Productivity and Quality in Health Care: Evidence from the Dialysis Industry*

Paul L.E. Grieco† Ryan C. McDevitt‡

March 2012

Abstract

Controlling the growth of health care costs is a long-term policy goal in many developed countries. We investigate whether policies that incentive cost reduction are likely to result in lower quality care. To do so, estimate a health provision production function in the US Dialysis industry where firms' quality of care is costly. We find a significant quality-quantity tradeoff for dialysis treatments. Controlling for unobserved productivity, we show that a dialysis facility can treat 1.7 percent more patients, holding inputs fixed, if it allows its rate of hospitalizations for septic infection to increase by one percentage point. Our approach also delivers productivity estimates which control for the quality of care provided, whereas traditional methods of productivity estimation would mis-attribute lower-quality output as high-productivity. We find a high degree of productivity dispersion in the dialysis industry even after controlling for quality provision. We also investigate the incentives to provide quality, we find that not-for-profit firms tend to offer higher quality of care, while there is little difference between the level of quality offered by firms in competitive versus monopoly markets.

JEL: D24
Keywords: productivity; quality variation

*We thank Mark Roberts for helpful comments.
†The Pennsylvania State University, Department of Economics, paul.grieco@psu.edu
‡University of Rochester, Simon Graduate School of Business, ryan.mcdevitt@simon.rochester.edu

1 Introduction

By treating more patients for a given level of inputs, a healthcare provider may reduce per-patient expenditures. If the provider reduces treatment quality to increase output, however, the overall gains from increasing "measured" productivity (i.e., capital to output or labor-to-output ratios) are ambiguous. With this as motivation, we extend existing methods for estimating firm production functions to measure the tradeoff between increasing output and reducing treatment quality for U.S. dialysis centers. In doing so, we estimate a production function for dialysis treatments while controlling for firms' endogenous input and quality choices. These estimates then allow us to assess the extent to which dialysis centers reduce costs by sacrificing the quality of care they provide.

While the role of the tradeoff between quality-of-care and cost-of-care is important to the entire health care industry, we focus on dialysis treatments — a process that cleans the blood of patients with end-stage renal disease (ESRD), or kidney failure — as our empirical setting for two reasons. First, payments to dialysis facilities comprise a substantial portion of Medicare's expenditures each year, making it an important area for policy analysis in its own right. Second, several features of the dialysis industry make it well suited for a study of quality provision in health care. For instance, dialysis is predominately conducted in standalone facilities, which creates a well-defined market for our analysis and reduces concerns that dialysis provision is cross-subsidizing other procedures, a common issue in healthcare settings. In addition, payments for treatment are largely uniform due to Medicare's prospective payment system and do not depend on quality. This allows us to isolate the effects of quality provision from price discrimination. Moreover, dialysis treatments follow a straightforward process related to stations and staff, which allows us to closely approximate a facility's production function. Facilities also choose input levels (i.e., staffing) and have observable differences in production (i.e., patient loads), which allow us to identify the relationship between inputs and outputs. Finally, facilities have observable quality differences (e.g., infection and death rates), which allow us to connect a firm's productivity to its output quality, the primary aim of our research.

To uncover the cost of providing higher quality care, we build on the structural methods first proposed by Olley & Pakes (1996), and later extended by Levinsohn & Petrin (2003) and Ackerberg et al. (2006). Conceptually, we adapt the insights from Ackerberg et al. (2006) to incorporate a "quality-choice" stage. After having acquired capital and having trained their labor force of nurses, managers observe an expected level of productivity and "choose" their quality of care. For example, managers may provide guidelines for the length of treatment, standards of cleanliness for equipment, and the degree of management oversight. They may also provide explicit or implicit incentives for either output or quality through compensation or promotion systems. While we do not directly observe managers quality choice, we do observe patient outcomes. If high quality care is likely to lead to better outcomes, than we can use outcomes as noisy proxies for quality. Since we have multiple measures of outcome (in our case, the septic infection rate and mortality rate) we can use an instrumental variable approach to recover the impact of quality-of-care on output.

We use our results to investigate why different dialysis centers choose different levels of quality. As part of Medicare's prospective payment system, dialysis centers have an efficiency incentive to minimize the costs of treating patients, which may include providing low-quality care. Counteracting this incentive, however, are plausible motivations for providing higher-quality care: centers must report quality statistics to Medicare which are then made public. Centers also face intermittent inspections by state regulators. In addition, patients can choose their dialysis provider, potentially leading centers to compete for patients by providing higher-quality care. Finally, some centers are not-for-profit and may have non-profit motives for providing high quality. Based on these conflicting forces, we investigate whether centers respond to incentives to provide high-quality care.

We find evidence of a substantial quality-quantity tradeoff for dialysis treatments: a center that reduces its treatment quality to allow a one percent higher expected septic infection rate increases its patient load by 1.7 percent, holding input levels and productivity constant. In addition, our approach allows us to recover estimates of total factor productivity for each firm that properly account for endogenous quality choices. We find substantial productivity dispersion in the industry that is not explained by differences in treatment quality. Furthermore, while we find large average productivity gains in the industry (between 5 and 8 percent per year), they have come through higher output, not lower infection rates. Finally, we investigate the determinants of quality in the industry. We find that market concentration does not significantly affect treatment quality, so it does not appear that competitive forces provide much incentive for quality. However non-profit dialysis centers provide a significantly higher quality of care: for-profit dialysis centers have a 1.3 percent higher infection rate than their non-profit counterparts. Overall, our results provide significant evidence that while profit-based incentives to reduce costs in dialysis treatments may lead to lower-quality care, competition has a limited effect on maintaining quality.

In addition to providing relevant policy analysis, this paper also contributes to the growing literature in empirical industrial organization related to the structural estimation of production functions. The estimation of production functions has a long history in economics, with many well-known econometric issues related to selection and simultaneity bias receiving considerable attention.[1] In light of this, recent work has developed structural techniques that use firms' observed input decisions to control for unobserved productivity shocks and overcome endogeneity problems. To adapt these methods to the dialysis industry, we explicitly incorporate the use of observable measures of output quality into the center's production function.

Firms' quality choices warrant special attention health care because failure to control for differences in product quality may bias estimates of productivity. As an example, consider the econometrician who observes factory A producing twice as many widgets as factory B using exactly the same measurable inputs, such as labor hours. If, in fact, widgets from factory A fail at a higher rate because its employees devote less attention to quality control, then the true difference in productivity is less than the factor of two suggested by a simple quantity measure. The productivity literature commonly assumes away this type of measurement error by arguing that higher-quality products (i.e., those with fewer defects) command higher prices, which will then be captured by revenue-denominated output measures. However, the mapping of price to quality has many shortcomings, however, as prices reflect more than just quality (Griliches & Mairesse 1995, De Loecker 2012).[2] Output quality is of first-order importance in many applications and estimation techniques that do not properly incorporate quality variation across firms result in misleading inference. Moreover, standard techniques to control for quality — such as price variation — may not adequately resolve the issue, particularly in healthcare settings such as dialysis where prices are set by Medicare rather than individual facilities and are completely unrelated to the quality of care provided (in fact, they are set before the quality of care is even chosen). Our approach illustrates how to directly incorporate measures of treatment quality into estimates of firm-level production functions. We believe this approach may be useful in a variety of healthcare settings beyond the dialysis industry.

The following section describes our institutional setting and data. Section 3 outlines our

[1]See Syverson (2011) for a recent review.

[2]A separate line of research, such as Fox & Smeets (2011), uses the quality of inputs (e.g., the education levels of employees) to explain productivity dispersion, which we will also consider in our application.

methods for estimating a production function in the presence of an endogenous quality choice. Section 4 presents our estimation results. Finally, Section 5 concludes with a discussion of our findings' implications.

2 Empirical Setting and Data Description

The demand for dialysis treatments comes from patients afflicted by end-stage renal disease (ESRD), a chronic condition characterized by functional kidney failure that results in death if not treated properly. Treatment options for ESRD include a kidney transplant or dialysis, a process which cleans the blood of waste and excess fluids. Patients can receive different dialysis modalities, the most popular being hemodialysis that circulates a patient's blood through a filtering device before returning it to the body.

The majority of patients undergoing dialysis in the United States receive care at free-standing dialysis facilities with coverage through Medicare. Medicare instituted its ESRD program in 1973 and now covers approximately 400,000 individuals; notably, all patients with ESRD become eligible for Medicare coverage, regardless of age. Today, Medicare spends more than $20 billion a year on dialysis care — about $77,000 per patient — which constitutes more than 6 percent of all Medicare spending despite affecting fewer than 1 percent of Medicare patients (ProPublica 2011). Beginning in 1983, Medicare has paid dialysis providers a fixed, prospective payment — the "composite rate" — for each outpatient treatment delivered, up to a maximum of three sessions per week per patient. Initially, the payment rate did not adjust for quality, length of treatment, dialysis dose, or patient characteristics. In 2005, Medicare began to adjust payments based on patient characteristics, but payments are still independent of patient outcomes.

Dialysis treatments are labor intensive by nature. Patients must remain connected to a dialysis machine for approximately 2-5 hours to filter impurities and remove excess fluid from the blood. Throughout this process, little scope for substituting labor exists. Prior to treatment, staff connect the machine to a patient by inserting two lines into a vascular access. During treatment, staff must continually monitor patients to evaluate their condition (e.g., blood pressure) and treat symptoms that arise (e.g., hypotension). Following treatment, staff disconnect a patient from the dialysis machine and assess his condition a final time before discharge. Staff must then sterilize equipment and prepare the station for the next patient. As a result of this hands-on care, the cost per patient treated necessarily increases with the average duration of treatment. Labor costs, which consist largely of nurses and technicians' wages, reflect this, accounting for approximately 70-75 percent of a facility's total variable costs (Ford & Kaserman 2000).

Facilities employ different types of labor, with registered nurses (RNs) constituting the majority of staff. Technicians, who have less-extensive training than RNs, also treat patients but can do so with only a high-school diploma and in-house training (though they must eventually pass a state or national certification test). Facilities also must have board-certified physicians as medical directors, but often have no physician on site. Medicare does not mandate a specific staffing ratio for dialysis centers, although some states do.

In our data, we observe the number of nurses working in a center and their rank, as well as the number of dialysis station available in the center annually. We believe that centers' capital and labor choices difficult to change in the short run. Our measure of the capital stock is the number of dialysis stations in the center. Investment in stations is extremely lumpy. The average number of dialysis stations run by a center is 18, so the purchase of a new machine is a significant investment. In our data, there is zero net investment over 90 percent of year-to-year transitions. The labor choice is less lumpy, but we still observe zero net hiring in roughly 18 percent of year-to-year transitions. Because of training and certification requirements, centers cannot quickly react to changes in productivity by hiring more workers. For this reason, we will use net hiring between periods t and $t+1$ to proxy for productivity in period t. In contrast to labor and capital choices, changes in "quality-targets" such as how much time and effort should be devoted to preventing infections in the center can be quickly adjusted by management. Therefore, it is important that our estimation method control for productivity differences between firms, since these differences may play a role in determining firms' quality of care.

The goal of our study is to assess the cost of providing high-quality care. Unfortunately, we do not observe direct measures of center policies as regards to the quality of care. Instead, we use proxies based on patient outcomes. Specifically, we use data on the rate of septic infection within the center and the ratio of the death rate to the expected death rate computed by Medicare based on individual patient characteristics. Patients undergoing dialysis are at a high risk of septic infection due to the exposure of their blood during treatment, the risk is strongly related to the cleanliness of the dialysis environment. The rate of hospitalization due to septic infection in our data was 12 percent per year, much higher than the rate in the general population.

Because a facility's payments per treatment do not vary by the duration of treatment under

Medicare's prospective payment system, a facility's profit per treatment decreases as treatment times — and, hence, labor costs — increase. At the same time, the effectiveness of dialysis increases with its duration; for instance, longer treatment cycles have been linked to lower mortality rates (ProPublica 2011). Centers thus face a straightforward tradeoff between increasing treatment quality and decreasing costs.[3] And though the costs of providing high-quality care are relatively clear, the benefits for dialysis centers are less apparent. First, demand-side incentives appear weak because dialysis provides life-sustaining functions for patients, making their demand for treatments inelastic. Second, patients typically have few dialysis centers to choose from in any given market — the mean market share across the United States is 0.457 — and, since ESRD immobilizes those affected by it, travel costs limit market choice. Finally, as discussed above, Medicare's payment system provides no direct financial incentive for providing high-quality care.

Firms may, however, still have several possible motives for delivering high-quality care. For instance, when a facility does face competition for patients, providing low-quality care may reduce demand if patients defect to other facilities that provide relatively better care and have excess capacity. Moreover, a facility that provides inadequate treatment may face increased regulatory scrutiny that further drives patients to competitors or results in decertification (the degree of regulatory scrutiny varies from state to state). Finally, some centers, particularly non-profit entities, may have motives to provide high-quality care unrelated to profitability.

Data Sources We use several sources of data for our analysis. Our primary dataset comes from the Centers for Medicare and Medicaid Services which contracts with the University of Michigan's Kidney Epidemiology and Cost Center to compile customized reports for each dialysis facility in the country. In December 2010, ProPublica, a non-profit organization dedicated to investigative journalism, obtained these reports under the Freedom of Information Act and posted them online. We systematically downloaded all 36,783 individual reports from 2002 — 2010 and constructed a usable dataset, which to our knowledge is the first time it has been used for research purposes. The data include detailed information on aggregated patient characteristics (e.g., age, gender, co-morbid conditions, body mass index, etc.) and facility characteristics (e.g., number of stations and nurses, years in operation, etc.).

[3] Critics allege that facilities may sacrifice quality of care in pursuit of efficiency, turning over three to four shifts of patients a day. And while policy makers contend that technicians should not monitor more than four patients at once, patient-to-staff ratios exceed this guideline in many facilities. At the extreme, inspection reports allege that

Table 1: Summary statistics, 2004-2009

Variable	**Mean**	**Std. Dev.**	**N**
Patient Years	50.002	32.356	19485
FTE Nurses & Techs	11.609	7.354	23934
Stations	18.331	7.978	23934
For-Profit	0.873	0.333	23934
Septic Infections	12.422	6.615	19459
Competing Facilities in HSA	8.097	13.436	23934
Market Share	0.457	0.388	23931

Table 1 presents selected summary statistics from the data, and several variables deserve note. First, Medicare analyzes individual patient records and calculates the number of patient-years each dialysis center serves (e.g., a patient treated at a center for six months is accounted for as one half of a patient-year). This figure provides an accurate record of dialysis provision that accounts for partial years of service due to death, transfers, transplants, newly diagnosed patients, and so forth. As discussed above, we use the number of full-time equivalent (a weighted mix of full-time and part-time) employees at each center and the number of dialysis stations available as our measure of labor and capital inputs.

Our primary proxy for quality is a center's hospitalization rate from septic (blood) infections as our primary measure of quality. Dialysis patients face a high risk of septic infections primarily because their blood may be exposed during treatment, with inattentive staff and unclean equipment substantially increasing the likelihood of infection. In addition to the septic infection rate, we use the ratio of deaths to expected deaths as an alternative measure of quality.[4] Importantly, we can also control for aggregate center-level patient characteristics which influence productivity and quality.

The competitive environment faced by dialysis centers is highly variable. Following the healthcare literature, we use hospital service areas (HSA) as our market boundaries for dialysis centers. While roughly 20 percent of dialysis centers are monopolies within their HSA, the average number of centers in an area is 8.1, but this figure is highly skewed. A center's average patient-weighted market share within an HSA is 0.45.

some clinics have allowed patients to soil themselves rather than interrupt dialysis (ProPublica 2011).

[4]The center-level expected death rate is calculated by Medicare using individual patient characteristics.

3 Measuring the Quality-Quantity Tradeoff in Dialysis

To measure the relationship between productivity and treatment quality, we propose and estimate a structural model of dialysis provision. In doing so, we account for both the standard endogeneity problems that arise when estimating production functions and the additional problem introduced by a firm's endogenous choice of output quality. The complication related to endogenous quality decisions stems from the unobserved (to the econometrician) choice made by firms that receive positive shocks to productivity: they may choose either to treat more patients, or to put more effort into treating current patients. If highly productive firms elect to provide higher-quality care for their patients, our estimates of the quality-quantity tradeoff will be biased downward, leading us to underestimate the true cost of quality.

To control for this potential source of bias, we extend the work of Olley & Pakes (1996) and Ackerberg et al. (2006) by incorporating endogenous quality targets. Since we only observe noisy measures of quality in our data our approach uses proxies for for patient quality—firm-level hospitalization rates for septic infection in our application. The use of a proxy may introduce attenuation bias due to measurement error in quality choices, which would lead us to underestimate the quality-quantity tradeoff. We correct for measurement error by using a second noisy measure of quality—the ratio of deaths to expected deaths—as an instrumental variable.

3.1 The Production Technology

We formalize the production function for dialysis treatments as follows. First, each period a center considers its current productivity level and mix of inputs before jointly determining (i) how many patients to treat and (ii) what quality of care to provide. Following the production function literature, we assume that the technology for dialysis treatments follows a Cobb-Douglas form:

$$Y_{it} = A_{it}(q_{it}) K_{it}^{\beta_k} L_{it}^{\beta_\ell}, \tag{1}$$

where Y_{it} is the number of dialysis treatments provided by center i in period t; the capital input, K_{it}, is the number of dialysis stations in center i, while the labor input, L_{it} is the full-time equivalent workforce at the center; and $A_{it}(q_{it})$ is a Hicks-neutral technology shifter that depends on the effort a center puts towards quality, q_{it}, which we model as one minus the center's infection rate (i.e., a lower infection rate reflects higher quality). Note that $A_{it}(q_{it})$ also depends on the productivity of the center, for which we assume the functional form,

$$A_{it}(q_{it}) = e^{\alpha_q q_{it} + \omega_{it} + \epsilon_{it}}, \tag{2}$$

where ω_{it} is a productivity shock that is known to center i in period t, whereas ϵ_{it} is an unanticipated productivity shock that is uncorrelated with all other variables. The parameter α_q measures the magnitude of the quality-quantity tradeoff, and is presumed to be negative.

By taking the logarithm of (1) and letting lower case letters stand for the logarithm of upper case letters, we arrive at the linear equation,

$$y_{it} = \alpha_0 + \alpha_q q_{it} + \beta_k k_{it} + \beta_\ell \ell_{it} + \omega_{it} + \epsilon_{it}. \tag{3}$$

Equation (3) makes apparent the well-known endogeneity problem endemic to production function estimation: because ω_{it} is observed by the firm but not by the econometrician, it may be correlated with the firm's capital and labor choices. Our approach adds an additional endogeneity problem: ω_{it} may affect the firm's quality target. As a result, OLS estimates of (3) are inconsistent. Classical methods of correcting for endogeneity involve applying instruments for capital, labor, and quality, or assuming productivity is fixed over time (i.e., $\omega_{it} = \omega_i$) and using a fixed effects estimator (Mundalk 1961). In practice, these approaches have had limited success. While input prices would seem to be appropriate instruments for capital and labor choices, they have performed poorly in practice and can be difficult to obtain. A valid instrument for quality targets that is uncorrelated with unobserved productivity would be even more challenging to find. Furthermore, while the fixed-effects assumption is relatively easy to implement, it is quite strong and would not resolve the endogeneity problems if changes in productivity are responsible for changes in input (or, in our case, quality) choices.[5]

To address these issues, Olley & Pakes (1996) propose an explicit structural approach to estimate the production process which uses observed firm decisions as proxies for unobserved productivity shocks.[6] The basic ideas of the structural approach have been extended by Levinsohn & Petrin (2003) and Ackerberg et al. (2006). In practice, the detailed assumptions required for the structural approach must be carefully evaluated to determine whether they fit the in-

[5] We have experimented with the fixed effect estimator as a robustness check on our preferred model. Generally, the fixed effect estimator produces similar estimates of the quality-quantity tradeoff.

[6] A second approach to production function estimation has come from the econometric dynamic panel literature (e.g., Blundell & Bond 2000), see Ackerberg et al. (2006) for a comparison of these approaches.

dustry and data under consideration. To our knowledge, our paper is the first attempt to apply this approach to a healthcare setting.

3.2 The Timing of Dialysis Center Decision Making

In their seminal paper, Olley & Pakes (1996) use capital investment as a proxy for unobserved productivity. While natural for their setting of telecommunications equipment, this approach is not appropriate for dialysis centers because investment in new stations is infrequent: investment is zero for over 90 percent of the firm-period observations in the data. In light of this, we focus on firms' hiring decisions. Nurses and technicians employed by dialysis centers require training and credentialing, which introduces costs and time lags to hiring and layoff decisions. Therefore, we regard a firm's hiring decision as a dynamic variable, which allows us to recover ω_{it}. While this assumption conflicts with OP's conception of labor representing an immediately flexible input, though the distinction is natural in our setting. In contrast to labor choices, we assume that a firm can quickly adjust the quality of care it provides. For example, to improve quality, a manager could advise a center's staff to take extra precautions when treating patients, or to reduce quality by placing less emphasis on cleanliness and safety or reducing the duration of treatments. A firms's managers choose new capital investment, new hiring, and its quality target based its knowledge of its capital, labor, productivity, and a vector of other observable characteristics, x_{it} at particular points in time. The vector x_{it} contains variabls which, while they do enter the production function directly, may affect firms' policy choices. These may include the level of competition in the market, the firms tastes for quality (via it's non-profit status), and other demand shifters or patient characteristics. The timing assumptions of our model are as follows:

1. At the end of period $t-1$, firms observe their productivity, $\omega_{i,t-1}$, and state, $x_{i,t-1}$; realize $y_{i,t-1}$; and make investment, $i_{i,t-1}$, and hiring, $h_{i,t-1}$, decisions.[7] Newly hired workers (and newly invested capital) do not become available until period t. That is, the transitions for labor and capital are:

$$k_{i,t} = k_{i,t-1} + i_{i,t-1}, \qquad \ell_{i,t} = \ell_{i,t-1} + h_{i,t-1}$$

2. At time $t-b$, which lies between periods $t-1$ and t, the firm discovers its new observable

[7]Strictly speaking, the investment decision may be made at or before time $t-1$.

state, $x_{i,t}$ (for example, it observes whether additional firms will enter between periods $t-1$ and t), and observes its "interim" productivity, $\omega_{i,t-b}$. The firm also chooses its quality target, $q_{i,t}$ for period t production.

3. At time t, the firm observes $\omega_{i,t}$, realizes production, $y_{i,t}$, and makes its hiring and investment decisions, $i_{i,t}$ and $h_{i,t}$.

In line with the literature, we assume productivity follows an exogenous Markov process between periods $t-1, t-b$, and t:

$$E[\omega_{it-b}|I_{it-1}] = E[\omega_{it-b}|\omega_{it-1}], \qquad E[\omega_{it}|I_{it-b}] = E[\omega_{it}|\omega_{it-b}],$$

where I_{it} represents firm i's information set at time t.

In this setting, unobserved productivity, ω, encompasses any factor that allows a center to treat more patients given its observable characteristics. For instance, a center's patients may follow treatment protocols more closely than others, which then frees the center either (i) to treat more patients because it devotes less time to dealing with complications that arise, or (ii) to spend the extra time treating existing patients more intensively, which ultimately improves outcomes but does not appear in productivity measures.

We follow Ackerberg et al. (2006) in introducting an "intermediate" period $t-b$ when the firm makes its quality decision before fully observing ω_{it}. This period is necessary to avoid multi-colinearity between the firms quality and hiring choices. While this assumption may appear initially artificial, it is natural in our setting. Effectively, we assume that centers make their quality choice prior to period t production being observed, but after they have prepared new capital and labor for production. Therefore, they have learned something about how their productivity has changed from the previous period, but do not perfectly know ω_{it}. As production is observed, they learn their productivity-level before choosing how to adjust labor and capital in preparation for period $t+1$.

3.3 Estimation

This section shows how to estimate the production function parameters and recover each firm's unobserved productivity. We assume that firms behave optimally given the information they have at the time of their decision. Therefore, the firm's hiring decision and quality target are functions of its current state. In addition to the production function variables, we assume that the firm's state includes a vector of center characteristics, x_{it}. That is, we denote the firm's hiring and quality policies as,

$$h_{it} = h(\omega_{it}, k_{it}, \ell_{it}, x_{it}), \qquad q_{it} = q(\omega_{it-b}, k_{it}, \ell_{it}, x_{it}),$$

where x_{it} is a state variable of other factors that affect firm policies but do not directly enter the production function.

In order to invert the hiring function and arrive at a firm's productivity level, we must explicitly control for all factors other than productivity that affect hiring. In our specification, we will include the following sources of variation in x:

For-profit status. Center's differ in their ownership type, with roughly 87.3 percent of the centers operating as for-profit entities and the remainder as non-profit. A center's ownership structure may affect its polices related to hiring and treatment quality, and we therefore control for this distinction by including a dummy variable for the center's for-profit status in x_{it}.

Competition. Because demand for dialysis centers is local, the extent of competition a center faces may affect its hiring and quality choices. For instance, centers in highly competitive markets may choose to improve quality or increase staff to attract patients. As such, we include the level of competition each center faces in x_{it} in the form of dummy variables for having 0, 1, 2, or 3+ competitors in an HSA. We assume that entry is realized at the beginning of the period, so the firm observes its level of competition when it makes its quality and hiring choices.

Vascular Access. Patients receive dialysis through three main types of vascular access: arteriovenous (AV) fistula, AV graft, and venous catheter. A patient's vascular access method influences the likelihood of developing a blood infection, with an AV fistula being significantly less likely to form clots or become infected. Centers vary in the proportion of patients with an AV fistula, which may affect their quality outcomes. And while properly formed fistulas tend to last many years, they require advanced planning because a fistula can take up to 24 months to form fully; as such, they do not represent perfectly variable patient characteristics. To control for these differences, we include the proportion of patients with an AV fistula in x_{it} when calculating the quality-quantity tradeoff. Again, we assume that centers learn their mix of patients for the coming period prior to choosing quality targets.

Under the assumption that a firm's hiring policy is monotonically increasing in productivity, we can invert the hiring function to recover a firm's productivity as a non-parametric function of observables:[8]

$$\omega_{it} = h^{-1}(h_{it}, k_{it}, \ell_{it}, x_{it}). \tag{4}$$

Note that k_{it} and ℓ_{it} have been determined already by virtue of the investment and hiring decisions at time $t-1$, while we assume that the observable state, x_{it}, is revealed to the centers by the intermediate period $t-b$. Since q_{it} is easily adjusted, it is not relevant to the hiring decision for the following period. Our timing assumptions imply that this decision is NOT collinear with the hiring function due to the innovation in productivity between time $t-1$ and $t-b$.[9] That is, q_{it} is a function of $\omega_{i,t-b}$, not $\omega_{i,t}$.

Substituting (4) into (3), we arrive at our first stage estimating equation,

$$\begin{aligned} y_{it} &= \alpha_0 + \alpha_q q_{it} + \beta_k k_{it} + \beta_\ell \ell_{it} + h^{-1}(h_{it}, k_{it}, \ell_{it}, x_{it}) + \epsilon_{it} \qquad (5) \\ &= \alpha_q q_{it} + \Phi(h_{it}, k_{it}, \ell_{it}, x_{it}) + \epsilon_{it}, \end{aligned}$$

where $\Phi(h_{it}, k_{it}, \ell_{it}, x_{it}) = \alpha_0 + \beta_k k_{it} + \beta_\ell \ell_{it} + h^{-1}(h_{it}, k_{it}, \ell_{it}, x_{it})$. As noted in Ackerberg et al. (2006), β_k and β_l cannot be identified directly from (5). Recall that ϵ is independent of all the observables in this equation by the assumption that it is not revealed to the firm when its hiring, investment, or quality decisions are made. If we had a perfect measure of quality, α_q could be consistently estimated from this equation. However, we do not directly observe quality targets, but instead observe only quality-related outcomes (e.g., infection rate, hospital admission rate, death rate, etc.). Since we observe multiple noisy measures of quality, we use one as an instrument to consistently recover α_q from a series estimator using linear instrumental variables. In practice, we use the septic infection rate as an error-ridden proxy for q_{it} and instrument it with the ratio of expected to actual deaths.[10]

We recover the remaining parameters in a second stage. Note that, given any $\beta = (\beta_k, \beta_\ell)$,

[8] There do appear to be some adjustment costs to hiring, as centers hire no workers in roughly 18 percent of center-years. We drop these observations since the hiring function will not be invertible in this range, in line with Olley & Pakes (1996).

[9] Ackerberg et al. (2006) emphasize the importance of this point.

[10] The expected death rate is calculated by Medicare based on individual level data which is not available in our data set.

we can compute the implied unobserved productivity level of each firm,

$$\omega_{it}(\beta) = \hat{\Phi}(h_{it}, k_{it}, \ell_{it}, x_{it}) - \beta_k k_{it} - \beta_\ell \ell_{it}.$$

So, for any β, we can non-parametrically estimate the productivity process,

$$\omega_{it}(\beta) = g(\omega_{it-1}(\beta)) + \xi_{it}(\beta),$$

where g is a non-parametric function of ω_{it-1} and ξ_{it} is a shock to productivity between time $t-1$ and t that is independent of the center's time-t information set.[11] Thus, we can estimate β from the moment condition:

$$E\begin{bmatrix} \xi_{it}(\beta)k_{it} \\ \xi_{it}(\beta)\ell_{it} \end{bmatrix} = 0. \tag{6}$$

We use (6) to estimate $\hat{\beta}$ using the generalized method of moments, which can then be used to recover firm-level productivity estimates, $\omega_{it}(\hat{\beta})$. Standard errors are calculated using the block bootstrap.

4 Results

4.1 Production Function Estimates and the Quality-Quantity Tradeoff

Results from estimates of dialysis centers' production functions are presented in Table 2. In each specification, we use a fourth-order polynomial with interactions to approximate $\Phi(\cdot)$ in the first stage, and a fifth-order polynomial to approximate $g(\cdot)$ in the second stage. Our results are essentially unchanged using a fifth-order polynomial in both stages.

The first column presents results from a specification that does not include the infection rate as a proxy for quality. These results indicate that, as we would expect, providing dialysis treatments is relatively labor intensive, which is in line with production function estimates from similar service industries (e.g., Fox & Smeets 2011). We also find some weak evidence for increasing returns to scale, although we cannot reject the hypothesis of constant returns.

Our first results that include quality in the first stage of the production function estimation appear in Column (II). These results do not use an instrument in the first stage and may be

[11] Since we normalize the mean of ω_{it} to be zero, the constant term of g is a consistent estimator of α_0.

Table 2: Preliminary Production Function Estimates. (These include vascular access)

	I	II	III	IV	V
Quality Effort, α_q		-0.017	-0.017	-0.016	-0.017
		(0.001)	(0.001)	(0.007)	(0.007)
Capital, β_k	0.441	0.443	0.437	0.448	0.439
	(0.113)	(0.093)	(0.136)	(0.089)	(0.069)
Labor, β_ℓ	0.690	0.674	0.686	0.665	0.682
	(0.046)	(0.043)	(0.033)	(0.043)	(0.032)
Controls For:					
Market Structure	NO	NO	YES	NO	YES
Non-Profit Status	NO	NO	YES	NO	YES
Vascular Access	NO	NO	YES	NO	YES
First Stage:	OLS	OLS	OLS	IV	IV

subject to attenuation bias if the realized infection rate is a noisy proxy for centers' quality targets. We find an economically and statistically significant effect of quality on production. The coefficient of -0.017 indicates that, holding inputs fixed, a firm that improves its quality enough that its expected infection rate falls by 1 percent will need to reduce overall patient hours by 1.7 percent.[12] In addition, the slight decline in β_ℓ from Column (I) to Column (II) suggests that firms with lower labor inputs, given capital and output levels, provide lower-quality care. That is, the labor coefficient is biased upwards in Column (I) because firms that appear to use labor more efficiently are actually providing lower-quality output.

The specification in Column (II) assumes that firms' policy functions depend only on their capital stock, employees, and current productivity. It does not include any other factors that could affect firm policies, such as the firm's competitive environment, which are accounted for in our model by x_{it}. As discussed earlier, dialysis centers may have different optimal policies based on the level of competition in their market, their for-profit status, and on their patients' characteristics.

Because our production function estimates rely on the inversion of the hiring policy function with respect to productivity, it is important to control for as many factors that affect hiring as possible. As such, we introduce controls for market structure, for-profit status, and vascular access into the estimation shown in Column (III). In particular, we estimate 8 (4×2) different

[12]We let q_{it} equal one minus the septic infection rate, so that higher q_{it} reflects higher quality.

$\Phi(\cdot)$ functions based on whether the firm faces 0, 1, 2, or 3+ competitors in its HSA and its for-profit status, while $\Phi(k, l, v)$ is specified as a fourth-order polynomial of three variables (capital, labor, and vascular access rate) and a full set of interactions. The production function estimates are essentially unchanged when we add these controls to the hiring policy, which implies that our results are robust to controlling for center-level characteristics.[13] Most importantly, the estimate of the quality-quantity tradeoff does not appear to be affected by the different incentives firms face to offer high or low quality due to their for-profit status, their level of competition, or their patients' characteristics.

As mentioned above, we use the provider's infection rate to proxy for service quality. In practice, providers may attempt to have a low infection rate but their actual rate will be subject to some forces beyond their control. Since effort is necessarily measured with error, attenuation bias affects our first stage estimates in Columns (II) and (III).[14] To correct for potential measurement error, we use a second error-ridden measure of quality, the ratio of deaths to expected deaths of the center's patients, as an instrument for the infection rate in columns (IV) and (V). While the precision of the first stage quality coefficient falls, using an instrument for quality has little effect on the point estimates.

Overall, we find that the level of service quality has a substantial impact on production, even after we control for differences in productivity across firms. A one percentage point increase in a center's infection rate increases production by 1.7 percent, and is robust to the inclusion of relevant controls.[15] These results clearly demonstrate that a quality-quantity tradeoff exists among dialysis service providers.

4.2 Productivity Dispersion, Growth, and Persistance

Having estimated the firm production function we are able to recover center-year (log) productivity levels for each firm in the dataset from

$$\hat{\omega}_{it} = y_{it} - \hat{\alpha}_0 - \hat{\alpha}_q q_{it} - \hat{\beta}_k k_{it} - \hat{\beta}_\ell \ell_{it}.$$

[13]We have also estimated the model controlling for market structure and for-profit status separately and find similar results. These results are available from the authors.

[14]This source of bias is separate from the possibility that quality targets are affected by firm productivity, which is the impetus for our structural estimation procedure.

[15]We have also estimated the model with a subsets of the controls presented here and found similar results

Table 3: Productivity Growth and Persistance

Year	Mean Growth Rate	Standard Deviation of Growth Rate	$\mathrm{Corr}(\omega_{i,t}, \omega_{i,t-1})$
2005	.0867	.4627	.6121
2006	.0820	.4616	.5326
2007	.0541	.4095	.6032
2008	.0567	.4262	.5720

This allows us to analyze the dispersion, growth, and persistence of productivity within the dialysis industry. Moreover, we are able to estimate the importance of productivity for firms' quality and hiring decisions.

To asses the degree of productivity dispersion, we first calculate the proportion of the variance in output that is explained by the production function outside of productivity differences:

$$R^2 = 1 - \frac{V(\hat{\omega}_{it})}{V(y_{it})}.$$

Our results indicate that the amount of productivity dispersion in the dialysis industry is substantial, with $R^2 = .559$, meaning that almost 44 percent of the variation in output is attributable to productivity differences across firms. For a basis of comparison, Fox & Smeets (2011) report R^2 statistics for service industries ranging from .438 (Accounting) to .739 (Computer Activities).

We can then use these productivity estimates to measure productivity growth and persistence within the dialysis industry, as reported in Table 3. On average, productivity growth is large, ranging between 5 and 8 percent per year; however, we observe a large degree of variation in productivity growth within the sample. Moreover, we find only a mild degree of persistence in productivity within a firm, as shown by a correlation in log productivity of approximately 0.6. This suggests that substantial year-to-year productivity shocks affect a center's output. These shocks could result from high staff turnover, patient turnover, or other factors that affect productivity.

4.3 The Sources of Quality

We next consider how controlling for a center's quality choices affects our measured distribution of productivity within the industry. First, we test whether failing to control for quality choices substantially affects industry-level estimates of productivity dispersion. To do so, we compare the R^2 from our preferred specification with one from a specification that does not control for output quality. Not controlling for firm-level quality choices causes the R^2 to fall slightly, from 0.556 to 0.546. Therefore, we conclude that, relative to other unobserved factors, variation in quality does not substantially drive variation in output. However, this is not because quality has no impact on output levels; rather, it is due to the substantial impact of unobserved factors on output.

Given the quality-quantity tradeoff we document in Section 4, we next consider which firms choose to target high levels of quality. Our estimates imply that the policy function for quality choice is a function of capital, labor, the number of local competitors, for-profit status, and productivity levels. To investigate this relationship, we consider the following reduced-form partially linear regression of infection rates on the indicators for for-profit status and the level of competition,

$$q_{it} = \gamma_{c(it)} + \delta_{fp(it)} + f(k_{it}, \ell_{it}, v_{it}, \hat{\omega}_{it}) + \epsilon_{it},$$

where h is a fourth order polynomial with a full set of interactions. In this regression, we use ω_{it} as a proxy for the firm's expected productivity at the time it chooses output quality, ω_{it-b}, which we are unable to recover directly.

We present the results in Table 4. The variables in h are normalized to mean zero, so that the estimates of γ_c are the expected infection rates for non-profit dialysis centers relative to the mean levels of the parameters of f. Surprisingly, we find that monopolist sites tend to offer slightly higher quality (lower infection rate) care than centers facing local competition, while the number of competitors appears to have at most a modest impact on a center's quality. Our results regarding for-profit relative to non-profit centers are more stark. We find that for-profit centers provide substantially lower quality care, with an average infection rate 1.7 percentage points (over ten percent) higher than infection rates at non-profit clinics. These results suggest that, while competitive pressure does not appear to lead firms to target lower quality levels, it does appear that profit motives may incentivize firms to provide lower levels of quality in order to increase output.[16]

[16] We have also investigated the relationship between productivity levels and quality, however there does not appear to be any significant relationship.

Table 4: Reduced form Estimation of Quality Policy Function

	Coeff.	Std. Error
Average Infection Rate:		
Monopolist	11.5047	0.1973
One-Competitor	12.1884	0.1992
Two Competitors	12.2575	0.2249
Three or more Competitors	11.7860	0.1798
For-Profit Dummy	1.3274	0.1523

5 Conclusion

By estimating center-level production functions that incorporate endogenous quality choices, we find evidence that centers reduce the quality of care they provide in order to increase the number of patients they treat, holding inputs and total factor productivity fixed. This result suggests that policies aimed at increasing efficiency may inadvertently affect health outcomes. Although we estimate considerable dispersion in productivity across firms, our findings imply that incentives to cut costs will lead to lower quality care, not greater efficiency. Similarly, our results on non-profit centers also provide evidence that firms react to cost incentives by adjusting their level of quality. Non-profits, which have less incentive to reduce costs, provide higher quality care to patients than their for-profit counterparts, on average.

We find little evidence that market forces provide incentives for high-quality care among dialysis centers by comparing monopoly markets to more competitive ones. While competition might be expected to provide a demand-side incentive to provide high quality, we find that firms in more competitive markets are not more likely to offer higher quality than monopolists, likely because demand for dialysis treatments is inelastic and two dominant providers command two-thirds of the market.

Because dialysis treatments comprise a large — and growing — cost for Medicare, controlling the cost of dialysis provision will likely concern policy makers for the foreseeable future. Our work informs policy by showing that, while productivity dispersion is high within the industry, cost-cutting provisions may simply lower the quality of care provided rather than increase efficiency.

References

Ackerberg, D. A., Caves, K. & Frazer, G. (2006), Structural identification of production function. UCLA, Deloitte and Touche, Rotman School of Management.

Blundell, R. & Bond, S. (2000), 'Gmm estimation with persistent panel data: an application to production functions', *Econometric Reviews* **19**, 321–340.

De Loecker, J. (2012), 'Product differentiation, multi-product firms and estimating the impact of trade liberalization on productivity', *Econometrica* **forthcoming**.

Ford, J. M. & Kaserman, D. L. (2000), 'Ownership structure and the quality of medical care: evidence from the dialysis industry', *Journal of Economic Behavior & Organization* **43**, 279–293.

Fox, J. & Smeets, V. (2011), 'Does input quality drive measured productivity differences in firm productivity', *International Economic Review* **52**(4), 961–989.

Griliches, Z. & Mairesse, J. (1995), Production functions: The search for identification. NBER Working Paper 5067.

Levinsohn, J. & Petrin, A. (2003), 'Estimating production functions using inputs to control for unobservables', *The Review of Economic Studies* **70**(2), pp. 317–341.
URL: *http://www.jstor.org/stable/3648636*

Mundalk, Y. (1961), 'Empirical production function free of management bias', *Journal of Farm Economics* **43**, 44–56.

Olley, G. S. & Pakes, A. (1996), 'The dynamics of productivity in the telecommunications equipment industry', *Econometrica* **64**(6), pp. 1263–1297.
URL: *http://www.jstor.org/stable/2171831*

ProPublica (2011), 'Dialysis: High costs and hidden perils of a treatment guaranteed to all'.
URL: *http://www.propublica.org/series/dialysis*

Syverson, C. (2011), 'What determines productivity?', *Journal of Economic Literature* **49**(2), 326–65.