Penn Institute for Economic Research
Department of Economics
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19104-6297
pier@econ.upenn.edu
http://economics.sas.upenn.edu/pier

PIER Working Paper 13-057

"Regulation and Capacity Competition in Health Care: Evidence from Dialysis Markets"

by

Mian Dai and Xun Tang

http://ssrn.com/abstract=2340061

Regulation and Capacity Competition in Health Care: Evidence from Dialysis Markets

Mian Dai* Xun Tang†

June 28, 2013

Abstract

This paper studies entry and capacity decisions by dialysis providers in the U.S. We estimate a structural model where providers make strategic continuous choices of capacities based on private information about own costs and beliefs about competitors' behaviors. We evaluate the impact on market structure and provider profits under counterfactual regulatory policies that increase per capacity cost or reduce per capacity payment. We find that these policies reduce the market capacity of dialysis stations. However, the downward sloping reaction curve shields some providers from negative profit shocks in certain markets. The paper also has a methodological contribution in that it proposes new estimators for Bayesian games with continuous actions, which differ qualitative from discrete Bayesian games such as those with binary entry decisions.

JEL Classification: L11, L13, I11

Keywords: Bayesian Games with Continuous Actions, U.S. Dialysis Market

*Department of Economics, LeBow College of Business, Drexel University, 3600 Market Street, Philadelphia, Pennsylvania,19104. E-mail: md598@drexel.edu. Phone: (215) 895-2536.

†Department of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, Pennsylvania 19104. Email: xuntang@econ.upenn.edu. Phone: (215) 898-7409.

1 Introduction

Dialysis is the major treatment for more than 630,000 patients in the U.S. who suffer from End Stage Renal Disease (ESRD). Medicare, the monopolistic buyer of dialysis services, has spent 8.6 billion dollars in 2007 on the treatment and medication of dialysis patients. While Medicare pays a fixed rate to dialysis providers, rising dialysis expenditure motivates a recent Medicare reform that aims at reducing dialysis cost and maintaining the quality of care.[1] The core component of the reform is a new payment system that incorporates the payments for the drugs and services furnished in a dialysis session into a single bundled rate. This effectively reduces the average per-treatment payment rate and lowers the per-patient margin for dialysis providers. On the other hand, insufficient payment may compromise dialysis accessibility. Dialysis capacity, as measured by the number of dialysis stations, is an important metric used by policy makers to evaluate the adequacy of dialysis payment. In this paper, we analyze the provision of dialysis capacity and use it to evaluate the implications of counterfactual dialysis payment policies. Our results offer some insights into the conduct of health care providers and helps us better understand the effectiveness of fixed-price regulation.

We build a model of static Bayesian games with *continuous* actions to examine the strategic interactions between U.S. dialysis providers in their capacity choices across Hospital Service Areas (HSA). We focus on three types of providers on the market: FMC, DaVita and all other non-Chain providers.[2] We estimate how the payoffs of these providers depend on exogenous market conditions (such as size, ESRD risks etc) as well as the endogenous choice of capacities. We then apply our estimates to address counterfactual policy questions, such as understanding the differential impact of payment policies that either raise or reduce the per-capacity margin. In particular, we are interested in how different providers would respond to positive or negative adjustment in reimbursement for dialysis and whether high capacity helps a dialysis provider maintain market presence with a lower margin.

The importance of capacity choices on the market structure and the intensity of competition have been emphasized under a variety of theoretical contexts (e.g. Dixit 1980, Gelman and Salop 1983, Kreps and Scheinkman 1983 etc). On the other hand, the empirical literature has largely ignored the strategic incentives in the continuous capacity choice and its impact on competitors' behavior. With little price and quality competition (Grieco and McDevitt 2013 and Cultler, Dafny and Ody 2012), these strategic incentives are important for dialysis providers.[3] In practice, a dialysis provider's

[1] The reform was proposed in 2008 and became effective in 2011. The full implementation of the policy is expected to complete in 2014.

[2] FMC and DaVita are the industry leaders with national footprints. They jointly own 2/3 of the dialysis facilities and treat more than 2/3 of all dialysis patients. All other non-FMC or non-DaVita providers are referred to as non-Chain hereinafter. Their operation scale and market penetration of these non-chain providers are not comparable to the leaders.

[3] Both Grieco and McDevitt (2013) and Cultler, Dafny and Ody (2012) find little effect of competition on various

margin decreases in the capacity of competitors, because additional capacity allows competitors to offer a more flexible treatment schedule, which is highly valued by patients.[4] Given the high operation and maintenance costs for each unit of capacity, the optimal capacity choices depend on tradeoffs between the market demand, the capacity costs and the competitive interactions between dialysis providers.

We characterize providers' strategies in a Bayesian Nash equilibrium by the first-order conditions of their constrained maximization problems (subject to the constraints of non-negative capacities). We show that these conditions take the form similar to censored regressions, except that expected capacities for competitors now enter as "generated" regressors. We propose a GMM estimator to identify and infer marginal effects of own and competitor capacities as well as market characteristics on a provider's profits.

Our GMM estimator has two desirable features. First, it fully exploits the structural relations in the model (i.e. fixed-point equations that characterize ex ante capacities in equilibrium). Hence, it's more efficient than the classic two-step estimators used for discrete Bayesian games, where the structural parameters are estimated from a reduced-form representation in the second step (e.g. Bajari, Hong, Krainer and Nekipelov 2010).[5] We first summarize the structural equilibrium relations between parameters of interests and observed outcomes into a likelihood function. We then include the first-order conditions for maximizing the likelihood as the first set of moments used in estimation. Second, our estimator retains the advantage of the two-step estimators over conventional nested fixed-point maximum-likelihood estimators. That is, we use data to guide the equilibrium selection by including the second set of moments that match the equilibrium predictions of expected capacities with those directly identifiable and estimated from data. Therefore, our estimator manages to remain agnostic about the equilibrium selection mechanisms.[6]

Our estimates conform to the empirical regularities that a dialysis provider's capacity choice is decreasing in those of the competitors' (i.e. reaction curve is downward sloping) and that competition is more intense between competitors with high capacities. We find that, all else being equal, a capacity increase of one unit decreases a rival's equilibrium choice of capacity by an average of 0.2-0.4 units and

measures of dialysis quality and patient outcomes.

[4] Given that a typical dialysis patient visits providers three times a week for a total of nine to twelve hours, the ease of scheduling is found to be more important than survival in patient's dialysis choice (Johansen 2011). It's easier for patients to find their preferred appointments with the providers capable of performing multiple concurrent dialysis sessions.

[5] In a classic two-step approach for estimating simultaneous entry games, players' choice probabilities in equilibrium is nonparametrically estimated from the data in the first step. Then in the second step, payoff parameters are estimated from the structural equation that characterizes players' optimal decisions, with the equilibrium probabilities therein replaced by first-step estimates.

[6] Nevertheless, we do require the conventional assumption that the data generating process is rationalized by a unique equilibrium.

decreases its entry probability by 3%-13%.[7] This suggests that focusing on the binary entry decisions has overlooked rather substantial strategic interactions between competitors' capacity choices. Our results are robust to various sample selection criteria and econometric specifications (such as two-stage, nested maximum-likelihood and GMM).

In our counterfactual analyses, we find that the average providers' capacity choices respond negatively to a reduction in profit margin that results from a more stringent reimbursement policy. The effects are reversed under more generous reimbursement policies that increase the profit margin. More interesting, the responses are heterogeneous across markets and providers. For example, FMC and non-chain independent providers respond most strongly to both negative and positive margin adjustments while DaVita reacts mildly in both cases. In a number of markets, DaVita reduces / increases its capacity when the margin widens / shrinks by the same portion as its competitors. This is driven by the downward sloping reaction curve. It incentivizes DaVita to decrease capacity if the rivals expand aggressively and to increase capacity if rival scales down substantially (despite the positive / negative margin adjustments). Overall, the shape of the reaction curve magnifies a firm's response to small changes in per-capacity margin and plays a crucial role in determining the effect of different reimbursement policies.

Apart from the empirical motivations and findings, our paper contributes to the literature on the econometrics of empirical games. We are not aware of any previous work that structurally estimates a model of static Bayesian games with *continuous* actions. The bulk of existing literature studies competitive effects in games with *discrete* actions in contexts such as entry games, where alternatives available to players are naturally finite (e.g. Mazzeo 2002, Seim 2006, Sweeting 2009 and Davis 2006). In our application, the discrete entry decision is closely associated with a continuous capacity choice, suggesting that even small incremental changes in capacities play an important role in determining firms' payoffs and market outcomes. Indeed, one of the main message of our paper is that overlooking the information revealed in the level of capacities may compromise researchers' understanding of the market mechanism.

The rest of this paper is organized as follows. Section 2 introduces the background of the dialysis industry. Section 3 specifies the econometric model for structural analyses. Section 4 discusses its identification and our methods for estimation. Section 5 describes our data. Section 6 presents the empirical findings. Section 7 concludes with implications for future research.

[7] One unit in capacity is equivalent to one dialysis machine.

2 Background

2.1 Dialysis and Capacity

Chronic Kidney Disease (CKD) affects more than twenty million adults in the United States. The advanced stage of CKD is known as End Stage Renal Disease (ESRD) and is most commonly caused by diabetes and high blood pressure. The only treatment for ESRD is dialysis or a kidney transplant. Given the limited supply of donor organ and the surgical risk due to comorbidity among a portion of population, the majority of ESRD patients rely on routine dialysis as the major treatment.

Dialysis is a therapy that removes waste (such as urea) and excess water from the body as a replacement for lost kidney function. Hemodialysis is the most common treatment modality and accounts for about 90% of the dialysis population in the U.S.[8] In a hemodialysis session, a dialysis machine pumps a patient's blood into the dialyzer, cleans it with dialysate (a solution that removes excess fluids and wastes) and injects cleaned blood back into the patient's body. It is impossible to provide dialysis without those machines.

Acquiring and operating a dialysis machine is quite costly. A new dialysis machine costs between \$10,000 and \$15,000 with a lifespan of five to seven years. There are other associated costs, such as dialysis chairs, private screens, etc. The industry experts estimate a cost of \$100,000 to maintain and operate one dialysis station. The dialysis capacity provided in a given market (HSA) is practically a permanent decision for dialysis providers, as data report little subsequent adjustment in capacities by providers following initial entry into a market. Grieco and McDevitt (2013) finds that dialysis capacity remains constant for over 90% of the dialysis facilities in the U.S. between 2004 and 2007.

2.2 Regulatory Background

The ESRD patients receive almost universal coverage under Medicare regardless of their age.[9] Around 80% of the dialysis population relies on Medicare as the primary payer (USRDS 2010). Under the old system, Medicare reimbursed three dialysis sessions per-week under a fixed rate (after adjusting for patient's case-mix, local wages and other factors associated with the cost of treatment). In addition, providers are also paid for separately billable services that are furnished during the in-center hemodialysis sessions (e.g. injectable drugs such as Epogen and diagnostic laboratory tests), which represent about 40 percent of total Medicare payments per dialysis treatment session. The generous

[8] Alternatively, about 10% of the dialysis population chooses peritoneal dialysis which is usually performed everyday by patients at home.

[9] ESRD was recognized as disability under the Medicare Reform Act in 1972. The legislation was signed into law and became effective in 1973.

reimbursement for separately billable services has raised concern that it may create distorted profit incentives for over-utilization. For example, one of the separately billable drugs known as Epogen, primarily used by ESRD patients, cost $2.1 billion in 2008 and has become Medicare's largest drug expenditure.[10] DaVita, one of the largest chain dialysis providers, was investigated for overbilling dialysis drugs.[11] Thamer et al. (2007) find that large for-profit chain facilities used larger dose of Epogen and suspect that the profit incentive is responsible for the outcome. The excessive use of drugs such as Epogen not only increases the Medicare expenditure, it also raises cardiovascular risk (e.g. heart attack, stroke etc.) and subsequently lowers the quality of life of ESRD patients.

In 2008, Medicare proposed a new payment system and eliminated the drug incentive by incorporating separate billable items into an expanded bundled payment.[12] Additionally, pay for performance quality incentive were introduced under the new system. Dialysis providers whose dialysis quality measures (namely, patient's hemoglobin and urea levels) [13] did not meet standards could be penalized with payment rate reduction of up to 2 percent. The new dialysis payment system became effective in 2011 and the full implementation is expected by 2014.

The new Medicare reimbursement rule could have a significant impact on dialysis providers. The Government Accountability Office estimates an $880 million saving on dialysis payments.[14] Our counterfactual experiments is motivated by this reform and investigate how dialysis providers respond to different adjustments in the reimbursement rate.

2.3 Dialysis Market

The dialysis market in the U.S can be characterized as a duopoly. In 2007, DaVita and Fresenius Medical Care (FMC), the two largest national chains, jointly treated over 66% of dialysis patients (31% by DaVita v.s. 35% by FMC) and owned around 66% of the dialysis facilities (30% by DaVita v.s. 36% by FMC).[15] The national chains grew significantly after a series of consolidations in the 2000s. In 2004, DaVita bought Gambro who owned over 550 facilities while in 2005, FMC bought Renal Care Group with more than 450 facilities.[16] Overall, the mergers between 2004 through 2006 consolidated the six largest chains into just two. The market structure was relatively stable between 2007 and 2010.

[10] Epogen treats anemia, a common complication of ESRD.

[11] In 2006, 25% of DaVita's revenue came from Epogen. The government decided not to pursue the case in 2011.

[12] For example, the Medicare base rate per dialysis session was $133.81 while it was $229.63 under the 2008 proposal. Note that the proposed base rate incorporates all separable billable services including lab test and injectable drugs.

[13] Hemoglobin is an indicator for whether patient's anemia is under management while urea reduction ratio is an indicator for dialysis adequacy.

[14] Source: http://www.ama-assn.org/amednews/2012/12/24/gvsd1226.htm

[15] Source: USRDS Atlas of ESRD 2009, Chapter 10.

[16] The merger between FMC and Renal Care Group was annunced May 2005 and completed in March, 2006.

In our model, we focus on three types of dialysis providers: FMC, DaVita and all other providers (referred to as non-Chain). Given the proximity in time frame when these consolidations and market reorganization occur between 2004 and 2006, we believe that the decisions of providers can be treated as made simultaneously.

3 A Model of Capacity Choices with Private Information

We now specify a model of simultaneous Bayesian games with continuous actions. Consider a market that is served by N providers competing through choices of their capacities. A provider (firm) i's profit from providing capacity K_i in market m is given by:

$$\Pi_{i,m}(K_{i,m}, K_{-i,m}, X_m, \varepsilon_{i,m}) = K_{i,m}\pi_i(X_m, K_{-i,m}, \varepsilon_{i,m}) - c_i(K_{i,m}).$$

The function $\pi_i(X_m, K_{-i,m}, \varepsilon_{i,m})$ is the per capacity revenue for i. It depends on the vector of competitors' capacities $K_{-i,m} \equiv (K_{j,m})_{j\neq i}$, the market characteristics X_m, and an idiosyncratic profit component $\varepsilon_{i,m}$ which is private information only known to provider i. We assume the private information $\varepsilon_{i,m}$ are independent from each other conditional on X_m. To simplify notation, we drop subscript m below.

A provider's revenue is proportional to its choice of capacity in the specification above. This is motivated by the observation that each dialysis machine receives a flat rate from Medicare for each treatment. A typical dialysis patient receives three treatments per week, each lasting for about four hours. An additional hour is needed for setting up and cleaning the machine per treatment. This sums to 15 hours per week. On average, a dialysis machine treats 3-5 patients per week. During a treatment, operating staff such as registered nurses and technicians have to overlook the patients and perform routine checks. Given the fixed dialysis price, it is plausible that a provider's revenue is approximately proportional to its dialysis capacity.

We adopt a linear specification of per capacity revenue π_i:

$$\pi_i(X, K_{-i}, \varepsilon_i) = X\beta_i + \textstyle\sum_{j\neq i}\gamma_{i,j}K_j - \varepsilon_i$$

where $\gamma_{i,j}$ are heterogeneous marginal effects of j's capacities on i's per capacity revenues. There are two reasons for this simple specification. First, this linear specification for π_i can be interpreted as a practical reduced-form approximation (regression) of actual per capacity revenue in the data-generating process. Using this as a benchmark helps us to better understand the strategic role of capacity choices in determining market outcomes and firms' profits. As we show in Section 6, this simple specification already manages to explain a large portion of the variation in capacity choices observed from the data. Second, focusing on this linear specification allows us to establish the identification of marginal effects of market characteristics on latent profits. In principle, we could extend our estimation algorithm in Section 4 to a richer structural model (such as one with market-level unobserved heterogeneity) to attain a better fit for data. Yet this is known to raise new challenges with identification of structural elements, including the marginal effects of capacities and market characteristics on profits.

The firm-specific fixed cost is specified as:

$$c_i(K_i) = a_i K_i^2 + b_i K_i, \text{ where } a_i > 0.$$

We adopt a quadratic cost specification for the following reasons. First, the assumption of constant scale of economy (i.e. costs are linear in capacity) is not plausible in the dialysis industry. Adding a dialysis station not only involves significant investment, but also requires additional space, maintenance and personnel (e.g. technician and nurses) etc. The supplies of these inputs are usually not very elastic. The quadratic form is intended to capture such diseconomies of scale. Second, quadratic cost takes a flexible nonlinear form, and can be considered as a second-order polynomial approximation to a more complicated cost structure. Our estimation algorithm below can be extended to allow for higher order polynomials in the specification. Finally, that there is no constant term in the quadratic function is due to the need for a location normalization: i.e., profits from no entry ($K_i = 0$) need to be zero.

To conduct structural analyses, we assume capacities observed from the data are rationalized by providers' pure-strategy Bayesian Nash equilibria. A pure strategy for a provider i is a mapping from its information set (X, ε_i) into the support for capacities (i.e. $\mathbb{R}_+$). A pair of pure-strategies $\{K_i^*(.)\}_{i \in N}$ forms a pure-strategy Bayesian Nash equilibrium (PSBNE) if:

$$K_i^*(X, \varepsilon_i) = \arg\max_{K_i \in \mathbb{R}_+} \mathbb{E}_{\varepsilon_{-i}} \left[\Pi_i(K_i, K_{-i}^*(X, \varepsilon_{-i}), X, \varepsilon_i) | X, \varepsilon_i \right] \tag{1}$$

where $K_{-i}^*(X, \varepsilon_{-i})$ is a shorthand for $\{K_j^*(X, \varepsilon_j)\}_{j \neq i}$. The case with $K_i = 0$ means provider i decides not to enter the market. Existence of a pure-strategy Bayesian Nash equilibrium in our model follows from Theorem 3 in Athey (2001) and the fact that the cross-derivatives of the *ex post* profits for i with respect to (K_i, K_j) and (K_i, ε_j) are constants.

We now derive the first-order condition of the equilibrium as a foundation for our maximum likelihood estimator. We maintain a popular regularity condition that the order of differentiation and integration in $\frac{\partial}{\partial K_i}\mathbb{E}_{\varepsilon_{-i}}\left[\Pi_i(K_i, K_{-i}(X,\varepsilon_{-i}), X, \varepsilon_i)|X,\varepsilon_i\right]$ can be switched for all i and vectors of admissible rival strategies $K_{-i}(X,\varepsilon_{-i})$.

Proposition 1 *Under the model assumptions above and a regularity condition that the order of differentiation and integration in* $\frac{\partial}{\partial K_i}\mathbb{E}_{\varepsilon_{-i}}\left[\Pi_i(K_i, K_{-i}(X,\varepsilon_{-i}), X, \varepsilon_i)|X,\varepsilon_i\right]$ *can be interchanged,*

$$K_i^*(X,\varepsilon_i) = \max\left\{0, \tfrac{1}{2a_i}\left(X\beta_i + \textstyle\sum_{j\neq i}\gamma_{i,j}\mathbb{E}_{\varepsilon_j}\left[K_j^*(X,\varepsilon_j)|X,\varepsilon_i\right] - b_i - \varepsilon_i\right)\right\} \tag{2}$$

in any PSBNE.

This proposition has a couple of key implications for estimation. First, it implies the scale of $a_i, b_i, \beta_i, \gamma_{i,j}$ and the distribution of ε_i cannot be jointly recovered from (2). Hence without loss of generality we set $a_i = 1/2$ by way of a scale normalization that is necessary for estimation. Second, the equilibrium condition in (2) is similar to a single-agent censored regression, except that a subvector of its regressors now consists of equilibrium objects $\{\mathbb{E}_{\varepsilon_j}[K_j^*|x,\varepsilon_i]\}_{j\neq i}$. Thus the model lends itself to standard maximum likelihood estimation of censored regressions with generated regressors.

We conclude this sections with further discussions that justify our modeling choices by the distinctive institutional details of the dialysis services market. First, providers' choices of capacity are essentially continuous. Our data show such capacities by a provider in a market range from zero to just under sixty. Thus it is infeasible to apply the typical multinomial choice model to analyze capacity decisions with such a large number of actions. Second, dialysis providers rarely adjust their capacities after the initial entry (Grieco and McDevitt 2013). This indicates that binary market entry decision is made practically in combination with a continuous choice of capacity. Both decisions are de facto permanent, based on a provider's expectation about market profitability. Finally, the game is one with incomplete information, because the dialysis providers have little information about idiosyncratic components in competitors' profits.

4 Econometric Methods

We sketch a proof of identification for coefficients $\beta_i, \gamma_{i,j}, b_i$ in (2) using a typical two-step argument. First, under the assumption that private information is independent across players conditional on market characteristics, player i's expectation for competitors' capacities in (2) is a function of commonly observed market characteristics X alone. With data rationalized by a single profile of equilibrium $\{K_i^*(.)\}_{i\in N}$, this function is directly identifiable as the expectation of K_j^* conditional on X.

With the scale normalization that $2a_i = 1$, (2) becomes

$$K_i^*(X, \varepsilon_i) = \max\left\{0, X\beta_i + \textstyle\sum_{j\neq i}\gamma_{i,j}\varphi_j(X) - b_i - \varepsilon_i\right\} \tag{3}$$

for all i, where $\varphi_i(X) \equiv \mathbb{E}_{\varepsilon_j}[K_j^*(X, \varepsilon_j)|X]$ is directly identifiable from the data. With the distribution of ε_i parameterized (e.g. as a normal or a logistic distribution), the joint identification of $\beta_i, \gamma_{i,j}, b_i$ and parameters in the distribution of ε_i follows from typical arguments for parametric Tobit models, provided the vector of $(X, \{\varphi_j(X)\}_{j\neq i})$ in equilibrium demonstrates sufficient variation (i.e. their joint support satisfies a mild full-rank condition).[17]

This identification strategy leads to the following two-step estimator: In the *first* step, nonparametrically estimate the expectation of competitors' equilibrium capacities by $\hat{\varphi}_i(X)$. This could be done using kernel estimators (with either the local constant or the polynomial approach), or using sieves estimators with polynomial basis. We adopt the latter approach in the empirical implementation in Section 6. That is,

$$\hat{\varphi}_j(x_g) \equiv \min_{\{\alpha_s\}_{0\leq s\leq S}} \tfrac{1}{G}\textstyle\sum_{g=1}^{G}\left[k_{g,j} - \sum_{s=0}^{S}\alpha_s x_g^s\right]^2 \tag{4}$$

where g is an index for the G independent games (markets) observed in data; and $k_{g,j}, x_g$ are realizations of K_j, X in market g. We set the order S to four in estimation. Alternatively, to reduce computational costs in estimation, one may choose to replace these first-step nonparametric estimates with those from a reduced-form Poisson regression. A Poisson regression works well in practice when the empirical distribution of dependent variables in the data are close in shape to some distributions from the exponential family. (See Cameron and Trivedi (1998) and Christensen (1997) for details.) In such cases, Poisson regressions are known to provide good approximations in terms of model fit especially when the dependent variables are continuous or count data, as is the case in our application.

In the *second* step, use a maximum likelihood estimator where the distribution of ε_i is parametrized (e.g., $\varepsilon_i\tilde{\ }N(0, \sigma_i^2)$ for all i, where σ_i^2 are constant parameters to be estimated). Specifically, let $\theta \equiv (\theta_i)_{i\in N}$ where $\theta_i \equiv (\beta_i, \{\gamma_{i,j}\}_{j\neq i}, b_i, \sigma_i)$. Our estimator is defined as

$$\hat{\theta}^{TS} \equiv \arg\max_\theta \tfrac{1}{G}\log \hat{L}_G(\theta)$$

[17] In fact, identification of coefficients $\beta_i, \gamma_{i,j}, b_i$ can be shown for (3) under nonparametric stochastic restrictions on ε_i instead of parametric assumptions. Examples of these stochatic restrictions include: independence between ε_i and X as in Buckley and James (1979) and Horowitz (1986); conditional symmetry as in Powell (1986); and median independence in Powell (1984).

where

$$\hat{L}_G(\theta) \equiv \prod_{g\leq G}\prod_{i\in N} \hat{f}_i(k_{g,i}|x_g;\theta_i); \text{ and} \tag{5}$$

$$\hat{f}_i(k_i|x;\theta_i) \equiv \left\{1-\Phi\left(\frac{x\beta_i+\sum_{j\neq i}\gamma_{i,j}\hat{\varphi}_j(x)-b_i}{\sigma_i}\right)\right\}^{1(k_i=0)}\left\{\frac{1}{\sigma_i}\phi\left(\frac{k_i-x\beta_i-\sum_{j\neq i}\gamma_{i,j}\hat{\varphi}_j(x)+b_i}{\sigma_i}\right)\right\}^{1(k_i>0)}.$$

With the number of basis used in the first step polynomial estimation expanding at an appropriate rate as the sample size increases, the preliminary estimate $\hat{\varphi}_j(.)$ converges to the true function uniformly at a rate fast enough to maintain the root-n asymptotic normality of the second step MLE estimators.

It is possible to improve estimation efficiency by further exploiting the fixed-point characterization of ex ante equilibrium capacities in the second step. To this end, we also propose a General Method of Moment (GMM) estimator, which exploits the structural fixed-point equation defining competitors' expected capacities in PSBNE: For all $i \in N$,[18]

$$\varphi_i(X) \equiv \mathbb{E}_{\varepsilon_i}[K_i^*(X,\varepsilon_i)|X] = \mathbb{E}_{\varepsilon_i}\left[\max\left\{0\ ,\ X\beta_i+\sum_{j\neq i}\gamma_{i,j}\varphi_j(X)-b_i-\varepsilon_i\right\}\Big|\, X\right]. \tag{6}$$

More specifically, the GMM estimator is:

$$\hat{\theta}^{GM} \equiv \arg\min_\theta \hat{M}_G'(\theta)\hat{W}_G\hat{M}_G(\theta)$$

where the empirical moments are:

$$\hat{M}_G(\theta) \equiv \left[\tfrac{1}{G}\nabla_\theta \log \hat{L}_G(\theta)\ \ ;\ \ \tfrac{1}{G}\sum_{g\leq G}\sum_{i\in N}\left[\tilde{\varphi}_i(x_g;\theta)-\hat{\varphi}_i(x_g)\right]^2\right] \tag{7}$$

where $\hat{L}_G(\theta)$ is defined in (5); $\hat{W}_G$ is a consistent estimator for the optimal GMM weight matrix; $\hat{\varphi}_i(x)$ is the first-stage non-parametric estimates for expected capacities in equilibrium (e.g. a polynomial approximation defined in (4)); and $\tilde{\varphi}_i(x;\theta)$ is a solution for $\{\varphi_i(x)\}_{i\in N}$ in (6) at $X = x$ given the vector of parameters θ. Alternatively, in order to reduce computational costs, one can also choose to

[18] Under normality assumption of ε, (6) has a closed form

$$\varphi_i = \Phi\left(\frac{x\beta_i+\gamma_i\varphi_{-i}-b_i}{\sigma_i}\right) * \left\{x\beta_i+\gamma_i\varphi_{-i}-b_i+\sigma_i\frac{\phi\left(\frac{x\beta_i+\gamma_i\varphi_{-i}-b_i}{\sigma_i}\right)}{\Phi\left(\frac{x\beta_i+\gamma_i\varphi_{-i}-b_i}{\sigma_i}\right)}\right\}$$

for $i \in N$, where φ_i and φ_{-i} are i and its rival's expected capacities.

replace $\hat{\varphi}_i(.)$ with the fitted values of expected capacities based on a reduced-form Poisson regression. We adopt this alternative in our estimation later. In Section 6, we first obtain initial GMM estimates by setting the weight matrix to be the identity matrix. We then estimate the optimal weight matrix $\hat{W}_G$ using the initial estimates, and apply it to weighted GMM to improve estimation efficiency.

For the purpose of comparisons, we also report in Section 6 the performance of a third estimator $\hat{\theta}^{FL}$, which is based on a full nested fixed-point maximum-likelihood approach. This estimator amounts to replacing the first-step estimates $\hat{\varphi}_i$ in the two-step estimator $\hat{\theta}^{TS}$ by $\tilde{\varphi}_i(x;\theta)$, i.e. solutions for $\{\varphi_i(x)\}_{i\in N}$ in (6) at $X = x$. Such an estimator is feasible under our specification of π_i and parametrization of the private information distribution. The nested maximum-likelihood estimator used in our case is analogous to that applied widely to dynamic discrete choice models (e.g. Rust 1987).

We conclude this section on econometric methods with several remarks regarding the issue of multiple equilibria and the comparison between the three estimators $\hat{\theta}^{TS}$, $\hat{\theta}^{FL}$ and $\hat{\theta}^{GM}$.

Remark 1. For any given market characteristics x and a vector of parameters θ, there could potentially be multiple solutions for $\{\varphi_i(x)\}_{i\in N}$ in (6). However, such multiplicity does not affect the validity of our two-stage estimator under the common assumption that the data are rationalized by a single Bayesian Nash equilibrium (BNE) given market characteristics. This is because, under this assumption, the expected capacities that enter the likelihood are directly identified from data rather than solved for from (6). On the other hand, the full nested fixed-point MLE estimator is susceptible to the issue of multiplicity in equilibria. While implementing this estimator, we use an algorithm of "Mathematical Programming with Constrained Maximization" (MPEC) for the maximization routine. This algorithm is known to implicitly deal with the multiplicity issue through an effective ad hoc procedure, which essentially always picks an equilibrium that maximizes the likelihood. More detailed discussions are included in Section 6.1.

Remark 2. The afore-mentioned multiplicity also does not affect the validity (consistency) of our GMM estimator under the assumption of a single BNE in the data, as we have incorporated the following ad hoc procedure in the calculation of $\hat{\theta}^{GM}$. Suppose for some (x_g, θ) the system of equations in (6) admits multiple solutions of the vector $\tilde{\varphi}(x_g;\theta) \equiv \{\tilde{\varphi}_i(x_g;\theta)\}_{i\in N}$ (which are often picked up by experimenting with multiple initial points while solving the nonlinear fixed-point equation in (6)). In such cases, choose the vector of $\tilde{\varphi}(x_g;\theta)$ that minimizes the empirical moments in (7) while evaluating the objective function of GMM. To see how such a procedure maintains the consistency of $\hat{\theta}^{GM}$ under multiple equilibria, note the second set of moments in (7) takes a form similar to the objective function of a minimum-distance estimator. Thus, this procedure is effectively using the directly identifiable

$\mathbb{E}[K_i^*(X, \varepsilon_i)|X]$ to guide our choices of equilibrium-implied expected capacities while implementing GMM.

Remark 3. The two-step estimator $\hat{\theta}^{TS}$ and the full nested fixed-point ML estimator $\hat{\theta}^{FL}$ have respective advantages. As explained in Remark 1, $\hat{\theta}^{TS}$ is robust to the issue of multiple equilibria but does not explicitly use the structure in the fixed-point characterization of ex ante capacities in equilibrium (i.e. the φ_i's). In contrast, $\hat{\theta}^{FL}$ is explicit in exploiting this structural relation defining $\{\varphi_i\}_{i \in N}$, but is potentially susceptible to the issue of multiple equilibria.[19] Therefore, choices between the two should depend on researchers' judgement about the possibility of multiple equilibria. The GMM estimator $\hat{\theta}^{GM}$, on the other hand, provides the benefits of both estimators. Due to the use of the second moments, the GMM estimator not only exploits the structural relations defining ex ante capacities, but also manages to deal with the issue of equilibrium multiplicity under the assumption that choices are rationalized by a single Bayesian Nash equilibrium in the data-generating process.

5 Data Description

5.1 Sample Construction and Market Definition

We construct our sample from the data on dialysis facility compare maintained by the Center for Medicare and Medicaid Services. The CMS receives monthly updates about the characteristics of each facility (e.g. facility name, address, chain affiliation, number of dialysis stations, date of certification etc.) and posts them online every quarter. A similar dataset has been used by several recent studies on the dialysis market including Ramanarayanan and Snyder (2011), Grieco and McDevitt (2013), and Cutler, Dafny and Ody (2012). The key variable of interest in our study is capacity, i.e. the number of dialysis stations.

The market for out-patient dialysis is local in nature. Dialysis patients usually receive three treatment sessions per week, each lasts for about four hours. They are in general unwilling or unable to travel too far. According to MedPAC, the median driving distance between patients and dialysis facility is six miles. Following several other studies on dialysis markets (e.g. Grieco and McDevitt 2013, Culter, Dafny and Ody 2012), we use Hospital Service Area (HSA) to delineate the local market. HSA is compiled by the Dartmouth Atlas from Medicare data on patient's hospital choice, which is relatively self-contained with respect to heath care services. The number of HSAs in the U.S. is roughly equal to the number of U.S counties; however, their boundaries don't generally overlap.

[19] The full-maximum-likelihood estimator should be more efficient than the two-step estimator, provided the identification of parameters holds under the parametrization, and the model always admits a unique PSBNE under all x and θ.

Unfortunately, demographics such as population, age and racial composition are not available at the HSA level. To obtain the market level profit and cost shifters, we assign each HSA to a county based on the population distribution within HSA.[20] Then we use the census's county demographics data (e.g. racial composition, age, income, poverty, size of business payroll etc.) to approximate the population characteristics within an HSA. To minimize the measurement error, we exclude the market if less than 60% of the HSA population is contained in a single county.[21]

We supplement the Census demographic data with hospital and physician capacity data obtained from the Dartmouth Atlas. While the Dartmouth Atlas provides this information only for 2006, it provides a good approximation for our sampling year in 2007, for there were no major shifts in the industry environment between these years. We include the age adjusted prevalence rate of diabetes as proxies for ESRD risks. We also use hospital beds and the number of nephrologists to control for the base demand and intensity of health care. These additional variables explain a significant portion of the variation in dialysis capacity in our later analysis.

Following Ford and Kaserman (1993), who showed the certificate of need (CON) regulation discourage entry by requiring additional regulatory procedure for providers to establish a market presence, we construct a binary indicator for the state level CON regulation. Finally, we use the distance between an HSA and the headquarters of the chains as an extra set of cost shifters.[22,23,24]

The distribution of dialysis capacity is highly skewed. While the average capacity in a given market is about 21, the total dialysis capacity in some markets can be as high as 1309. The high capacity outliers usually lie in heavily populated cities such as Chicago and Los Angeles. In these markets, chain providers often operate multiple branches in close proximity to each other. The nature of competitive interactions are clearly different from an average market. For our analysis, we choose to focus on areas with population between 40,000 and 800,000.[25] We exclude another 124 outliers with total market capacity greater that 60.[26] Our results are robust to alternative sample selection criterion as shown

[20] We decompose each HSA into a collection of zip codes and obtain the population for each zip from the Census. We assign HSA to a county if that county contains the largest proportion of HSA population.

[21] This eliminates 288 HSAs.

[22] We use the HSA boundary file from Dartmouth Atlas to pin point the centroid for the market. The distance from the geographical center of each HSA to the headquarter of either chain is calculated using Haversine formula.

[23] Note that Davita's headquarters changes from El Segundo, CA to Denver, CO. We compute the distance variable based on both locations. Our tables reports the results based on the Denver headquarters. Using CA headquarters doesn't change our results.

[24] While other facility specific cost shifters (e.g. number of FTE employees, compensations etc.) are made available in the Medicare Cost Reports, they generally vary substantially over time and are not likely to influence the time-invariant capacity decision.

[25] This eliminates 1646 HSAs, most of which are sparsely populated areas.

[26] Our current cutoff is approximately the 90th percentile in the capacity distribution. Our results are robust to alternative cutoff values (e.g. the 95th or the 85th percentile in capacity distribution).

in Appendix A.[27]

Our final sample contains 1320 HSA and 1287 facilities in the 48 contiguous states in 2007. In these markets, chain dialysis providers usually open a single branch if entry occurs. One potential concern is that capacity decisions of chain providers are correlated across markets. However, such concern should be minimal for our analysis since our sample selection criterion ensures a set of relatively isolated markets. A back of the envelope spatial analysis shows that the average distance from an in sample FMC facility to its closest FMC neighbor is 12.5 miles while that of DaVita is 12.3 miles.[28] Most of the facilities (287 out of 353 for DaVita and 242 out of 490 for FMC) are not within 10 miles radius of another facility of the same chain. Overall, we don't believe correlated capacity decision to be a significant concern for our analysis.

5.2 Descriptive Statistics

Table 1 summarizes the capacity choices of FMC, DaVita and non-Chain providers. Two significant empirical regularities motivate our model. First, dialysis capacities vary substantially over a wide support. For example, the average capacity of FMC is 6.4 followed by non-Chain and DaVita with the standard deviation ranging between 9.22 and 11.4. While one may still apply a multinomial choice model with fewer choices by arbitrarily grouping capacity levels into relabeled categories (e.g. a multinomial choice model with three alternatives being low, medium and high capacities), this overlooks the information contained in the rich variation of capacities. We believe the continuous choice model better characterizes dialysis providers' capacity decisions.

Second, there is a substantial number of markets in which some providers choose not to enter. FMC, the largest provider, has a presence in about 31% of the markets (followed by non-chain 25% and DaVita 23%). This pattern makes the overall *unconditional* expectation of capacities much smaller than the mean capacity *conditional on entry*. Nevertheless, the standard deviation of capacity choice is similar to that conditional on entry. This suggests that the variation in capacity choices is mostly driven by the specific choices of capacities upon entry.

The correlation between entry and capacity decisions in Table 2 presents descriptive evidence for strategic interactions. Both market presence and dialysis capacities are negatively correlated with that of the rival's. For example, the correlation between FMC and DaVita's entry decisions is –0.12, while the correlation between FMC's capacity and DaVita's entry decision is -0.14. This suggests

[27] As a robustness check, we perform the analysis in larger samples that include almost all HSAs using the two-stage estimator. We obtain results very similar to what we report in the paper. However, it is very computationally costly to apply ML and GMM estimators on larger samples. These results are reported in Table A1 in the appendix.

[28] The distances calculation is based on both in sample and out of sample facilities.

DaVita and FMC generally enter different markets, and DaVita is even less likely to enter when FMC chooses a larger capacity. However, we cannot infer from these aggregate correlation patterns alone that all providers' response curves to competitors' capacities are downward sloping, for this would risk overlooking the heterogeneity in providers' profit and cost structures. The observed aggregate correlation pattern in Table 2 could be driven by a subset of providers and thus not be representative of the other providers. Indeed, a key motivation for our structural analyses in Section 4 is the need to account for such heterogeneity in the reaction curve across providers.

Table 3 presents summary statistics of the market demographics. An average market in our sample is populated by 75,531 residents, 1030 miles away from FMC's headquarter and 1100 miles away from Davita's headquarter. About 23 percent of the markets are located in the northeast region, 24 percent in the mid-west, 17 percent in the west and the remaining 31 percent in the south. The CON regulation is effective in 21 percent of the markets. We employ a parsimonious set of profit and cost shifters including percent of population over 65, racial composition etc. In alternative specifications, we experiment with a larger sets of variables such as poverty rate, income, population density, size of business payroll, number of uninsured, the number of hospital registered nurses, the size and racial mixes of Medicare enrollees etc. These variables do little to explain the variation in dialysis capacity and therefore we did not include them in our main specification.

6 Results

6.1 Details in implementing the estimators

We estimate the model using the estimators described in Section 4. Table 4 presents the results. Panel A, B and C present the results from the two-step estimator $\hat{\theta}^{TS}$, the maximum likelihood (nested fixed point) estimator $\hat{\theta}^{FL}$, and the GMM estimator $\hat{\theta}^{GM}$ respectively. We now provide some further details in the implementation / calculation of these estimators.

In the first step of the two-step estimation, we adopt the Poisson regression approach to estimate the expected capacities in equilibrium. That is, we fit the observed capacity choices to a Poisson distribution, and use it to estimate providers' expected capacities. A Poisson regression is advantageous because the observed capacity choice is non-negative with many observations censored at zero. In a finite sample with moderate size such as ours, a Poisson yields a better fit to the data than the nonparametric alternative of polynomial approximation.[29] In the second step, we use a tobit model

[29] The R^2 in Poisson regression are 31%, 30% and 23% for FMC, DaVita and non-Chain while they are 22%, 22% and 18% in polynomial approximation.

to estimate the effect of $\widehat{\mathbb{E}}(K^*_{-i}|X)$ and X on K_i, where $\widehat{\mathbb{E}}(K^*_{-i}|X)$ is the predicted capacity obtained from the first step. We follow a standard bootstrap procedure to calculate standard errors of the estimates, based on 300 bootstrap samples of size G.

We estimate the nested fixed-point MLE using an optimization strategy called Mathematical Program with Equilibrium Constrain (MPEC), which was introduced by Judd and Su (2011). The standard maximization routines for calculating the nested fixed-point MLE is computationally demanding as it requires solving for equilibrium outcomes defined by the fixed-point mapping (6) in every market for every iteration of parameter values throughout the maximization routine. Besides, the issue of multiplicity arises in such routines, for a given parameter may well admit more than one Bayesian Nash equilibria in general.

In comparison, the MPEC algorithm maximizes the likelihood with respect to both model parameters and providers' strategies (as characterized by expected capacities) in each market subject to the constraints that the expected capacity choices constitute equilibrium defined by our model. As shown by Judd and Su (2011), the solution to the constrained maximum likelihood is equivalent to the solution of the nested fixed point. Thus MPEC is computationally more feasible than nested fixed-point MLE, since the constrained maximization in MPEC doesn't require solving for the non-linear fixed point mapping in every market. For our application, there are 1320 markets with three choice variables in each market. The likelihood is maximized with respect to 3960 more parameters (in addition to the covariates of our empirical specification) subject to 3960 constraints defined by (6). Besides, MPEC is also known to have dealt with the issue of multiple equilibria implicitly: That is, evaluating the likelihood at an MPEC solution is equivalent to evaluating the likelihood at a nested fixed-point MLE solution when the equilibrium selection mechanism is degenerate at the one that yields the highest likelihood. (For more details, see Proposition 1 in Judd and Su (2011) and the subsequent discussions.) The standard errors are obtained through the Hessian of the likelihood function evaluated at our estimates.

To implement our GMM estimator, we use two sets of moment conditions as defined in (7) in Section 4. The first set consists of the first-order condition of the likelihood as defined by (5). The second set of moments matches the model predicted capacity to the ones identified from the data. To reduce computational costs, we use fitted values from a Poisson regression as the first-step estimates $\hat{\varphi}_i$ and use the MPEC algorithm to find the maximizer of our GMM objective function. We then follow a standard two-step approach to estimate GMM: That is, first obtain an initial GMM estimate by setting the weight matrix to be the identity matrix and then use it to compute the optimal weight matrix. We then re-estimate the model by substituting the optimal weight matrix into the GMM objective function. The standard errors for the two-step GMM estimator $\hat{\theta}^{GM}$ are then calculated using the classical approach as in Section 6 of Newey and McFadden (1994).[30]

6.2 Estimates

The estimates from each panel of Table 4 are close in magnitude. The standard errors on many of the GMM estimates are significantly smaller than those on the two-step estimates (e.g. population, nephrologist, diabetes rate, register nurse, CON regulation etc.), especially the standard errors on coefficients of strategic variables. This seems to point to some gains in estimation efficiency from exploiting the equilibrium structure of our model. Since GMM is advantageous over two-stage and fixed point maximum likelihood, we will focus our discussion based on GMM estimates (unless otherwise indicated).

The strategic effect of rival's capacity is strongly significant with a negative sign. This is very robust across different econometric models. The magnitudes of the strategic coefficients suggest that FMC competes aggressively with all other providers. For example, according to the GMM estimates, the effect of FMC on DaVita is about 50% larger than the effect of non-Chain (-2.17 vs. -1.45) while it is 30% larger than the effect DaVita on non-Chain (-0.82 vs. -0.61). Nevertheless, there is no significant evidence that competition is more intense between chain providers.

The strategic effects are quite large. A quick calculation of the marginal effects (similar to those in a Tobit model) suggests that holding other factors at their mean, a one unit increase in DaVita's capacity decreases FMC's capacity by 0.24 units and a one unit increase in non-Chain's capacity decreases FMC's capacity by 0.27 unit. For both DaVita and non-Chain, FMC poses a stronger competitive pressure. A one unit increase in FMC's capacity reduces DaVita's capacity by 0.37 units while it reduces non-Chain's capacity by 0.21 units.

The strategic interactions between dialysis providers also imply a strong effect of capacity choices on rival's entry probability. The marginal effect on the entry probability (i.e. the probability of being uncensored) can be calculated as:

$$\frac{\partial \Pr(K^* > 0|x)}{\partial x} = \frac{\beta}{\sigma}\phi\left(\frac{x\beta + \gamma E(K^*|x)}{\sigma}\right)$$

Our estimate implies that FMC poses the strongest competition to both DaVita and non-Chain. A one unit increase in FMC's capacity decreases the entry probability of DaVita and non-Chain by

[30] This classical approach essentially amounts to stacking all moments used for estimation together (i.e. including those used for estimating the linear coefficients in the Poisson regressions in the first step), and then estimating the covariance matrix based on such an augmented set of moments.

0.03 and 0.01 in an average market, which translates into 12.5 and 3.3 percentage point decreases.[31] Since each provider generally chooses different levels of capacity, a rival's presence should have a non-uniform effect across markets and the competition should be more intense when the rival chooses a higher capacity level. A model that focuses on discrete market entry decisions would have overlooked these heterogeneous competitive interactions.

Most of the coefficients are significant with expected signs. The market size, the extent of diabetic population, the supply of local nephrologists and registered nurses are positively associated with the capacity of all providers while entry barriers such as CON regulation is negatively associated with capacity choice. Some market conditions may be important for one firm but not for others. This implies that dialysis firms target different demographics and it is important to account for firm heterogeneity. Distance to headquarters doesn't seem to be very important to the capacity choice of the chains.[32] But non-chain providers are more likely to offer positive capacity when the market is further away from either chain's headquarters.

To better understand the magnitude of these estimates, Table 5 reports the change in equilibrium capacity and the number of markets with high entry probabilities (i.e. greater than 50%) when some market level variables change. To derive the effect on capacity, we resolve the equilibrium based on equation (6) after adjusting the market level variables upward by 10%. The entry probability of a provider on a market characterized by x is computed as

$$\Pr(K_i^* > 0|x) = \Phi[(x\beta_i + \gamma_i \mathbb{E}[K_{-i}^*|x] - b_i)/\sigma_i] \tag{8}$$

and $E(K^*|X)$ is the new equilibrium capacity after the adjustment of market variables. We also calculate a measure of provider-specific "market penetration" under these hypothetical adjustments. Such a measure is defined as the proportion of markets in the sample where a provider would enter with probability greater than 50% under the adjustment considered.

The effect of market size is quite substantial. When population increases by 10%, the total capacity unambiguously increases by 7%, 9% and 8% for FMC, DaVita and non-Chain while the market penetration measure increases by 12%, 19% and 15%. There are mild increases in capacity and market penetration measures when nephrologists increase by 10%. The effect on non-Chain provider is the smallest. This is probably not too surprising since many non-chain dialysis facilities are owned and operated by the local nephrologist group. By contrast, chain facilities benefit more from a larger

[31] The percentage points are computed based on the mean entry probability of 0.24 and 0.30 for DaVita and non-Chain. The 0.03 points decrease in probability is equivalent to 0.03/0.24=12.5% in percentage points.

[32] Distance to FMC's headquarter is marginally significant for DaVita. All other distance variables are insignificant for chains.

nephrologist stock since such stocks make it easier to locate medical directors with an established referral base. The increase in the prevalence rate of diabetes has heterogeneous effects on different providers. While FMC increases the capacity substantially, non-chain providers reduce the capacity while the capacity of DaVita remains almost constant. This is driven by the intense competitive pressure that FMC exerts on the rivals. Besides, for each additional unit increases in FMC's capacity resulting from the larger diabetic patient base, the competition felt by non-chain providers is strong enough to warrant the capacity reduction through the downward sloping reaction curve. A similar intuition explains that both DaVita and non-Chain respond mildly to the repeal of CON regulation by increasing capacity while FMC increases capacity substantially.

As robustness checks, we perform our analysis on a larger sample with 3129 HSA markets by applying our two-stage estimator and report the results in Panel A of Appendix Table A1. This includes 92% of all HSA in the contiguous US.[33] In addition, we experimented with alternative sample selection criterion by focusing on markets with total capacity less than 80 and 40. The results in Panel B and Panel C of Table A1 remain qualitatively similar to those of Table 4.[34]

6.3 Model Fit

We compute several predictions from our model and compare them to the observed outcomes in the data. Overall, these predictions suggest that our model fits the data quite well.

First, we evaluate the model predicted entry probabilities for each provider and compare them to the observed entry rate. The predicted entry probability is calculated using (8) where $\mathbb{E}[K_i^*|x]$ solves the fixed point mapping defined by (6) evaluated with our estimates in Panel C of Table 4.

Panel A of Table 5 reports the mean and standard deviation of these probabilities. The predicted entry probabilities are very close to the observed entry rates. For example, the entry rate of DaVita is 0.23 while our model predicts an average entry probability of 0.24. We also impute binary entry decision using $I\{Pr(K_i^*|x) > 0) \geq 0.5\}$. Based on such imputations, we find our model correctly predicts the entry decisions in 70%, 78% and 73% of the markets for FMC, DaVita and non-Chain. Overall, our model did well in capturing the censored capacity choice.

Second, we compute the model implied expected capacity choice $\mathbb{E}(K^*|X)$. These expectations are the solution to the fixed mapping evaluated at the estimated parameters. We find the predicted expected capacities are close to the data average. The average observed capacities are 6.4, 4.4 and

[33] We drop the markets where the physician, hospital capacity or demographical information is missing. We also drop about 200 large markets where the total market capacity is greater than 60.

[34] As explained previously, directly estimating with GMM or ML using these (larger) samples exceeds our computational capability.

5.32 for each provider while our model predicts 6.06, 4.23 and 5.08 respectively. This suggests that our model did well in explaining the average capacity choices.

Finally, we extrapolate from our sample and use our estimates to predict the capacity choice in 1809 medium to small HSAs that are not currently included in our sample.[35] It turns out that our model did a good job in the out of sample prediction as well. Appendix Table A2 presents the comparison between predicted and the observed outcomes in the out of our sample markets.

6.4 Counterfactual

Our counterfactual experiment is motivated by the policy debate on the rapid growth of dialysis expenditure. As a result, Medicare started to implement a new bundled dialysis payment system in 2011. The new system incorporates the formerly separate billable items into a new bundled flat rate.[36] This effectively lowers the dialysis provider's per-treatment margin. To capture this policy effect, let λ be a factor of profit and Δ be a factor of cost, the counterfactual payoff under the alternative profit and cost factor is:

$$\Pi_i = \lambda K_i(x\beta_i + \sum\nolimits_j \gamma_{ij} K_j^* - \varepsilon_i) - \Delta(a_i K_i^2 + b_i K_i)$$

under the similar conditions derived in Section 3, this give rise to a counterfactual capacity choice of:

$$K_i^*(x, \varepsilon_i, \lambda, \Delta) = \max\left\{0, \frac{\lambda}{\Delta}\left(x\beta_i + \sum\nolimits_j \gamma_{ij} K_j^*(x, \varepsilon_j, \lambda, \Delta) - \varepsilon_i\right) - b_i\right\} \tag{9}$$

Let $\varphi_i = \mathbb{E}(K_i^* | x, \varepsilon_i, \lambda, \Delta)$ be the counterfactual capacity choice defined by the following equilibrium condition for each provider i:

$$\varphi_i = \Phi\left(\frac{x\beta_i + \gamma_i\varphi_{-i} - b_i\frac{\Delta}{\lambda}}{\sigma_i}\right) * \frac{\lambda}{\Delta} * \left\{x\beta_i + \gamma_i\varphi_{-i} - \frac{\Delta}{\lambda}b_i + \sigma_i \frac{\phi\left(\frac{x\beta_i + \gamma_i\varphi_{-i} - b_i\frac{\Delta}{\lambda}}{\sigma_i}\right)}{\Phi\left(\frac{x\beta_i + \gamma_i\varphi_{-i} - b_i\frac{\Delta}{\lambda}}{\sigma_i}\right)}\right\} \tag{10}$$

Under the status quo when $\lambda = 1, \Delta = 1$, (10) gives the same prediction as Table 6. If the per-capacity cost increases by 20% and per-capacity revenue decreases by 10%, the counterfactual capacity distribution can be obtained by adjusting $\frac{\lambda}{\Delta} = \frac{0.9}{1.2} = 0.75$ in (10). A similar counterfactual has been

[35] These markets are used to perform robustness analysis in Panel A of Appendix A1. There are 3129-1320=1809 markets where 1320 is the number of markets included in our final sample.

[36] The new system also incorporates some pay-for-performance incentives. Penalties will be imposed on providers whose dialysis quality measures (namely, patient's hemoglobin and urea levels) did not meet standards. The maximum payment reduction is up to 2 percent. We didn't explicitly investigate the pay-for-performance incentive in our analysis for two reasons. First, the incentive is relatively small. Second, several existing papers (Grieco and McDevitt 2012, Cutler, Dafny and Ody 2013) find that dialysis quality is not sensitive to competition.

applied by Schaumans and Verboven (2011) to investigate the market for healthcare professionals. According to MedPAC, the base composite rate under the old system is about \$142 per patient in 2012 (after excluding the \$20 drug add-on payment). Since the separate billable drugs account for approximately 40% of total Medicare payment for dialysis, this leads to an estimated average of \$257 per treatment payment.[37] Given the new composite payment base rate of \$235 in the same year, there is approximately a 5% reduction in the per-patient payment. Due to the lack of the detailed patient level payment and cost data, we cannot precisely measure the reduction on the per-patient margin as a result of reduction in the payment rate. Instead, we qualitatively investigate the new policy by lowering $\frac{\lambda}{\Delta}$ to different levels. In Table 6, we report the new equilibrium outcome when $\frac{\lambda}{\Delta}$ is reduced by 2%, 5% and 8%. Possibilities also remain that the margin goes up for facilities who don't rely too much on the separate billable drugs. We also simulate the outcome when $\frac{\lambda}{\Delta}$ increases by 2%. Finally, we simulate the heterogeneous response to the policy by adjusting $\frac{\lambda}{\Delta} = 0.98$ for DaVita while holding the profit to cost ratio constant for other providers. This is motivated by the well-known fact that DaVita relies heavily on drug revenue.

Table 6 presents the results for different counterfactual scenarios. The "capacity" column presents the mean capacity choice and aggregated expected capacity across all markets observed in data. The "entry" column presents the provider-specific mean entry probabilities and the number of markets where a provider would enter with probability greater than 50% in the counterfactual environment.

All providers reduces their capacity choice in response to the reduction in margin. However, the magnitudes of the responses are quite different across providers. The magnitude of capacity withdraw from DaVita is the smallest, followed by FMC and non-Chain. Thus the local independent providers were affected most by the reform assuming that it reduces the margin for all providers by the same proportion. On the other hand, one striking finding is that the capacity adjustment of DaVita is negatively correlated with that of FMC. In fact, DaVita increases its capacity stock in 108 markets when $\frac{\lambda}{\Delta} = 95\%$. The downward sloping reaction curve, arising from the strategic interactions in capacity choices, could be the key reason. When a rival reduces its capacity, the downward sloping reaction curve would incentivize a provider to increase its capacity. It also helps to preserve the margin for providers in the face of any negative profit shocks.

The entry probabilities respond in a similar manner when the margin for all providers decreases by the same proportion. Non-Chain providers respond most strongly followed by FMC and then DaVita. Though the average entry probability for DaVita decreases unambiguously with lower and lower profit to cost ratio, the number of market with high entry probability increases substantially for DaVita

[37] It is computed as 20+142/0.6.

when $\frac{\lambda}{\Delta} = 98\%$ and remain the same same as that under the status quo when $\frac{\lambda}{\Delta} = 95\%$.

Since the policy reform can imply a positive profit shock for providers who are less reliant on drug revenues, we also investigate the outcome when $\frac{\lambda}{\Delta} = 102\%$ for all providers. The results closely mirror those from a profit ratio reduction. FMC responds most aggressively by increasing its capacity stock. While non-Chain also increases the capacity, DaVita slightly cuts back its capacity stock due to the expansion of both rivals and the downward sloping reaction curve. Overall, when the providers face the same profit shock (that is, the same percentage change in margin), DaVita is better at absorbing the negative profit shock and is less responsive to the positive profit shock. This may arise from the asymmetries in providers' profit and cost structures as well as the difference in their strategic responses to their rival's capacity choice.

Finally, we simulate the heterogeneous effect of the policy reform by assuming that it implies a negative profit shock for DaVita while having no effect on FMC and non-Chain providers. We set $\frac{\lambda}{\Delta} = 0.95$ for DaVita and $\frac{\lambda}{\Delta} = 1$ for FMC an non-Chain.[38] Our model implies that DaVita reduces the capacity stock substantially while both rivals slightly increase capacity. The negative profit shock for DaVita is magnified by the downward sloping reaction curve. When the margin decreases, a direct response for DaVita's is to reduce capacity. This motivates both FMC and non-Chain to increase their capacity through the reaction curve, which in turn incentivizes DaVita to further reduce capacity. Overall, our model suggests that with the downward sloping reaction curve, even a seemingly small asymmetric profit shock could induce significant change in dialysis providers' capacity choice.

7 Conclusion

This article is motivated by two features of the U.S. dialysis market: Dialysis providers choose capacities and rarely change them after initial entry, and capacity choices vary substantially over a wide support. To capture these empirical regularities, we propose a new structural model of Bayesian games with continuous actions and use it to examine strategic capacity choices of U.S. dialysis providers. We estimate our model using a couple of estimators, one of which is a GMM estimator fully exploiting the structural relationship without assuming arbitrary equilibrium selection. Then we use the model to investigate counterfactual policy interventions that either increase or reduce per-capacity margin.

Our estimates suggest Cournot-like competition between providers with downward sloping reaction curves. Dialysis providers' capacities decrease with that of competitors". A unit increase in a

[38] As we explained previously, this is motivated by the observation that DaVita relies heavily on drug revenue. In 2007, New York Times reported that 40% of DaVita's revenue comes from dialysis related drugs. By constrast, 25% of FMC's revenue depends on drug.

competitor's capacity can reduce a provider's entry probability by 3-13%. This suggests that conventional models of the binary entry decisions would overlook the heterogeneity of strategic effects. Our econometric method is of interests in its own right and can be applied to a wider class of Bayesian games with continuous choices.

Table 1. Summary Statistics of Capacity Distribution

Variable	Definition	Mean	Std.	min	max
K_{fmc}	FMC's Capacity Decision	6.40	11.4	0	60
K_{dav}	DaVita's Capacity Decision	4.40	9.22	0	54
$K_{nonchain}$	Non-Chain's Capacity Decision	5.32	10.0	0	58
I_{fmc}	$I(K_{fmc} > 0)$	0.31	0.46	0	1
I_{dav}	$I(K_{dav} > 0)$	0.23	0.42	0	1
$I_{nonchain}$	$I(K_{nonchain} > 0)$	0.28	0.45	0	1
$K_{fmc} \vert I_{fmc} = 1$	FMC's Capacity Decision Conditional on Entry	20.4	11.2	2	60
$K_{dav} \vert I_{dav} = 1$	DaVita's Capacity Decision Conditional on Entry	19.3	9.24	7	54
$K_{nonchain} \vert I_{nonchain} = 1$	Non-Chain's Capacity Decision Conditional on Entry	18.9	10.0	1	58

Note: obs=1320

Table 2. Capacity Correlations

	Correlation					
	I_{fmc}	I_{dav}	I_{other}	K_{fmc}	K_{dav}	$K_{nonchain}$
I_{fmc}	1					
I_{dav}	-0.12	1				
I_{other}	-0.13	-0.07	1			
K_{fmc}	0.83	-0.12	-0.12	1		
K_{dav}	-0.14	0.88	-0.04	-0.13	1	
$K_{nonchain}$	-0.13	-0.09	0.85	-0.13	-0.06	1

Note: obs=1320

Table 3. Summary Statistics of Market Variables

Variable	Definition	Mean	Std Dev
pop*	Total HSA pop from DA	75,531.17	65,734.43
black	Percent black from Census	0.07	0.09
white	Percent white from Census	0.88	0.11
latino	Percent latino from Census	0.09	0.11
asian	Percent asia from Census	0.02	0.03
age1	Percent pop age between 22 and 44 from Census	0.33	0.04
age2	Percent pop age between 44 and 65 from Census	0.34	0.04
age3	Percent pop age 65+ from Census	0.14	0.03
neph	Number of nephrologist per 1000 pop from DA	1.51	1.00
bed*	Number of hospital bed per 1000 pop from DA	2.59	0.89
rn	Number of registered nurse per 1000 pop from DA	1.32	0.27
dbrate	Prevelence rate of diabetes	8.42	1.63
conreg	CON regulation indicator	0.21	0.41
NE	Northeast region indicator	0.23	0.42
MW	Midwest region indicator	0.24	0.42
West	West region indicator	0.17	0.37
dfmc	Distance to FMC's headquarter in 1000 miles	1.03	0.73
dfmc2	dfmc sqaured	1.60	2.04
ddav	Distance to DaVita's headquarter in 1000 miles	1.10	0.38
ddav2	ddave sqaured	1.34	0.83

Note: obs=1320. Variables labeled by * enter the estimation in logs and are reported without logging in this table. Explanatory variables come from the DA (Dartmouth Atlas) or Census when indicated and otherwise are constructed by the authors.

Table 4. Strategic Capacity Model

	Panel A: Two-Stage Estimation ($\hat{\theta}^{TS}$)						Panel B: Fixed Point ML ($\hat{\theta}^{FL}$)						Panel C: GMM ($\hat{\theta}^{GM}$)					
	FMC		Dav		Non-Chain		FMC		Dav		Non-Chain		FMC		Dav		Non-Chain	
	EST	SE	EST	SE	EST	SE	EST	SE	EST	SE	EST	SE	EST	SE	EST	SE	EST	SE
K_{fmc}			-1.06	0.24***	-0.73	0.24***			-1.88	0.19***					-2.17	0.09***	-0.82	0.10***
K_{dav}	-1.07	0.26***			-0.68	0.26***	-2.32	0.25***			-1.86	0.22***	-0.76	0.10***			-0.61	0.08***
$K_{non-chain}$	-1.43	0.26***	-1.04	0.30***			-1.44	0.20***	-1.73	0.26***	-1.58	0.23***	-0.86	0.17***	-1.45	0.16***		
lpop	25.0	2.20***	24.9	2.43***	21.7	2.10***	25.0	1.59***	28.3	2.23***	26.9	2.11***	20.1	1.22***	35.1	1.19***	20.9	0.78***
neph	2.46	0.76***	3.38	0.99***	1.23	0.87	2.11	0.76***	2.34	0.89***	2.11	0.93**	2.00	0.11***	4.36	0.11***	2.00	0.05***
lbed	1.39	4.19	-5.34	4.51	4.04	4.65	-0.37	4.00	-0.55	4.80	-0.04	4.63	3.86	0.97***	4.74	0.80***	1.06	0.89
rn	9.87	4.45**	11.3	4.84**	3.44	5.04	10.8	4.33***	12.1	5.12***	10.7	4.95**	6.39	2.08***	9.59	1.80***	5.35	0.78***
dbrate	2.21	0.76***	1.35	0.93	0.44	0.84	1.47	0.76**	1.47	0.92*	0.91	0.91	2.03	0.68***	2.43	0.55***	0.51	0.67
conreg	-4.84	1.92**	-2.66	1.99	-3.54	2.11*	-3.93	1.90**	-3.91	2.25**	-4.08	2.16**	-4.08	0.42***	-5.10	0.38***	-2.60	0.15***
NE	-8.16	4.04**	-10.0	5.07**	7.68	4.99	-6.17	4.03*	-5.38	4.92	-0.66	4.70	-9.45	3.13***	-16.5	2.88***	5.70	0.98***
MW	-8.80	2.54***	-4.71	3.26	-5.35	2.96*	-7.35	2.61***	-8.09	3.18***	-7.65	3.09***	-6.74	0.98***	-8.54	0.87***	-6.56	0.36***
West	-8.49	4.00**	-10.8	5.18**	10.5	4.10**	4.78	4.35	8.54	5.55*	12.7	4.64***	-7.24	0.88***	-13.4	0.78***	9.10	1.23***
dfm	10.9	9.18	30.7	10.7***	33.8	11.7***	22.7	10.1**	30.4	12.1***	37.8	11.4***	-1.41	6.10	7.91	5.66*	28.3	2.93***
dfmc2	-3.96	3.03	-8.22	3.69**	-12.3	3.65***	-8.91	3.30***	-11.6	3.98***	-14.4	3.60***	-0.02	1.75	-1.17	1.60	-10.7	0.89***
ddav	12.6	14.2	7.70	16.9	72.9	13.6***	2.51	13.3	10.7	16.0	32.6	15.8**	-4.25	8.39	5.98	6.73	63.2	10.6***
ddav2	-6.27	7.12	2.99	8.38	-33.3	6.85***	-0.73	6.37	-3.74	7.69	-13.6	7.47**	0.00	4.69	-0.62	3.71	-30.3	3.89***
black	60.2	22.6***	22.5	22.5	81.0	44.2*	41.8	19.4**	43.2	20.4**	65.9	32.2**	70.7	14.9***	75.6	9.88***	82.9	88.2
white	22.8	20.7	-14.1	20.3	69.2	43.3	15.6	18.3	15.8	18.4	42.4	30.9*	28.8	18.8*	12.7	12.6	69.1	88.9
latino	38.1	8.56***	7.69	10.3	16.1	10.2	14.7	8.65**	13.5	11.2	12.6	10.1	27.6	16.8*	22.7	13.1**	16.1	10.3*
asian	-86.6	50.1*	21.5	47.5	87.9	56.4	25.3	42.3	40.7	43.8	64.8	49.2*	-30.8	72.1	-11.2	48.9	65.2	110
age1	29.0	45.9	-98.2	52.7*	-12.0	47.8	-7.01	46.3	-18.4	53.3	-9.17	50.4	12.9	175	-78.8	138	20.0	112
age2	81.9	39.7**	-40.7	46.8	7.13	42.9	-3.77	40.3	-16.7	45.7	-18.8	44.0	41.0	131	-29.1	103	35.0	86.6
age3	53.4	37.0	-45.7	43.9	28.2	41.3	90.4	39.7**	90.7	47.9**	86.2	44.9**	44.5	112	10.9	88.6	39.4	74.9
cons	-384	45.6***	-270	55.0***	-382	61.0***	-330	39.6***	-375	51.4***	-395	51.9***	-298	129**	-416	106***	-384	146***
sigma	23.0	0.71***	24.1	0.86***	22.9	0.74***	20.4	0.80***	22.3	1.15***	22.5	1.03***	21.3	0.19***	25.1	0.04***	22.7	0.07***

Note:***p<1%,**p<5%,*p<10%. obs=1320. In panel A, the estimates are obtained using Two-Stage estimator. The expected capacity choice of each provider is estimated in the first step using a Poisson regression with regressors being the polynomials of market characteristics, and the standard errors are calculated using a bootstrap procedure. In panel B, the nested fixed point problem is recast into MPEC and estimated using Knitro. In Panel C, the estimates are obtained through GMM. K_{fmc},K_{dav},$K_{non-chain}$ are the expected capacity choice of FMC, DaVita and Non-Chain respectively. lpop and lbed are logged population and logged number of hospital bed. The definition of other variables are presented in Table 3.

Table 5. Equilibrium Response to Market Variables

	Capacity			No. of HEP markets		
	FMC	Dav	Non-Chain	FMC	Dav	Non-Chain
base case	8003	5586	6700	330	115	214
popuplation increases 10%	8560	6097	7238	368	137	246
nephropoligst increases 10%	8066	5784	6766	337	126	216
prevalence rate of diabetes increases 10%	8996	5593	6565	395	107	202
CON regulation =0 for all markets	8336	5593	6795	354	114	216

Note: obs=1320. Capacity is derived from equation (6) after adjusting market variables. The capacity column reports the sum of expected capacity choice for each provider across markets. The number of HEP (i.e. "high entry probability") markets for a provider is defined as the number markets in which that provider's entry probability is estimated to be greater than 0.5.

Table 6. Model Fit

	Observed			Predicted		
	FMC	Dav	Non-Chain	FMC	Dav	Non-Chain
Panel A: Entry Pr(K>0)						
Mean	0.31	0.23	0.28	0.35	0.24	0.30
Std dev	0.46	0.42	0.45	0.21	0.18	0.18
Panel B: Capacity E(K|x)						
Mean	6.40	4.40	5.32	6.06	4.23	5.08
Std dev	11.39	9.23	10.03	5.33	4.46	4.16

Note: obs=1320. "Mean" reports the sample average. "std dev" reports the standard deviation.

Table 7. Counterfactual

		Capacity			Entry		
		FMC	Dav	Non-Chain	FMC	Dav	Non-Chain
status quo $\frac{\lambda}{\Delta}=1$	mean	6.1	4.2	5.1	0.35	0.24	0.3
	total	8003	5886	6700	330	115	214
$\frac{\lambda}{\Delta}$=98%	mean	4.6	4.0	3.3	0.29	0.23	0.23
	total	6069	5280	4408	218	141	72
$\frac{\lambda}{\Delta}$=95%	mean	2.8	3.1	1.6	0.21	0.18	0.13
	total	3727	4135	2116	90	115	8
$\frac{\lambda}{\Delta}$=92%	mean	1.6	2	0.7	0.13	0.13	0.06
	total	2060	2650	859	10	61	0
$\frac{\lambda}{\Delta}$=102%	mean	7.7	4.1	7.4	0.41	0.24	0.39
	total	10148	5470	9722	446	86	406
$\frac{\lambda}{\Delta}$=95% for Dav	mean	6.9	1.8	5.4	0.38	0.13	0.32
status quo for others	total	9095	2427	7139	403	6	251

Note: obs=1320. The Mean capacity row reports the average expected capacity across markets. The total capacty row reports the sum of expected capacity across markets. The mean entry row reports the average entry probability across markets. The total entry probability row reports the total number of market with high entry probability (i.e entry probability greater than 0.5).

A Supplemental Tables

Table A1. Additional Robustness

	Panel A						Panel B						Panel C					
	FMC		Dav		Non-Chain		FMC		Dav		Non-Chain		FMC		Dav		Non-Chain	
	EST	SE	EST	SE	EST	SE	EST	SE	EST	SE	EST	SE	EST	SE	EST	SE	EST	SE
FMC			-0.82	0.17***					-0.69	0.17***					-0.75	0.23***		
Dav	-1.14	0.20***			-0.88	0.17***	-0.84	0.16***			-0.64	0.16***	-1.46	0.28***			-1.05	0.23***
Other	-1.18	0.29***	-0.60	0.28**	-0.55	0.18***	-0.92	0.25***	-0.22	0.25	-0.35	0.15**	-1.45	0.37***	-0.93	0.37**	-0.93	0.28***
lpop	25.1	1.64***	19.7	1.77***	20.7	1.48***	25.1	1.58***	19.6	1.83***	20.5	1.48***	21.6	1.61***	17.0	1.79***	18.6	1.45***
black	51.5	15.1***	9.54	12.4	0.76	11.7	58.7	16.2***	10.0	13.0	-1.53	12.1	42.4	13.0***	8.45	11.9	4.78	10.9
white	10.1	15.2	-17.6	11.8	-23.3	11.4**	15.0	16.2	-20.3	12.6	-25.6	12.0**	4.95	13.1	-13.4	11.2	-20.7	10.6*
latino	26.5	6.98***	3.23	7.85	14.8	6.92**	26.3	7.02***	-1.42	8.20	14.6	7.08**	19.4	6.59***	-1.22	8.04	5.67	7.13
asian	-85.2	33.1***	7.93	29.2	-0.75	27.0	-72.2	32.3**	16.6	29.7	-2.27	27.6	-77.3	31.6**	32.4	29.2	-1.86	26.8
age1	28.1	36.8	-60.9	38.2	-39.5	36.1	33.0	37.5	-65.2	40.1	-39.3	37.2	22.4	34.3	-78.1	38.1**	-47.4	35.9
age2	38.1	33.3	-54.1	34.2	-20.4	31.9	36.1	34.2	-65.5	36.0*	-24.6	33.0	21.6	30.7	-69.7	33.1**	-26.1	30.7
age3	36.0	32.8	-20.4	34.7	24.7	32.1	46.9	33.4	-20.4	36.3	27.7	33.0	32.1	30.6	-26.4	34.4	12.9	31.6
neph	2.23	0.55***	2.54	0.59***	1.32	0.59**	1.98	0.57***	2.50	0.63***	1.14	0.61*	2.26	0.50***	2.16	0.59***	1.41	0.57**
lbed	5.53	3.08*	-2.14	3.32	6.12	3.11**	6.86	3.16**	-2.76	3.50	7.04	3.20**	4.17	2.86	-4.25	3.30	4.41	2.99
rn	10.4	3.47***	6.35	3.69*	3.59	3.47	10.3	3.54***	6.62	3.87*	3.02	3.58	8.68	3.23***	6.70	3.61*	3.26	3.35
dbrate	1.67	0.64***	1.13	0.69	0.59	0.66	1.47	0.65**	0.89	0.72	0.27	0.68	1.52	0.59***	1.15	0.69*	0.75	0.65
conreg	-3.42	1.56**	-2.37	1.70	-3.68	1.59**	-3.76	1.60**	-2.59	1.78	-3.88	1.63**	-2.79	1.46*	-1.60	1.69	-3.70	1.54**
NE	-7.63	3.87**	-12.4	4.07***	7.76	3.64**	-10.2	3.95***	-17.3	4.22***	9.05	3.72**	-9.05	3.67**	-11.3	4.15***	3.77	3.73
MW	-6.10	2.18***	-3.09	2.36	-1.94	2.26	-6.23	2.19***	-3.46	2.43	-1.18	2.29	-6.49	2.03***	-4.47	2.34*	-1.43	2.23
West	-10.0	3.64***	-15.0	4.10***	11.8	3.11***	-10.8	3.65***	-16.6	4.29***	13.2	3.20***	-6.85	3.57*	-11.1	4.14***	11.5	3.03***
fmcdist	13.1	8.44	14.9	8.79*	33.1	8.17***	7.29	8.36	10.3	9.02	34.2	8.18***	8.99	7.98	16.8	8.85*	24.6	8.40***
fmcdist2	-3.53	2.82	-2.16	2.94	-12.5	2.57***	-1.83	2.80	-0.64	3.04	-13.1	2.59***	-2.80	2.64	-3.83	2.94	-9.58	2.60***
davdist	2.38	11.5	-21.5	11.6*	58.7	11.0***	-1.29	11.4	-29.0	12.0**	61.9	11.2***	-3.53	10.4	-19.5	11.2*	43.8	10.5***
davdist2	0.35	5.69	15.4	5.72***	-26.0	5.45***	1.22	5.64	19.4	5.87***	-27.3	5.52***	2.82	5.28	13.5	5.55**	-19.1	5.29***
cons	-358	35.8***	-192	35.9***	-260	31.0***	-360	36.3***	-177	37.1***	-257	31.4***	-293	32.9***	-151	35.0***	-218	29.3***
sigma	23.0	0.71***	23.5	0.83***	23.0	0.75***	23.9	0.71***	25.0	0.83***	24.1	0.76***	20.2	0.70***	21.9	0.87***	21.1	0.77***
obs	3129						3204						2930					

Note:***p<1%,**p<5%,*p<10%. Results are obtained using a two-stage estimator and standard errors are computed based on 300 bootstraps. Panel A is based a sample of 3129 markets where only large markets with capacity greater than 60 are dropped. Panel B and Panel C are based on alternative cutoffs when markets with capactiy greater than 80 and greater than 40 are dropped. The market capacity of 80 corresponds to approximately the 95th percentile in capacity distribution and 40 corresponds to the 85th.

Table A2. Model Fit: Out of Sample Prediction

	Observed			Predicted		
	FMC	Dav	Other	FMC	Dav	Other
Panel A: Capacity $E(K\|x)$						
Mean	2.84	2.53	2.55	3.43	2.16	2.36
Std dev	7.57	7.06	7.02	4.96	4.12	3.49
Panel B: Entry $Pr(K > 0)$						
Mean	0.16	0.15	0.15	0.21	0.12	0.15
Std dev	0.36	0.35	0.36	0.21	0.17	0.17
No markets:1809						

Note: The out of sample prediction is made based on the GMM estimates reported in Table 4 and the large sample in Panel A of Table A1. We report the results after excluding 1320 markets that are currently included in the main specifications. This left 3129-1320=1809 markets.

B Proof of Proposition 1

Proof of Proposition 1.. Suppose the solution of the maximization problem is in the interior (i.e. $K_i^*(X, \varepsilon_i) > 0$). Then the first-order condition of (1) implies that in equilibrium:

$$K_i^*(X, \varepsilon_i) = \tfrac{1}{2a_i} \left\{ X\beta_i + \textstyle\sum_{j \neq i} \gamma_{i,j} \mathbb{E}_{\varepsilon_j} \left[K_j^*(X, \varepsilon_j) | x, \varepsilon_i \right] - b_i - \varepsilon_i \right\}, \tag{11}$$

where the right-hand side of (11) is necessarily strictly positive. Since $a_i > 0$, the second-order condition for such an interior solution would be satisfied automatically. Next, suppose the solution is on the corner, i.e. $K_i^*(X, \varepsilon_i) = 0$. Because $a_i > 0$ and $\Pi_i(0, K_{-i}, x, \varepsilon_i) = 0$ regardless of K_{-i}, the first-order condition of (1) evaluated at $K_i^*(X, \varepsilon_i) = 0$ and $K_{-i}^*(.)$ is necessarily negative. (Otherwise the solution would be interior.) Therefore, with a corner solution in equilibrium, we have:

$$K_i^*(X, \varepsilon_i) = 0 \text{ and } X\beta_i + \textstyle\sum_{j \neq i} \gamma_{i,j} \mathbb{E}_{\varepsilon_j} \left[K_j^*(X, \varepsilon_j) | x, \varepsilon_i \right] - b_i - \varepsilon_i < 0. \tag{12}$$

Combining (11) and (12), we conclude that (2) holds in any PSBNE. *Q.E.D.* ■

References

[1] Athey, S. "Single Crossing Properties and the Existence of Pure Strategy Equlibria in Games of Incomplete Information." Econometrica, Vol. 69 (2001), pp. 861-890.

[2] Bajari, P., Hong H., Krainer, J. and Nekipelov,D. "Estimating Static Models of Strategic Interactions." *Journal of Business & Economic Statistics* Vol 28 (2010), pp. 469-482.

[3] Buckley, J. and James, I. "Linear Regression with Censored Data." *Biometrika*, 66 (1979),429-436.

[4] Cameron, A.C. and P.K. Trivedi *Regression analysis of count data*, Cambridge University Press 1998.

[5] Christensen, R. *Log-linear models and logistic regression. Springer Texts in Statistics (Second ed.)*. New York: Springer-Verlag 1997.

[6] Cultler, D., Dafny, L. and Ody, C. "How Does Competition Impact the Quality of Health Care? A Case Study of the U.S. Dialysis Industry." (2012), unpublished manuscript.

[7] Davis, P. "Spatial Competition in Retail Markets: Movie Theaters." *The RAND Journal of Economics* Vol. 37 (2006), pp. 964-982.

[8] Dixit, A. "The Role of Investment in Entry-Deterrence." *The Economic Journal* Vol. 90 (1980), pp. 95-106.

[9] Gelman, J. and Salop, S. "Judo Economics: Capacity Limitation and Coupon Competition." *The Bell Journal of Economics* Vol. 14 (1983), pp. 315-325.

[10] Horowitz, J.L. "A Distribution-Free Least Squares Estimator for Censored Linear Regression Models." *Journal of Econometrics*, Vol. 32 (1986), pp. 59-84.

[11] Judd, K. and Su, C.L. "Constrained Optimization Approaches to Estimation of Structural Models." *Econometrica* Vol. 80 (2012), pp.2213-2230.

[12] Kreps, D. and Scheinkman, J. "Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes." *The Bell Journal of Economics* Vol. 14 (1983), pp. 326-337.

[13] Ford, J. and Kaserman, D. "Certificate-of-Need Regulation and Entry: Evidence from the Dialysis Industry." *Southern Economic Journal*, Vol. 59 (1993), pp.783-791.

[14] Grieco, P. and McDevitt, R."Productivity and Quality in Health Care: Evidence from the Dialysis Industry." (2013) unpublished manuscript.

[15] Johansen, K. "Choice of Dialysis Modality in the United States." *Archives of Internal Medicine*, Vol. 171 (2011), pp.107-109.

[16] Mazzeo, M., "Product Choice and Oligopoly Market Structure." *The Rand Journal of Economics*, Vol. 33 (2002), pp.221-242.

[17] Newey, W. and McFadden, D. "Large Sample Estimation and Hypothesis Testing." Chapter 36, *Handbook of Econometrics*, Volume IV (1994), Edited by R.F. Engle and D.L. McFadden. Elsevier

[18] Powell, J.L. "Least Absolute Deviations Estimation for the Censored Regression Model." *Journal of Econometrics*, Vol. 25 (1984), pp. 303-325.

[19] Powell, J.L. “Symmetrically Trimmed Least Squares Estimation of Tobit Models.” *Econometrica*, Vol. 54 (1986), pp.1435-1460.

[20] Ramanarayanan, S. and Snyder, J. "Reputations and Firm Performance: Evidence From the Dialysis Industry." (2011) unpublished manuscript.

[21] Rust, J. "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher." *Econometrica* Vol. 55 (1987), pp.999-1033.

[22] Seim, K. "An Empirical Model of Firm Entry with Endogenous Product-Type Choices." *The RAND Journal of Economics*, Vol. 37 (2006), pp. 619-640.

[23] Sweeting, A. "The Strategic Timing Incentives of Commercial Radio Stations: An Empirical Analysis Using Multiple Equilibria." *The RAND Journal of Economics* , Vol. 40 (2009), pp. 710-742.

[24] Schaumans, C. and Verboven, F. "Entry and Regulation: Evidence from Health Care Professions." *The RAND Journal of Economics*, Vol. 39 (2008), pp.949–972.

[25] Thamer, M., Zhang, Y., Kaufman, J., Cotter, D., Dong, F., and Hernán, M. A. "Dialysis Facility Ownership and Epoetin Dosing in Patients Receiving Hemodialysis." *Journal of American Medical Association*, Vol. 297 (2007), pp.1667-1674.

[26] U.S. Renal Data System, USRDS 2009 Annual Data Report: Atlas of Chronic Kidney Disease and End-Stage Renal Disease in the United States, National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD, 2009.