.V.

Goldman, Dana, and John A. Romley. 2008. "Hospitals as Hotels: The Role of Patient Amenities in Hospital Demand." National Bureau of Economic Research Working Paper 14619.

Grumbach, Kevin, Janet Coffman, Karen Vranizan, Noelle Blick, and Edward H. O'Neil. 1998a. "Independent Practice Association Physician Groups in California." *Health Affairs* 17 (3): 227–37.

Grumbach, Kevin, Denis Osmond, Karen Vranizan, Deborah Jaffe, and Andrew B. Bindman. 1998b. "Primary Care Physicians' Experience of Financial Incentives in Managed-Care Systems." *New England Journal of Medicine* 339 (21): 1516–21.

Hammelman, Eric, Narda Ipakchi, Jennifer Snow, and Bob Atlas. 2009. "Reforming Physician Payments: Lessons from California." California Healthcare Foundation. www.chcf.org/publications/2009/09/reforming-physician-payments-lessons-from-california (accessed October 8, 2014).

Ho, Katherine. 2006. "The Welfare Effects of Restricted Hospital Choice in the US Medical Care Market." *Journal of Applied Econometrics* 21 (7): 1039–79.

Ho, Katherine, and Ariel Pakes. 2011. "Do Physician Incentives Affect Hospital Choice? A Progress Report." *International Journal of Industrial Organization* 29 (3): 317–22.

Ho, Kate, and Ariel Pakes. 2014. "Hospital Choices, Hospital Prices, and Financial Incentives to Physicians: Dataset." *American Economic Review.* http://dx.doi.org/10.1257/104.12.3841.

Kessler, Daniel P., and Mark B. McClellan. 2000. "Is Hospital Competition Socially Wasteful?" *Quarterly Journal of Economics* 115 (2): 577–615.

Ketcham, Jonathan D., Pierre-Thomas Léger, and Claudio Lucarelli. 2012. "Standardization Under Group Incentives." Unpublished.

Limbrock, Frank. 2011. "Pecuniary and Non-Pecuniary Incentives in Prescription Pharmaceuticals: The Case of Statins." *B.E. Journal of Economic Analysis and Policy* 11 (2): 1–28.

Luft, Harold S., Deborah W. Garnick, David H. Mark, Deborah J. Peltzman, Ciaran S. Phibbs, Erik Lichtenberg, and Stephen J. McPhee. 1990. "Does Quality Influence Choice of Hospital?" *Journal of the American Medical Association* 263 (21): 2899–906.

McClellan, Mark B. 2011. "Reforming Payments to Health Care Providers: The Key to Slowing Healthcare Cost Growth While Improving Quality?" *Journal of Economic Perspectives* 25 (2): 69–92.

Melichar, Lori. 2009. "The Effect of Reimbursement on Medical Decision Making: Do Physicians Alter Treatment in Response to a Managed Care Incentive?" *Journal of Health Economics* 28 (4): 902–7.

Melnick, Glenn A., and Jonathan D. Ketcham. 2008. "Have HMOs Broadened Their Hospital Networks? Changes in HMO Hospital Networks in California, 1999–2003." *Medical Care* 46 (3): 339–42.

Muhlestein, David. 2013. "Continued Growth of Public and Private Accountable Care Organizations." *Health Affairs Blog.* http://healthaffairs.org/blog/2013/02/19/continued-growth-of-public-and-private-accountable-care-organizations/ (accessed October 8, 2014).

Neyman, J., and Elizabeth L. Scott. 1948. "Consistent Estimates based on Partially Consistent Observations." *Econometrica* 16 (1): 1–32.

Office of Statewide Health Planning and Development. 2003. "Patient Discharge Private Use Dataset."

Office of Statewide Health Planning and Development. 2003. "Private Use Birth Cohort File."

Office of Statewide Health Planning and Development. 2003. "Annual Financial Data Reports." http://oshpd.ca.gov/HID/DataFlow/HospData.html (accessed October 2014).

Pakes, Ariel. 2010. "Alternative Models for Moment Inequalities." *Econometrica* 78 (6): 1783–822.

Pakes, Ariel, Jack Porter, Kate Ho, and Joy Ishii. 2011. "Moment Inequalities and Their Application." Unpublished.

Robinson, James C. 2003. "Hospital Tiers in Health Insurance: Balancing Consumer Choice with Financial Motives." *Health Affairs* W3: 135–46.

Robinson, James C., and Lawrence P. Casalino. 2001. "Reevaluation of Capitation Contracting in New York and California." *Health Affairs Supplemental Web Exclusives:* W11–9.

Rosenbaum, Paul R., and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70 (1): 41–55.

Rosenthal, Meredith B., Richard G. Frank, Joan L. Buchanan, and Arnold M. Epstein. 2001. "Scale and Structure of Capitated Physician Organizations in California." *Health Affairs* 20 (4): 109–19.

Rosenthal, Meredith B., Richard G. Frank, Joan L. Buchanan, and Arnold M. Epstein. 2002. "Transmission of Financial Incentives to Physicians by Intermediary Organizations in California." *Health Affairs* 21 (4): 197–205.

Song Zirui, Dana Geib Safran, Bruce E. Landon, Yulei He, Randall P. Ellis, Robert E. Mechanic, Matthew P. Day, and Michael E. Chernew. 2011. "Health Care Spending and Quality in Year 1 of the Alternative Quality Contract." *New England Journal of Medicine* 365 (10): 909–18.

Sorensen, Alan T. 2003. "Insurer-Hospital Bargaining: Negotiated Discounts in Post-deregulation Connecticut." *Journal of Industrial Economics* 51 (4): 469–90.

State of California Department of Managed Health Care. 2003. "Annual Financial Reporting Forms." https://www.dmhc.ca.gov/LicensingandReporting/HealthPlanFilingsandReporting.aspx#.VDVSS_IdV8E (accessed November 2013).

Town, Robert J., and Gregory S. Vistnes. 2001. "Hospital Competition in HMO Networks." *Journal of Health Economics* 20 (5): 733–53.

Yegian, Jill M. 2003. "Tiered Hospital Networks: Reflections from the California HealthCare Foundation." *Health Affairs* W3: 147–53.

This article has been cited by:

1. Jessica Calfee Stahl. 2016. Effects of Deregulation and Consolidation of the Broadcast Television Industry. *American Economic Review* **106**:8, 2185-2218. [Abstract] [View PDF article] [PDF with links]

2. Liran Einav, Amy Finkelstein, Heidi Williams. 2016. Paying on the Margin for Medical Care: Evidence from Breast Cancer Treatments. *American Economic Journal: Economic Policy* **8**:1, 52-79. [Abstract] [View PDF article] [PDF with links]

3. Martin Gaynor, Kate Ho, Robert J. Town. 2015. The Industrial Organization of Health-Care Markets. *Journal of Economic Literature* **53**:2, 235-284. [Abstract] [View PDF article] [PDF with links]

4. Gautam Gowrisankaran, Aviv Nevo, Robert Town. 2015. Mergers When Prices Are Negotiated: Evidence from the Hospital Industry. *American Economic Review* **105**:1, 172-203. [Abstract] [View PDF article] [PDF with links]