REGRESSION DISCONTINUITY DESIGNS WITH AN ENDOGENOUS FORCING
VARIABLE AND AN APPLICATION TO CONTRACTING IN HEALTH CARE

Patrick Bajari
Han Hong
Minjung Park
Robert Town

Regression Discontinuity Designs with an Endogenous Forcing Variable and an Application
to Contracting in Health Care
Patrick Bajari, Han Hong, Minjung Park, and Robert Town
NBER Working Paper No. 17643
December 2011
JEL No. D82,I10,L14

ABSTRACT

Regression discontinuity designs (RDDs) are a popular method to estimate treatment effects. However, RDDs may fail to yield consistent estimates if the forcing variable can be manipulated by the agent. In this paper, we examine one interesting set of economic models with such a feature. Specifically, we examine the case where there is a structural relationship between the forcing variable and the outcome variable because they are determined simultaneously. We propose a modi…ed RDD estimator for such models and derive the conditions under which it is consistent. As an application of our method, we study contracts between a large managed care organization and leading hospitals for the provision of organ and tissue transplants. Exploiting "donut holes" in the reimbursement contracts we estimate how the total claims filed by the hospitals depend on the generosity of the reimbursement structure. Our results show that hospitals submit significantly larger bills when the reimbursement rate is higher, indicating informational asymmetries between the payer and hospitals in this market.

Patrick Bajari
Professor of Economics
University of Minnesota
4-101 Hanson Hall
1925 4th Street South
Minneapolis, MN 55455
and NBER
bajari@econ.umn.edu

Han Hong
Landau Economics Building
579 Serra Mall
Stanford, CA 94305
doubleh@stanford.edu

Minjung Park
Haas School of Business
UC Berkeley
545 Student Services Building #1900
Berkeley
CA 94720-1900
minjungp@gmail.com

Robert Town
Health Care Management Department
The Wharton School
University of Pennsylvania
3641 Locust Walk
Philadelphia, PA 19104
and NBER
rtown@wharton.upenn.edu

1 Introduction

Regression discontinuity design is a commonly used method to estimate treatment effects in a non-experimental setting. In an RDD, the researcher searches for a forcing variable that shifts the regressor of interest discontinuously at a known cutoff. If such a variable can be found, an RDD generates a consistent estimate so long as the continuity condition holds (see Hahn, Todd, and Van der Klaauw, 2001). A number of researchers have pointed out that this condition may fail if the forcing variable can be manipulated by an agent (see McCrary, 2008; Lee and Lemieux, 2010; Imbens and Lemieux, 2008). Urquiola and Verhoogen (2009) demonstrate that previously proposed RDD identification strategies in hedonic regressions and public finance may not be valid because economic theories of sorting may predict failure of the continuity condition.

In this paper, we examine one interesting set of economic models with such a feature. They are models where there is a structural relationship between the forcing variable and the outcome variable because the two are determined simultaneously. For instance, suppose that a hospital chooses the optimal level of health care expenditure for a patient given a piecewise linear reimbursement schedule. If a researcher is interested in how reimbursement rates affect the hospital's health care provision, she could potentially exploit discontinuous changes in marginal reimbursement rates to identify the effect. Since marginal reimbursement rates are a function of health care expenditure, health care expenditure would be both the forcing variable and the outcome variable while the reimbursement rate would be the regressor of interest. This is an extreme case of simultaneity since the forcing variable and the outcome variable are identical. Here the forcing variable can clearly be manipulated by the hospital and a standard RDD would not yield consistent estimates.

In this paper, we propose a modified RDD strategy that can be applied to such a model and discuss a set of conditions under which our estimator is consistent. A key idea is that for many choice models, including the one considered in our paper, the optimal

solution implies a strictly monotonic relationship between the type of the agent and the agent's choice, except at a point in which different types of agents will behave identically and will bunch. We exploit this monotonicity to recast the problem such that the type of the agent is seen as the forcing variable. Our estimator is an RDD-style estimator applied to this reformulated problem, and the key idea behind it is that a discontinuous change in the density of the outcome variable at the cutoff can tell us something about the incentive effects.

We apply our estimator to understand a fundamental question in health economics–the responsiveness of health care providers to financial incentives. Hospitals, physicians and other health care providers possess more information about the appropriateness and necessity of care than the patient or, importantly, their insurer. This fact combined with the likelihood that health care providers are concerned with their own financial well-being implies that first-best contracts may be difficult to implement. Understanding the magnitude of this agency problem is a requisite step to both assessing the welfare consequences of provider agency and designing the optimal contracts in health care settings. A central parameter in understanding the importance of provider agency is the responsiveness of providers to changes in reimbursements. Physicians and hospitals control most of the flow of resources in the health care system and medical care expenditures are a large component of most industrialized countries' GDP—in the U.S., health care expenditures are currently over 16% of GDP (Congressional Budget Office, 2008). Thus, the welfare gain from better aligning incentives in these contracts with societal objectives is potentially very large. Despite the importance of this issue and the existence of a large theoretical literature (McGuire, 2000), the convincing empirical literature examining the role of the reimbursement contract structure in affecting provider behavior is relatively sparse.

We have collected a unique data set on contracts between hospitals and one of the largest U.S. health insurers for organ and tissue transplants for all of the hospitals in its network. Organ and tissue transplants are an extremely expensive and rare procedure. In 2007, 27,578 organs were transplanted in the U.S. and the average total billed charges for

kidney transplantation in our data, the least expensive and most commonly transplanted organ, exceed $140,000. The infrequency and complexity of the procedures likely lead to informational asymmetry between hospitals and insurers, making organ transplants an interesting place to examine provider agency. To the best of our knowledge, no other study in the literature has assembled a panel data set of reimbursement contracts between a major private insurer and hospitals of this scope and detail.

The form of the contracts in our data is fairly simple. As a hospital treats patients, it uses its information system to keep track of all reimbursable expenses, which include, but are not limited to, drugs and nights in the hospital. Our hospitals have standard "list prices" for each of these reimbursable expenses. The sum of all of these list prices times the reimbursable items is referred to as "charges." The contract specifies what fraction of the charges submitted by the hospital for each patient will be reimbursed by the insurer. A key feature of the reimbursement schedules is that the total reimbursement amount for each patient follows a piecewise linear schedule: the marginal reimbursement rate changes discontinuously when certain levels of expenditure are reached. This generates discontinuities in the marginal price received by the hospital for its provision of health care.

Using a model of hospitals' optimal health care provision, we verify that in the model the key conditions required for our estimator are satisfied at one of the discontinuity points in the reimbursement schedule. We then apply our estimator to that discontinuity point to estimate the sensitivity of health care provision to the reimbursement rate. Our results clearly show that hospitals will submit significantly larger bills if they face a higher reimbursement rate. When the marginal reimbursement rate changes from 0% to 50%, a magnitude of change typically found in our contracts, the marginal increase in hospitals' expenditures for a given increase in patients' illness severity becomes 2 to 14 times larger. These results suggest that hospitals' behavior is strongly influenced by financial incentives.

The rest of this paper proceeds as follows. In Section 2, we present a model of hospitals' health care choice. In Section 3, we propose our estimation strategy and discuss

4

its sampling properties. Section 4 provides a literature review on agency problems in health care markets. Section 5 describes our data and Section 6 presents model estimates. Section 7 concludes the paper.

2   Model

2.1   The Agency Problem

Consider a health insurer (the principal) that designs compensation contracts for the provider of a medical service (the agent). The insurer's enrolled patients arrive at the hospitals and need potentially costly treatment. Patients differ in their severity of illness which is denoted $\theta \geq 0$, where θ is a random variable with a continuous density function $f(\theta)$ and cdf $F(\theta)$. The health shock, which is determined prior to admission, captures patient heterogeneity in the demand for health care. Patients are passive players in this framework. A central assumption is that patients' heterogeneity is one-dimensional, fully captured by θ.[2] The provider then chooses a level of treatment $q \geq 0$. The value of the health outcome to the patient is given by $v(q, \theta)$, which is twice continuously differentiable. The cost of providing treatment at level q is given by $c(q)$.

The agent (the hospital) observes θ and chooses the level of health care q. The principal (the insurer) cannot observe θ but can observe the hospital's choice of q. Hence, the principal cannot directly contract on the optimal level of q, and instead must rely on a compensation scheme to the agent of the general form $r(q)$ in order to implement the desired q.

To continue, one needs to specify the payoff functions of the principal and the agent. Naturally, the cost of treatment is borne by the agent, and $r(q)$ is paid to the agent by the principal. We assume that the net monetary benefits of the principal are $k - r(q)$,

---

[2]In health settings, patients' heterogeneity is likely to be multi-dimensional. Our analysis assumes that we can summarize the multi-dimensional heterogeneity into a single index.

where k is some fixed payment that he receives from the patient (insurance premium). We assume that the agent's net monetary benefits are just $r(q) - c(q)$. Furthermore, we assume that each party receives a non-pecuniary benefit that is proportional to the patient's payoff. This captures the idea that both the principal and the agent benefit from successful health outcomes.[3] We also assume quasi-linear utility functions so that there are no income effects. We can write the payoffs of the principal and the agent as

$$u^p = \gamma^p v(q, \theta) + k - r(q),$$
$$u^a = \gamma^a v(q, \theta) - c(q) + r(q).$$

Thus, the agent maximizes $\gamma^a v(q, \theta) - c(q) + r(q)$ and the FOC is (for now, ignoring potential non-differentiability in $r(q)$),

$$\gamma^a \frac{\partial v(q, \theta)}{\partial q} = c'(q) - r'(q). \tag{1}$$

The equality in (1) has a simple economic interpretation: the left hand side is the agent's *marginal benefit* from treatment while the right hand side is her *net marginal cost* (total marginal costs less marginal reimbursement).

2.2 Assumptions

We shall assume that the payoffs obey the following conditions:

---

[3]For example, the hospital will value positive patient outcomes if for no other reason than concerns over attracting future patients or deflecting scrutiny by regulators.

$$\frac{\partial v(q,\theta)}{\partial q} > 0 \tag{2}$$

$$\frac{\partial^2 v(q,\theta)}{\partial^2 q} < 0 \tag{3}$$

$$\frac{\partial v(q,\theta)}{\partial \theta} < 0 \tag{4}$$

$$\frac{\partial^2 v(q,\theta)}{\partial\theta\partial q} > 0 \tag{5}$$

$$\frac{\partial c(q)}{\partial q} > 0 \tag{6}$$

$$\frac{\partial^2 c(q)}{\partial^2 q} \geq 0 \tag{7}$$

Assumptions (2) and (3) state that the value of the health outcome to the patient is increasing and strictly concave in q. Assumption (4) implies that health shocks adversely affect utility. Assumption (5) implies that the value of the health outcome to the patient exhibits strictly increasing differences in (q,θ): the marginal utility of health care increases as agents receive more adverse health shocks. According to assumptions (6) and (7), the cost of providing treatment is an increasing and (weakly) convex function in q.

This structure captures the intuitive idea that (i) extra treatments lead to a better health outcome, and the marginal benefit of extra treatments becomes lower as the level of treatment increases; (ii) a more severe condition has a higher marginal benefit of extra treatments; and (iii) providing more treatment costs more money, and marginal treatments are (weakly) more expensive. As a result, when a patient's condition is more severe she should be offered more treatment.

When the agent consumes q dollars of health care to treat a patient, the agent is reimbursed $r(q)$ by the principal. As we discussed in the introduction, we are interested in situations where the constraint set faced by the agent displays kinks. Reflecting the typical reimbursement schedules used by the health insurer in our data, we shall assume

that $r(q)$ satisfies:

$$r(0) = 0 \tag{8}$$

$$r'(q) = \delta_1 \text{ for } 0 < q < q_1 \tag{9}$$

$$r'(q) = 0 \text{ for } q_1 \leq q \leq q_2 \tag{10}$$

$$r'(q) = \delta_2 \text{ for } q > q_2. \tag{11}$$

This assumption implies that the amount of reimbursement for each patient is piecewise linear. For expenditures between 0 and q_1, the hospital is reimbursed δ_1 for every dollar spent to treat the patient. Once expenditures exceed q_1, the hospital hits what is called the donut hole and is forced to bear all of its health care expenses at the margin. Finally, for expenditures above q_2, the hospital is reimbursed δ_2 for every dollar spent. Figure 1 illustrates a reimbursement scheme implied by assumptions (8)–(11). The region $[q_1, q_2]$ is often referred to as the "donut hole." Donut holes are observed in other health care settings as well, most notably Medicare Part D and high deductible health plans with an attached health savings account.

[Figure 1 about here]

In this paper, our main interest lies in understanding hospitals' behavioral responses to the reimbursement structure, not in understanding what the optimal reimbursement scheme should look like. Although the question of if and why the observed contract differs from the optimal one is a very interesting topic,[4] we abstract away from the optimal contract design problem faced by the principal and just condition on the existence of donut holes to learn about the impact of financial incentives on hospital behavior. We note that in reality we might observe an incentive scheme that departs from the optimal

---

[4]For instance, researchers have argued that optimal health contracts should not have donut holes as they pose excessive risk and there are better ways of dealing with moral hazard (Rosenthal, 2004).

one for various reasons, such as institutional constraints or complexity in implementing the optimal contract.[5]

2.3 Optimal Decision Rule

Under the assumptions written above, the optimal decision rule of an agent who treats a pool of patients exhibits the following features:

1. There will be bunching at q_1.

2. There will be a gap near q_2 and the size of the gap crucially depends on the shape of $u^a(q, \theta)$.

3. The optimal choice of q is strictly increasing in θ except for bunching at q_1.

Figure 2 illustrates these observations. In drawing the figure, we assume that $0 < \delta_2 < \delta_1 < 1$, which is what we typically observe in the data. The marginal benefit curve for a given level of θ is decreasing in q, and is given by $\gamma^a \frac{\partial v(q,\theta)}{\partial q}$. The lower is γ^a, the flatter are the marginal benefit curves. A higher θ is associated with a marginal benefit curve that is more to the right. The net marginal cost curve is just $c'(q) - r'(q)$. For this figure, we assume that $c'(q)$ is constant, which is not crucial for any of our results but simplifies the graphical analysis.

[Figure 2 about here]

Imagine a level of θ that corresponds to an optimal choice below q_1. As θ increases, the optimal choice will also increase until some level θ_1 at which it will be exactly q_1. Given the kink in the incentive scheme, there is an upward jump in the net marginal cost

---

[5] In case of Medicare Part D, a donut hole was introduced due to limited government budget available for the program.

curve, causing bunching at q_1 for levels higher than θ_1. At some point, however, high enough levels of θ above θ_1 will cause the marginal benefit curve to shift enough so that optimal choices will exceed q_1 and be on the part of the net marginal cost curve that is $c'(q)$ (i.e., $r'(q) = 0$). The choice of q then continues to rise monotonically with θ until we hit a gap in choices just around q_2, where the net marginal cost drops. To see why we have a gap, consider the level θ^* that is depicted in Figure 2. For this level of severity the agent is indifferent between choosing two levels of health care, one strictly below q_2 (say q_L) and another strictly above (say q_H). By the monotonicity of $q(\theta)$ which follows from the assumption of increasing differences in (q, θ), there will not be any choices of treatment that correspond to expenditures within the interval (q_L, q_H). Finally, for all $\theta > \theta^*$, $q(\theta)$ is strictly increasing.

In Figure 3, an agent with a higher level of γ^a is depicted. The marginal benefit curves of this agent will be shifted up and right compared to those in Figure 2, and they become steeper (a consequence of $v(q, \theta)$ being multiplied by γ^a). This implies that all choices will be shifted to the right (higher levels of q for any given θ), and the steepness of the marginal benefit curves implies that the gap will be small. In this particular example, the size of the gap $[q_L, q_H]$ is almost negligible. The graphical analysis of Figures 2 and 3 offers a complete treatment of what the agent's behavior would be in face of a kinked incentive scheme as described in Figure 1.

[Figure 3 about here]

Throughout our discussion, we have assumed that the agent cannot "cheat" and fraudulently announce costs that were not incurred. If this can happen, then we might observe patterns that are not implied by the optimal decision rule. Although such a fraudulent reporting is not impossible, we think it is uncommon among the hospitals in our data because they are large, established hospitals that are subject to regular audits.

3 Estimation

In this section we propose an estimator that will yield consistent estimates of the agent's behavioral responses when the forcing variable is endogenously chosen by the agent. We first discuss the key intuition behind our approach and then outline our estimation procedures.

3.1 Using Discontinuous Changes for Identification

At the two discontinuity points q_1 and q_2, the marginal reimbursement rate faced by the hospital changes discontinuously. These discontinuities seem to present a natural setting for an RDD. Our problem, however, differs from typical RDD settings because q is both the forcing variable (the level of q determines the marginal reimbursement rate) and the dependent variable (our goal is to estimate how the level of q responds to the marginal reimbursement rate). In this canonical choice model, the forcing variable is clearly endogenous.

One might think that an RDD estimator might be still consistent if there is "optimization error" that prevents agents from precisely controlling the forcing variable (see Lee and Lemieux, 2010). This solution to the problem of an endogenous forcing variable does not work if the forcing variable is structurally related to the dependent variable. If the forcing variable and the outcome variable are structurally related (in our problem, they are the same), when we add optimization error to the forcing variable, we are adding the same optimization error or some transformation of it to the dependent variable. Thus, patients who are on the left hand side of a discontinuity and patients who are on the right hand side of the discontinuity will have systematically different optimization errors added to their outcomes, which will lead to inconsistent estimates under standard RDD estimation.

In our paper, we propose an alternative solution to the problem. A key step in our

approach is to transform the problem so that we make the type of the patient θ a forcing variable. From the earlier discussion, and more generally the monotone comparative statics literature of Topkis (1978) and Milgrom and Shannon (1994), we know that the assumption of strictly increasing differences in (θ, q) implies that the optimal health care provision q is a strictly increasing function of patient type θ, with the exception of where there is bunching at q_1. As a result, the percentiles of q will identify θ. That is, if we see a patient with the 5th percentile of health expenditure within a hospital, that patient will have the 5th percentile of health shock within that hospital. This means that for all practical purposes, the health shocks are observable to the econometrician. Since q is only weakly increasing in θ around the first discontinuity point due to the presence of bunching, the econometrician cannot infer θ from the cdf of q in the region. Hence, our estimation procedure can be applied to the second kink, but not the first one.

Once we reformulate the problem so that the patient type θ is viewed as a forcing variable (which is exogenously endowed and cannot be manipulated), a shift in the patient type θ determines whether the hospital's choice of q for that patient will be on the left hand side or right hand side of the second discontinuity point. This then generates an exogenous change in the marginal price faced by the hospital, allowing for identification of the hospital's response to incentives. Our approach boils down to estimating a variant of regression discontinuity models in the empirical quantile function of the hospital's choice q.

Figure 4 illustrates the idea behind our approach. Among patients who come to the hospital with a realization of health shock θ, there will be a value of θ at which the hospital is indifferent between choosing q_L ($< q_2$) and q_H ($> q_2$). Let θ^* denote the level of severity which leads to such an indifference. Then for all patients whose θ is greater than θ^*, the hospital will choose q larger than q_H and will face a marginal reimbursement rate of δ_2. For patients whose θ is smaller than θ^*, the hospital will choose q smaller than q_L and will face a marginal reimbursement rate of 0. Thus, the hospital's supply of health care services will be more responsive to an increase in θ on the right hand side of θ^* than on the

12

left hand side of θ^* as long as the hospital is price sensitive. Therefore, by comparing how quickly q rises with an increase in θ for values of θ just below and just above θ^*, we can infer how the total claims filed by the hospital depend on the reimbursement structure. We let ϕ_{SLOPE} denote the change in the slope of the quantile function, $q'(\theta)$, at θ^*. Note that the slope of the quantile function $q'(\theta)$ is by definition equal to the inverse of density.

[Figure 4 about here]

The figure also shows a possibility of gap $\phi_{GAP} = q_H - q_L$ at θ^*. As discussed earlier, for a high value of γ^a the gap could be very small, while for a low value of γ^a the gap might be large.

It is worthwhile to note key underlying assumptions behind our approach. A key identifying assumption is that the slope of the quantile function would be the same at θ^* from both sides if there were no change in the marginal reimbursement rate at θ^*. In other words, the density of q would be continuous at θ^* in the absence of a discrete change in incentives at θ^*. This assumption allows us to attribute any discrete change in the slope of the quantile function at θ^* to a discrete change in the financial incentive. A similar type of continuity assumption is found in the conventional RDD (Hahn, Todd, and Van der Klaauw, 2001) and in McCrary (2008) who proposes a test for manipulation of a forcing variable by examining discontinuity in the density function of the forcing variable at the cutoff point.

In order to correctly map q to θ, we also require that there be no error in the hospital's choice of $q(\theta)$. If $q(\theta)$ contains error in it, we cannot infer patient type θ from the observed q. In reality, $q(\theta)$ is very likely to have error in it since hospitals cannot perfectly control the level of treatment: there is lumpiness in treatments, there could be some unforeseen events that make it more costly to treat a less sick patient, etc. Since such error would lead to measurement error in inferred θ, we expect that our estimator could suffer from a downward bias. Intuitively, when θ is measured with error, the difference between the

slopes of the quantile function on the LHS and RHS of the threshold will be smoothed out, leading to a downward bias. In the extreme case, if q is determined entirely randomly without any relation to patient sickness and the hospitals have no control over q, we would not observe any difference in the slopes of the quantile function between the two sides of θ^*.

3.2   Estimation

We consider several different estimation approaches that deal with different levels of hospital heterogeneities. The first method applies to individual hospital data. The second method makes use of a global parametric assumption to pool information from data across all hospitals.

3.2.1   Individual Hospital Estimates

Suppose that there are $i = 1, ..., n$ individuals treated in the hospital under consideration. Let q_i denote the health expenditure of individual i. Let $\hat{F}(\cdot)$ denote the empirical distribution of the observed q's for the hospital. We propose estimators for ϕ_{GAP} and ϕ_{SLOPE} at the upper regression discontinuity point of q_2 and derive the asymptotic distribution of the estimators.

The incentive scheme is such that for θ approaching a cutoff value θ^*, $q(\theta)$ approaches a limit value q_L. As soon as θ moves to the right of θ^*, $q(\theta)$ takes a discrete jump at the point of θ^* by an amount $\phi_{GAP} > 0$ to q_H.

By normalization, θ is estimated as the empirical CDF of the observed q. Hence θ^* is estimated by

$$\hat{\theta}^* = \hat{F}(q_2) = \frac{1}{n} \sum_{i=1}^{n} 1(q_i \leq q_2),$$

where $\hat{F}(\cdot)$ is the empirical distribution of the observed q's. Given that we define $\theta^* =$

$F(q_2)$ where $F(\cdot)$ is the true distribution function of q, the asymptotic distribution of $\hat{\theta}^*$ is immediate:

$$\sqrt{n}(\hat{\theta}^* - \theta^*) \xrightarrow{d} N\left(0, F(q_2)(1 - F(q_2))\right).$$

We are interested in estimating the magnitude of the discontinuity ϕ_{GAP}. This is estimated by

$$\hat{\phi}_{GAP} = \hat{q}_H - \hat{q}_L = \min\{q_i : q_i > q_2\} - \max\{q_i : q_i \le q_2\}.$$

The goal is to derive the asymptotic distribution of $\hat{\phi}_{GAP} - \phi_{GAP}$. It suffices to show that the joint distribution of $n(\hat{q}_L - q_L)$ and $n(\hat{q}_H - q_H)$ are independent exponential distributions. To see this, note that

$$
\begin{aligned}
P(n(\hat{q}_L - q_L) &\le -x, n(\hat{q}_H - q_H) \ge y) \\
&= P(q_i \le q_L - x/n, q_i \ge q_H + y/n, \forall i) \\
&= (1 - P(q_L - x/n \le q_i \le q_H + y/n))^n \\
&= \left(1 - f^- x/n - f^+ y/n + o(1)/n\right)^n \xrightarrow{n \to \infty} e^{-f^- x - f^+ y}.
\end{aligned}
$$

In other words, $n(\hat{q}_L - q_L)$ and $n(\hat{q}_H - q_H)$ converge to two independent (negative and positive) exponential random variables with hazard rates $f^- = f(q_L)$ and $f^+ = f(q_H)$, where we have used f^- and f^+ to denote the (left and right) densities at q_L and q_H. The limiting distribution of $n(\hat{\phi}_{GAP} - \phi_{GAP})$ is therefore the sum of these two independent exponential random variables.

Next we turn to the estimation of difference between the slopes of $q(\theta)$ at q_H and q_L, defined as $\phi_{SLOPE} = \lim_{\theta \to \theta_+^*} q'(\theta) - \lim_{\theta \to \theta_-^*} q'(\theta)$. Note that $\phi_{SLOPE} = \frac{1}{f^+} - \frac{1}{f^-}$. Hence it suffices to obtain consistent nonparametric estimators for f^+ and f^-. This can be done using standard one sided kernel smoothing methods.

Define

$$\hat{f}^- = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h}k(\frac{\hat{q}_L - q_i}{h})1(q_i \leq \hat{q}_L),$$

and

$$\hat{f}^+ = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h}k(\frac{q_i - \hat{q}_H}{h})1(q_i \geq \hat{q}_H).$$

In the above, $k(\cdot)$ is a one-sided density function supported on $(0, \infty)$, and h is a sequence of bandwidth parameters used in typical kernel smoothing. It is straightforward to show that as long as $nh \longrightarrow \infty$ and $nh^3 \longrightarrow 0$,

$$\sqrt{nh}\left(\hat{f}^- - f^-\right) \xrightarrow{d} N\left(0, f^-\int K(u)^2 du\right),$$

and

$$\sqrt{nh}\left(\hat{f}^+ - f^+\right) \xrightarrow{d} N\left(0, f^+\int K(u)^2 du\right)$$

and that they are asymptotically independent. Therefore

$$\sqrt{nh}\left(\hat{\phi}_{SLOPE} - \phi_{SLOPE}\right) \xrightarrow{d} N\left(0, \frac{1}{f^{-3}}\int K(u)^2 du + \frac{1}{f^{+3}}\int K(u)^2 du\right).$$

3.2.2   A Parametric Model Using Multiple Hospital Data

Now we consider how to extend the previous method to allow for pooling information from data across multiple hospitals. Consider first $\phi_{GAP} = q_H - q_L$. We define $y_i = q_i 1(q_i \leq q_2)$ and $z_i = q_i 1(q_i > q_2) + M1(q_i \leq q_2)$. In the homogeneous case, we have defined $\hat{q}_L = \max\{y_i\}$ and $\hat{q}_H = \min\{z_i\}$, where M is a number that is larger than any of the data points. This definition of the estimators can be extended to incorporate heterogeneous data from all hospitals.

With cross-hospital data, the observed threshold value q_2 can be hospital dependent, which we will denote as $q_2(t)$, where we have used t to index hospitals. Suppose hospital heterogeneity is captured by covariates x_t, where x_t can be $q_2(t)$ itself. Let I_t be the

number of patient observations for each hospital. We specify the following parametric assumption that

$$q_L(t) \equiv q_L(x_t) = g_L(x_t, \beta_L) \quad \text{and} \quad q_H(t) \equiv q_H(x_t) = g_H(x_t, \beta_H).$$

In the above, we can use a flexible series expansion functional form of $g_L(x_t, \beta_L)$ and $g_H(x_t, \beta_H)$ so that they are linear in the parameters β_L and β_H. The structure of this problem fits into the boundary parameter estimation method studied in the literature. Possible estimators include the linear programming approach and nonstandard likelihood estimator (c.f. Donald and Paarsch, 1996; Chernozhukov and Hong, 2004) and the extreme quantile regression approach of Chernozhukov (2005). We describe these alternatives in the following.

The linear (or quadratic, etc.) programming approach estimates the parameters by

$$\hat{\beta}_L \quad = \quad \arg\min_{\beta_L} \sum_{t=1}^{T} I_t^L g_L(x_t, \beta_L), \quad \text{where} \quad I_t^L = \sum_{i=1}^{I_t} 1(y_i > 0),$$
$$\text{such that} \quad y_i \leq g_L(x_t, \beta_L), \forall i = 1, ..., I_t^L, t = 1, ..., T, \text{ (ordering implicitly understood)}$$

and

$$\hat{\beta}_H \quad = \quad \arg\max_{\beta_H} \sum_{t=1}^{T} I_t^H g_H(x_t, \beta_H), \quad \text{where} \quad I_t^H = \sum_{i=1}^{I_t} 1(z_i < M),$$
$$\text{such that} \quad z_i \geq g_H(x_t, \beta_H), \forall i = 1, ..., I_t^H, t = 1, ..., T. \text{ (ordering implicitly understood)}$$

The objective functions $\sum_{t=1}^{T} I_t^L g_L(x_t, \beta_L)$ and $\sum_{t=1}^{T} I_t^H g_H(x_t, \beta_H)$ can be replaced by

$$\sum_{t=1}^{T} \sum_{i=1}^{I_t} 1(y_i > 0)(y_i - g_L(x_t, \beta_L))^2 \text{ and } \sum_{t=1}^{T} \sum_{i=1}^{I_t} 1(z_i < M)(z_i - g_H(x_t, \beta_H))^2$$

or other types of penalization functions. The linear programming approach however seems to be the easiest to implement.

Alternatively, β_L and β_H can be estimated by the extreme quantile regression method of Chernozhukov (2005):

$$\hat{\beta}_L = \arg\min_{\beta_L} \sum_{t=1}^{T} \sum_{i=1}^{I_t} 1(y_i > 0) \rho_{\tau_L}(y_i - g_L(x_t, \beta_L)),$$

where $\rho_\tau(u) = (\tau - 1(u \le 0))u$ is the *check function* of Koenker and Bassett (1978), such that

$$\tau_L \to 1 \quad \text{as} \quad n_L = \sum_t I_t^L \to \infty.$$

Similarly, $\hat{\beta}_H = \arg\min_{\beta_H} \sum_{t=1}^{T} \sum_{i=1}^{I_t} 1(z_i < M) \rho_{\tau_H}(z_i - g_H(x_t, \beta_H))$, where

$$\tau_H \to 0 \quad \text{as} \quad n_H = \sum_t I_t^H \to \infty.$$

The quantile regression approach has the advantage of being robust against a certain fraction of outliers in the data. On the other hand, the programming estimators always satisfy the constraints of the relation between y_i and $g_L(x_t, \beta_L)$ and between z_i and $g_H(x_t, \beta_H)$.

By adopting a parametric functional form on $q_L(x_t)$ and $q_H(x_t)$ we are maintaining a strong specification assumption which can potentially be tested by the data. An implicit assumption of the parametric functional form is that $g_L(x_t, \beta_L^0) \le q_2(t) \le g_H(x_t, \beta_H^0)$ for all t at the true parameters β_L^0 and β_H^0. Of course their estimates introduce sampling noise, but we still expect that it should be largely true for most t:

$$g_L(x_t, \hat{\beta}_L) \le q_2(t) \le g_H(x_t, \hat{\beta}_H).$$

The approximate validity of this condition can be used as the basis of a model specification test.

Then $\phi_{GAP}(x_t)$ will be estimated consistently by

$$\hat{\phi}_{GAP}(x_t) = g_H(x_t, \hat{\beta}_H) - g_L(x_t, \hat{\beta}_L).$$

Conducting statistical inference on $\hat{\phi}_{GAP}(x_t)$ requires the limiting *joint* distribution of $\hat{\beta}_L$ and $\hat{\beta}_H$. They converge to a nonstandard distribution at a fast $1/n$ rate for $n = \sum_t I_t$. The limiting distribution can be obtained by simulation which we will describe below in the context of the parametric likelihood approach.

In fact we can also adopt a maximum likelihood approach. This will be useful in case we are interested in the shape of the distribution of q_{it} in order to conduct counter-factual welfare calculations. To this end, assume that

$$\epsilon_{it}^L = g_L(x_t, \beta_L) - y_{it} \sim f_L(\epsilon_{it}^L, x_t, \alpha_L) \text{ for } y_{it} \le g_L(x_t, \beta_L),$$

and

$$\epsilon_{it}^H = z_{it} - g_H(x_t, \beta_H) \sim f_H(\epsilon_{it}^H, x_t, \alpha_H) \text{ for } z_{it} \ge g_H(x_t, \beta_H).$$

The maximum likelihood estimator for α_L, α_H and β_L, β_H can then be written as

$$
\begin{aligned}
(\hat{\alpha}_L, \hat{\beta}_L) &= \arg\max_{\alpha_L, \beta_L} \sum_{t=1}^{T} \sum_{i=1}^{I_t} 1(y_{it} > 0) \log f_L(g_L(x_t, \beta_L) - y_{it}, x_t, \alpha_L) \\
&\quad \text{such that} \quad y_{it} \le g_L(x_t, \beta_L), \forall i = 1, ..., I_t, t = 1, ..., T,
\end{aligned}
$$

and

$$
\begin{aligned}
(\hat{\alpha}_H, \hat{\beta}_H) &= \arg\max_{\alpha_H, \beta_H} \sum_{t=1}^{T} \sum_{i=1}^{I_t} 1(z_{it} < M) \log f_H(z_{it} - g_H(x_t, \beta_H), x_t, \alpha_H) \\
&\quad \text{such that} \quad z_{it} \ge g_H(x_t, \beta_H), \forall i = 1, ..., I_t, t = 1, ..., T.
\end{aligned}
$$

In fact the linear programming estimator is a special case of the above maximum likelihood estimator when the densities $f_L(\epsilon_{it}^L, x_t, \alpha_L)$ and $f_H(\epsilon_{it}^H, x_t, \alpha_H)$ are exponential distribution

with a homogeneous hazard rate parameter: $f(\epsilon) = \lambda e^{-\lambda \epsilon}$. In this case, in addition to obtaining $\hat{\beta}_L$ and $\hat{\beta}_H$ from the linear programming estimators, we also estimate the hazard parameters by

$$1/\hat{\lambda}_L = \frac{\sum_{t=1}^{T} \sum_{i=1}^{I_t} 1(y_{it} > 0) \left(g_L(x_t, \hat{\beta}_L) - y_{it} \right)}{\sum_{t=1}^{T} \sum_{i=1}^{I_t} 1(y_{it} > 0)}$$

and

$$1/\hat{\lambda}_H = \frac{\sum_{t=1}^{T} \sum_{i=1}^{I_t} 1(z_{it} < M) \left(z_{it} - g_H(x_t, \hat{\beta}_H) \right)}{\sum_{t=1}^{T} \sum_{i=1}^{I_t} 1(z_{it} < M)}$$

Even though $\hat{\beta}_L$ and $\hat{\beta}_H$ converge at $1/n$ rate to a nonstandard limit distribution, $\hat{\alpha}_L$ and $\hat{\alpha}_H$ are still root n consistent and asymptotically normal, as long as there is no functional relations between α and β.

To estimate $\phi_{SLOPE}(x_t)$, we can use

$$\hat{\phi}_{SLOPE}(x_t) = \frac{1}{f_H(0, x_t, \hat{\alpha}_H)} - \frac{1}{f_L(0, x_t, \hat{\alpha}_L)}.$$

Since $\hat{\phi}_{SLOPE}(x_t)$ is root n consistent and asymptotically normal, its limiting distribution can be obtained by the delta method combined with the standard sandwich formula, or by simulation or bootstrap, in which $\hat{\beta}_L$ and $\hat{\beta}_H$ can be held fixed because they do not affect the asymptotic distribution.

The joint asymptotic distribution for $\hat{\beta}_L$ and $\hat{\beta}_H$ can be obtained by parametric simulations. Given the assumption that the parametric model is correctly specified, it is possible to simulate from the model using the estimated parameters $\hat{\beta}_L$, $\hat{\beta}_H$, $\hat{\alpha}_L$ and $\hat{\alpha}_H$. The approximate distribution can be obtained from repeated simulations. Instead of recomputing the maximum likelihood estimator at each simulation, it suffices to recompute weighted programming estimators of β_L and β_H at each simulation:

$$\tilde{\beta}_L = \arg\min_{\beta_L} \sum_{t=1}^{T} \sum_{i=1}^{I_t} f_L(0, x_t, \hat{\alpha}_L) \frac{\partial g_L(x_t, \hat{\beta}_L)'}{\partial \beta_L} \beta_L$$
$$\text{such that} \quad y_{it} \le g_L(x_t, \beta_L) \quad \forall i, t,$$

20

and

$$\tilde{\beta}_H = \underset{\beta_H}{\arg\max} \sum_{t=1}^{T} \sum_{i=1}^{I_t} f_H(0, x_t, \hat{\alpha}_H) \frac{\partial g_H(x_t, \hat{\beta}_H)'}{\partial \beta_H} \beta_H$$
$$\text{such that} \quad z_{it} \geq g_H(x_t, \beta_H) \quad \forall i, t.$$

We can also consider the possibility that $g_L(x_t, \beta_L)$ and $g_H(x_t, \beta_H)$ are correctly specified but $f_L(\epsilon_{it}^L, x_t, \alpha_L)$ and $f_H(\epsilon_{it}^H, x_t, \alpha_H)$ are misspecified. In this case, each of the above methods (linear and quadratic programmings, extreme quantile regression, (pseudo) maximum likelihood estimation) will still deliver consistent estimates of β_L and β_H and hence ϕ_{GAP}. But the estimates for α_L, α_H and hence ϕ_{SLOPE} are clearly inconsistent.

In this case, if we are willing to impose parametric assumptions on ϕ_{GAP} through $g_L(x_t, \beta_L)$ and $g_H(x_t, \beta_H)$, but are not willing to make parametric assumption on ϕ_{SLOPE}, we can still estimate ϕ_{SLOPE} using nonparametric density estimators. We can also use nonparametric density estimators to perform semiparametric simulations for consistent inference about $\hat{\phi}_{SLOPE}$. Suppose x_t is continuously distributed with dimension d. Let

$$\hat{f}^-(x) = \frac{1}{n} \sum_{t=1}^{T} \sum_{i=1}^{I_t} \frac{1}{h} w(x_t, x) k(\frac{g_L(x_t, \hat{\beta}_L) - q_{it}}{h}) 1(q_{it} \leq g_L(x_t, \hat{\beta}_L)),$$

and

$$\hat{f}^+(x) = \frac{1}{n} \sum_{t=1}^{T} \sum_{i=1}^{I_t} \frac{1}{h} w(x_t, x) k(\frac{q_{it} - g_H(x_t, \hat{\beta}_H)}{h}) 1(q_{it} \geq g_H(x_t, \hat{\beta}_H)),$$

where

$$w(x_t, x) = k^d(\frac{x_t - x}{h}) / \sum_{t=1}^{T} k^d(\frac{x_t - x}{h}).$$

Then we can form the estimate $\hat{\phi}_{SLOPE}(x_t) = 1/\hat{f}^+(x_t) - 1/\hat{f}^-(x_t)$.

The limiting distribution of the MLE's $\hat{\beta}_L$ and $\hat{\beta}_H$ can be obtained by recomputing

the following weighted programming estimators with simulated data:

$$\bar{\beta}_L \;=\; \arg\min_{\beta_L} \sum_{t=1}^{T} \sum_{i=1}^{I_t} f^-(x_t) \frac{\partial g_L(x_t, \hat{\beta}_L)'}{\partial \beta_L} \beta_L$$
$$\text{such that} \quad y_i \leq g_L(x_t, \beta_L) \quad \forall i, t,$$

and

$$\bar{\beta}_H \;=\; \arg\max_{\beta_H} \sum_{t=1}^{T} \sum_{i=1}^{I_t} f^+(x_t) \frac{\partial g_H(x_t, \hat{\beta}_H)'}{\partial \beta_H} \beta_H$$
$$\text{such that} \quad z_i \geq g_H(x_t, \beta_H) \quad \forall i, t.$$

As before, the simulated distributions of $n(\bar{\beta}_L - \hat{\beta}_L)$ and $n(\bar{\beta}_H - \hat{\beta}_H)$ should approximate the limit distributions of the maximum likelihood estimates $n(\hat{\beta}_L - \beta_L^0)$ and $n(\hat{\beta}_H - \beta_H^0)$.

4 Literature on Provider Agency in Health Care Markets

Health care markets are rife with informational asymmetries which can be leveraged by health care providers to increase incomes relative to full information, first-best equilibrium. Arrow (1963) noted that a first-best insurance contract would specify a state-dependent payment. However, these contracts are not generally negotiated because health states are not readily observed. Since the work of Arrow (1963), a large theoretical literature has arisen characterizing the optimal payment contract under different informational, preference and market structure scenarios.[6] However, most current provider-insurer contracts do not correspond to the structures generally prescribed by theory.[7] The failure

---

[6]McGuire (2000) provides an excellent review of this literature.

[7]For example, private insurers generally pay physicians on a fee-for-service or percentages of billed charges basis, while Medicare pays physicians on a fee-for-service basis and hospitals by groupings of

of insurers to negotiate first-best contracts suggests that there is meaningful scope for provider agency. Given the size of the health care sector (approximately 16% of GDP in the U.S.), the potential welfare consequences of provider agency are extremely large.

Empirical analyses of the magnitude of the agency problem date to the work on physician induced demand of Fuchs (1978). Strong circumstantial evidence that provider behavior deviates from the socially optimal one is provided by the *Dartmouth Atlas Project*. They find that there are large geographic variations in utilization by Medicare enrollees that are unrelated to health status. The *Dartmouth Atlas Project* suggests (but only provides limited econometric evidence) that the geographic differences are driven by geographic variation across providers in demand inducement. Following Fuchs (1978) a literature has arisen that attempts to estimate the degree of physician demand inducement, but the identification strategies employed in these papers are generally suspect.[8] There are important exceptions, however. Gruber and Owings (1996) find that within state declines in fertility are associated with increases in cesarean sections. Yip (1997) shows that cardiac surgeons responded to payment reductions by increasing the number of procedures they performed. During the 1980s, Medicare changed its hospital reimbursement system from retrospective to prospective using Diagnostic Related Groups (DRG) as the basis of the payment. The incentive under prospective payment is to reduce the length of stay of Medicare beneficiaries and the policy appeared to have had the expected impact (Hodgkin and McGuire, 1994) without dramatically impacting the quality of care (Cutler, 1995). Even within the DRG system, hospitals appear to leverage their superior information into more generous payments. Dafny (2005) finds that when Medicare changed the structure of the DRG payment generosity, hospitals responded by upcoding patients into more generous payment groups.

There has been little detailed analysis of the response of providers to the specific in-

diagnoses (Diagnosis Related Groups). Generally, these contract structures are not optimal under most contracting models.

[8] See Dranove and Wehner (1994) for a discussion of the limitation of the attempts to estimate physician agency.

centives embedded in reimbursement and remuneration contracts.[9] Gaynor and Gertler (1995) find that physicians reduce their effort when faced with lower powered incentives. Gaynor, Rebitzer and Taylor (2004) analyze a model of physician behavior under group incentives and test the predictions using detailed contract and financial data from a network HMO. They rely on variation in the size of the panel over which the group incentive is implemented to identify the parameters. They find that the HMO's incentive contract provides a typical physician with an increase, at the margin, of $0.10 in income for each $1.00 reduction in medical utilization expenditures. The presence of these high powered incentives reduced medical expenditures by 5%. More recently, Ketcham, Léger and Lucarelli (2011) examine the impact of physician gainsharing on the use of medical devices. In a gainsharing arrangement hospitals pay physician groups if they reduce the hospital's cost of medical device acquisition and utilization. They find that gainsharing arrangements yield significant reductions in hospitals' medical device costs.

5 Data

The focus of our empirical analysis is organ and tissue transplants, and this choice is motivated by the following considerations. For the purpose of our study, we need reimbursement schedules that have discrete changes in marginal reimbursement rates. Furthermore, we need to study the type of procedures that have potentially significant informational asymmetries between hospitals and insurers and thus leave ample opportunity for provider agency. Finally, we need detailed data on the shape of the reimbursement schedule that each hospital faces as well as data on patients treated in each hospital. Organ and tissue transplants fit all these criteria—piecewise linear reimbursement schedule, significant informational asymmetry due to the complexity of the procedures, and availability of data on contracts and patients from a leading insurer.

---

[9] The state of the empirical literature is well summarized by Glied (2000), "There is very little empirical evidence on the behavior of physicians paid using different payment arrangements."

Our estimation uses two data sets from the same source. The first data set comes from one of the largest private health insurers in the U.S. and has information on its contracts with 127 hospitals which specify reimbursement schedules for transplants. The second data set contains information on the set of patients who received organ or tissue transplants in each of the 127 hospitals. We merge the two, and the resulting data set has (*i*) claim-level information, such as the admission and discharge dates of the patient, the type of transplant received by the patient, the size of the bill submitted by the hospital to the insurer and the reimbursement amount paid by the insurer, as well as (*ii*) hospital-level information, such as the name and location of the hospital and the reimbursement schedule the hospital faces for each type of organ transplant surgery it performs. The data run from 2004 through 2007.

The insurer uses this network of hospitals for its own enrollees and also sells access to this network to other health insurers and self-insured employers. This insurer is a major player in the organ transplant market, with 80% market share among private vendors.

There are various types of organ and tissue transplants covered by the contracts, major ones being bone marrow transplant (BMT), kidney transplant, liver transplant, heart transplant and lung transplant. Organ and tissue transplants are a rare and exceedingly expensive procedure. In 2007, 27,578 organs were transplanted in the U.S. and the average total billed charges for kidney transplantation in our data, the least expensive and most commonly transplanted organ, exceed $140,000. Between 2005 and 2008, the cost of organ transplant rose at an annual rate of 14%—a rate that is larger than general health care cost inflation. An organ transplant is an extremely challenging and complex procedure taking anywhere from 3 (kidney) to 14 hours (liver). Organ transplants usually require significant post-operative care (up to 3 weeks of inpatient care) and careful medical management to prevent rejection. The infrequency of the procedures, the complexity of the treatments and the large variation across patients in their response to transplantation make it difficult for insurers to determine the appropriateness of the care for a given episode. That, in turn, implies that hospitals are in a position to engage in agency behavior in response to

the incentives embodied in their contracts.

The insurer negotiates a separate contract with each individual hospital, instead of having one common contract applied to all participating hospitals. As a result, the reimbursement schedule differs across hospitals. For about 75% of hospitals in our original sample, the reimbursement schedule takes a form as shown in Figure 1 with two kinks (the marginal reimbursement rate starts at positive, becomes zero for a certain range and then becomes positive again). The remaining 25% of hospitals have contracts that have only one kink (the marginal reimbursement rate starts at positive and then remains at zero above a certain expenditure level). Under the second type of contract, the maximum amount of reimbursement is capped at a fixed level, while the maximum reimbursement increases with billed charges under the first type of contract. As a result, hospitals are exposed to greater risk under the second type of contract. Even among hospitals that have the first type of contract, there is a large variation in the locations of the first kink (q_1) and the second kink (q_2), the marginal reimbursement rate for each of the segments (δ_1 and δ_2) and the height of the donut hole ($\delta_1 q_1$). These differences in the contract type and contract terms likely reflect variation in bargaining power as well as heterogeneity in the patient pool across hospitals. For instance, we find that larger hospitals (presumably with greater bargaining power) are more likely to have the first type of contract. Also, conditional on having the first type of contract, larger hospitals are likely to have higher marginal reimbursement rates δ_1 and δ_2 (see Ho (2009) for a nice discussion of hospitals' bargaining power and markups). In our analysis, we focus on hospitals whose reimbursement schedules display two kinks.

Our empirical measure of q is charges that hospitals submit to the insurer, which is the sum of list prices times the quantities of all reimbursable items.[10] Since the list prices

---

[10]To be precise, q measures charges incurred between admission and 90 days after discharge. This period includes most of the major components related to transplant care, such as organ procurement, transplant operation, inpatient care and necessary follow-ups within 90 days post discharge. Typically, 75% of the total costs associated with transplant care occur during this period. The reimbursement schedules we examine apply to charges incurred during this period only, and there are separate provisions for charges incurred prior to admission or after more than 90 days post discharge.

of reimbursable items are set above their marginal costs, the charges are higher than the costs that hospitals incur to treat the patients. While q represents "quantity" in our model in Section 2, our empirical measure of q is in dollars. As a result, one might worry about potential discrepancies between the two. For instance, it is well known that the chargemaster (the file in which list prices are kept) varies significantly across hospitals, in which case we would observe different charges for two patients in two different hospitals even if they received the same level of treatment. Or if newer drugs cost more even though they treat less severe conditions, the ordering of patients based on our empirical q would differ from the ordering based on patient sickness, violating the monotonicity condition required for our approach. In our analysis, ordering of patients will be done separately for each hospital and organ type in order to partially address the concern.

One practical issue we encounter is that the number of patients who receive a certain type of organ transplant within a hospital is typically small. To deal with this issue, we pool observations across years for a given hospital and organ type (as long as the reimbursement structure does not change over time) since it seems plausible to expect that a given hospital's price sensitivity does not change during the short sample period. To further reduce the potential bias arising from the small number of patients, we restrict our attention to (hospital, organ) pairs that have enough observations.[11] Table 1 presents summary statistics for our estimation sample.

[Table 1 about here]

From the table, it is clear that there is a huge variation in charges. A simple regression shows that about 15%-25% of variation in charges is explained by hospital dummies for each of the organ types. This could be due to differences across hospitals in patient pool, list prices or innate resource use intensity. Ideally, we would closely examine the various components of the charges—the costs of organ procurement, hospitalization, tests, drugs,

---

[11]In our result tables below we report how many patients each hospital has.

etc. for each patient. Our data essentially is the information that the insurer receives from the hospital and such detail is not transmitted to the insurer and is generally not available.

The lack of information on detailed components of charges also prevents us from empirically examining what hospitals do in practice to adjust their level of care q in face of financial incentives. However, we believe that there are margins that can be moved around to control charges. Note that the q measure in our application includes post-operative care for some period of time. During this time period, hospitals have significant discretion over q. Hospitals can release patients earlier or later, depending on how sick the patient is, and also potentially depending on the reimbursement structure. Hospitals have case managers who are keenly aware of the reimbursement structure and monitor how long patients have stayed, the associated costs, etc. Transplant surgeons we interviewed highlighted that there is significant variation in resource use that is attributable to testing and many of these tests are, in fact, discretionary with modest expected benefit. A hospital staff also mentioned another interesting example of an action hospitals take that affects the charges for a given patient. Hospitals often contract with nearby hotels and step-down facilities and place transplant patients in the advanced stages of their recovery in them instead of keeping inpatient setting. The staff also noted that the utilization of those facilities is often discretionary and financial incentives can affect their use.

6 Results

We apply our proposed estimators to the second discontinuity point q_2 in order to estimate how the amount of health care provision depends on the reimbursement structure. In the first set of results, we apply maximum likelihood estimation to data pooled across multiple hospitals. The maximum likelihood estimation will yield $\hat{\alpha}_L, \hat{\alpha}_H, \hat{\beta}_L$ and $\hat{\beta}_H$, and these allow us to obtain the size of gap, $\hat{\phi}_{GAP}(x_t) = g_H(x_t, \hat{\beta}_H) - g_L(x_t, \hat{\beta}_L)$, and the change in the slope of the quantile function, $\hat{\phi}_{SLOPE}(x_t) = \frac{1}{f_H(0, x_t, \hat{\alpha}_H)} - \frac{1}{f_L(0, x_t, \hat{\alpha}_L)}$, at the discontinuity

point for each hospital characterized by x_t. We use exponential distribution for densities f_L and f_H with hazard rate parameter $\lambda_L(x_t, \alpha_L)$ and $\lambda_H(x_t, \alpha_H)$, respectively. All contract variables that potentially differ across hospitals, such as the locations of the first and second kinks (q_1 and q_2) and the marginal reimbursement rates (δ_1 and δ_2), are included in x_t. We also include higher-order polynomials of these variables in x_t to flexibly capture the distribution of q for multiple hospitals. To compute standard errors, we use parametric bootstrap using 500 simulations. We apply MLE to each organ type separately.

In Tables 2-4, we report maximum likelihood estimates of ϕ_{GAP} and ϕ_{SLOPE} for each hospital in the data, along with hospital characteristics.[12] We report the number of patients treated in each hospital as well. Since our maximum likelihood estimation is applied to pooled observations across hospitals within a given organ type, the relevant number of observations used in estimation is much higher than that indicated by each individual hospital. Table 2 reports estimates for BMT, Table 3 for kidney transplants, and Table 4 for liver transplants.

[Tables 2, 3 and 4 about here]

From the results in Tables 2-4, we see that $\hat{\phi}_{SLOPE}$ is positive and statistically significant for almost all cases. This suggests that for a given increase in the severity of patient health shock, hospitals tend to increase their health care spending by a larger amount when they face a positive marginal reimbursement rate than when the marginal reimbursement rate is zero. To interpret the magnitude of the coefficients, take the results for hospital 1 in Table 2. The hospital increases its bone marrow transplant spending by $252.1 for one percentile increase in patient illness severity when it is on the LHS of the kink (marginal reimbursement rate = 0%), while it increases its spending by $494.5 for one percentile increase in illness severity when it is on the RHS of the kink (marginal reimbursement rate = 50%). This amounts to approximately two times larger sensitivity

_____
[12] We know hospital names as well but we are not allowed to divulge them.

of the hospital's health care spending to BMT patients' health condition due to the hike in the reimbursement rate. Similarly, take the results for hospital 1 in Table 3.[13] The hospital increases its kidney transplant spending by $113 for one percentile increase in illness severity when it is on the LHS of the kink (marginal reimbursement rate = 0%), while it increases its spending by $515.4 for one percentile increase in illness severity when it is on the RHS of the kink (marginal reimbursement rate = 55%). This amounts to approximately four and a half times larger sensitivity of the hospital's health care spending to kidney transplant patients' health condition due to the increase in the reimbursement rate. Similar results hold for liver transplants as well.

Overall, the sensitivity of health care spending to patient illness is 2 to 13 times larger on the RHS than on the LHS for bone marrow transplants, 2 to 7 times larger on the RHS than on the LHS for kidney transplants, and 3 to 14 times larger on the RHS than on the LHS for liver transplants. What is also interesting is that $\hat{\phi}_{SLOPE}$ tends to be larger when δ_2 is larger, which again suggests that hospitals are sensitive to reimbursement rates in their health care provision decision. For instance, In Table 2, $\hat{\phi}_{SLOPE}$ is largest for Hospital 3, which has largest δ_2, and $\hat{\phi}_{SLOPE}$ is smallest for Hospital 1, which has smallest δ_2. Similar patterns hold for liver transplants (Table 4) although the picture is less clear for kidney transplants (Table 3). Due to lack of data on patient outcomes, we cannot convert these numbers to the final health outcomes of patients. Nonetheless, these results strongly suggest that financial incentives matter for hospitals' decision on resource use.

Another pattern we observe in Tables 2-4 is that $\hat{\phi}_{GAP}$ is always positive, although insignificant most of the time. The fact that $\hat{\phi}_{GAP}$ is always positive alleviates concerns about possible model misspecification. As we discussed in Section 3.2.2, an implicit assumption of the parametric functional form is that $g_L(x_t, \beta_L^0) \leq q_2(t) \leq g_H(x_t, \beta_H^0)$ for all t at the true parameters β_L^0 and β_H^0. Since we find that $g_L(x_t, \hat{\beta}_L) < g_H(x_t, \hat{\beta}_H)$ holds for all hospitals in the data, there is no evidence of model misspecification. In terms

---

[13]The kth hospital in Table 2 is different from the kth hospital in Table 3 or 4.

of sheer magnitudes, the estimates of $\hat{\phi}_{GAP}$ are quite large. For instance, the size of gap for hospital 1 in Table 2 is \$8610. But due to large standard errors, most estimates of $\hat{\phi}_{GAP}$ are statistically indistinguishable from zero. In light of these results, we mainly focus on interpretation of $\hat{\phi}_{SLOPE}$ in the remainder of this paper.

In order to test the robustness of our results, we perform our analysis at a finer level of aggregation: at the individual hospital level. In this second set of results, we apply kernel estimator as discussed in Section 3.2.1 to estimate ϕ_{SLOPE} separately for each pair of hospital and organ. This approach also allows us to use only local variation around the cutoff for identification of incentive effects. We do not estimate ϕ_{GAP}, taking our earlier results into account.[14] In our estimates, half-normal kernels with various choices of bandwidth were used to construct the weights. We use Silverman's plug-in estimates for bandwidths and also try using twice and half the plug-in estimates to test robustness. Table 5 reports kernel estimates of ϕ_{SLOPE} for each hospital and each organ type.[15]

[Table 5 about here]

The results in Table 5 are similar to our earlier results, although the magnitudes and significance differ. The fact that our global estimator (MLE) and local estimator (kernel) lead to similar conclusions is reassuring.

An overall picture that consistently appears in all these results is that hospitals tend to submit much larger bills when marginal reimbursement rates are higher. Although we cannot determine whether this is mainly due to underprovision below the threshold (necessary care is withheld) or overprovision above the threshold (unnecessary care is provided) due to lack of information on components of the final charges, the finding that hospitals are highly sensitive to financial incentives in their health care decisions is

---

[14]As the discussion in Section 3.2.1 makes clear, ignoring ϕ_{GAP} does not affect our kernel estimation of ϕ_{SLOPE}.

[15]Since estimation is done separately for each hospital, only those hospitals with sufficient numbers of observations are used in Table 5. As a result, the number of hospitals reported in Table 5 is smaller than those in Tables 2-4. In Table 5, we report the number of patients treated by each hospital.

very interesting. We also emphasize that our estimates are valid only locally around the threshold point, and do not allow us to infer the general price sensitivity of hospitals over a wide range of expenditures. This limitation of local validity is an inherent feature of RDD-type estimators.

If we had infinitely many data points, our approach outlined in Section 3.1 would suggest that we look for a "break" in the slope of the quantile function in an arbitrarily small neighborhood around θ^*. Due to the sparseness of our data, however, it is hard to tell from the raw data whether the density function has a discontinuous change at θ^*. Thus, essentially our estimates simply tell us that the average density on the RHS is smaller than the average density on the LHS within a small window around the cutoff point. Then a question that could potentially arise is whether we can interpret the observed change in the density as a result of the change in the marginal reimbursement rate. This kind of interpretational issue often arises in RDD applications since researchers frequently need to deal with small data.

To address this potential concern, we run the same type of analysis for a "control group." We have a set of hospitals whose contracts have only one kink point (they have the first kink point, but after the first kink point, the marginal reimbursement rate is always zero). We then impose an artificial cutoff point, similar in location to the cutoff point for our estimation sample of hospitals, and perform similar analysis as in Table 5. If our earlier results are an artifact of e.g. the right-skewed distribution of expenditures or something else unrelated to financial incentives, we might expect to find similar results for this control group. Estimation results for this control group of hospitals are reported in Table 6.

[Table 6 about here]

A comparison of Table 6 against Table 5 indicates that our earlier results were likely reflective of hospitals' true behavioral responses to financial incentives. In Table 5 we

saw that the estimates of ϕ_{SLOPE} were positive and statistically significant for 6 out of 7 hospitals, while the only negative estimate was not statistically significantly different from zero. In contrast, we see that 3 out of 4 hospitals have negative estimates for ϕ_{SLOPE} in our control group. Thus, we conclude that our kernel estimates of ϕ_{SLOPE} in Table 5 mostly reflect true behavioral responses of hospitals to reimbursement structures.

7 Conclusion

In this paper, we propose a modified RDD estimator that will be consistent when the forcing variable can be manipulated by agents. Our proposed estimator can be applied to many interesting settings that have been considered to be outside of the RDD framework. For instance, it is not possible to use standard RDDs to recover consumers' price sensitivity in the presence of non-linear budget constraints or workers' labor supply elasticity in the presence of higher marginal tax rates for higher income brackets, because agents optimally choose their forcing variable. Our paper shows that a modification to the standard RDDs allows us to consider these types of problems within an RDD-style framework. An application of our estimator to contracts in the health care market reveals that hospitals' health care spending is significantly influenced by financial incentives.

The assumptions required for our estimator are unlikely to hold for all settings, and thus it is important for researchers to examine whether the assumptions hold for their problems of interest. A key assumption is the strict monotonicity between the type and the dependent variable. This is likely to be violated if the type is multi-dimensional or if there is optimization error. In future work, we plan to investigate the performance of our estimator under more general conditions and improve our estimator to make it robust against these complications.

References

[1] **Arrow, Kenneth**. 1963. "Uncertainty and the Welfare Economics of Medical Care." *American Economic Review,* 53: 941-973.

[2] **Chernozhukov, Victor**. 2005. "Extremal Quantile Regression." *The Annals of Statistics*, 33(2): 806-839.

[3] **Chernozhukov, Victor and Han Hong**. 2004. "Likelihood Inference for a Class of Nonregular Econometric Models." *Econometrica*, 72(5): 1445-1480.

[4] **Cutler, David**. 1995. "The Incidence of Adverse Medical Outcomes Under Prospective Payment." *Econometrica*, 63(1): 29-50.

[5] **Dafny, Leemore**. 2005. "How Do Hospitals Respond to Price Changes?" *American Economic Review*, 95(5): 1525-1547.

[6] **Donald, Stephen and Harry Paarsch**. 1996. "Identification, Estimation, and Testing in Parametric Empirical Models of Auctions within Independent Private Values Paradigm." *Econometric Theory*, 12: 517-567.

[7] **Dranove, David and Paul Wehner**. 1994. "Physician-Induced Demand for Childbirths." *Journal of Health Economics,* 13(1): 61-73.

[8] **Fuchs, Victor**. 1978. "The Supply of Surgeons and the Demand for Operations." *Journal of Human Resources*, 13: 121-133.

[9] **Gaynor, Martin and Paul Gertler**. 1995. "Moral Hazard and Risk Spreading in Partnerships." *RAND Journal of Economics*, 26(4): 591-613.

[10] **Gaynor, Martin, James Rebitzer and Lowell Taylor**. 2004. "Physician Incentives in Health Maintenance Organizations." *Journal of Political Economy*, 112(4): 915-931.

[11] **Glied, Sherry**. 2000. "Managed Care." in *Handbook of Health Economics,* eds. A. Cuyler and J. Newhouse, North-Holland, 707-753.

[12] **Gruber, Jonathan and Maria Owings**. 1996. "Physician Financial Incentives and Cesarean Section Delivery." *Rand Journal of Economics*, 27(1): 99-123.

[13] **Hahn, Jinyong, Petra Todd, and Wilbert Van der Klaauw**. 2001. "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design." *Econometrica*, 69(1): 201-209.

[14] **Ho, Katherine**. 2009. "Insurer-Provider Networks in the Medical Care Market." *American Economic Review*, 99(1): 393–430.

[15] **Hodgkin, Dominic and Thomas McGuire**. 1994 "Payment Levels and Hospital Response to Prospective Payment." *Journal of Health Economics*, 13: 1-29.

[16] **Imbens, Guido and Thomas Lemieux**. 2008. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics*, 142(2): 615-635.

[17] **Ketcham, Jonathan, Pierre Léger and Claudio Lucarelli**. 2011 "Standardization Under Group Incentives." Working Paper.

[18] **Koenker, Roger and Gilbert Bassett**. 1978. "Regression Quantiles." *Econometrica*, 46: 33-50.

[19] **Lee, David and Thomas Lemieux**. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature*, 48(2): 281-355.

[20] **McCrary, Justin**. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics*, 142(2): 698-714.

[21] **McGuire, Thomas**. 2000. "Physician Agency." in *Handbook of Health Economics,* eds. A. Cuyler and J. Newhouse, North-Holland, 467-536.

[22] **Milgrom, Paul and Chris Shannon**. 1994. "Monotone Comparative Statics." *Econometrica*, 62(1): 157-180.

[23] **Rosenthal, Meredith. 2004**. "donut-Hole Economics." *Health Affairs*, 23(6): 129-135.

[24] **Topkis, Donald**. 1978. "Minimizing a Submodular Function on a Lattice." *Operations Research*, 26: 305-321.

[25] **Urquiola, Miguel and Eric Verhoogen**. 2009. "Class-Size Caps, Sorting, and the Regression-Discontinuity Design." *American Economic Review*, 99(1): 179-215.

r(q)

q₁ q₂ q
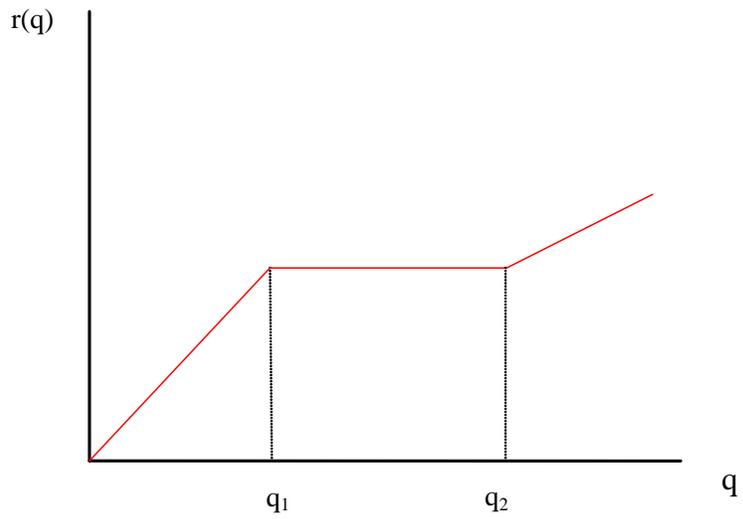
Figure 1: A Typical Reimbursement Scheme

Marginal
Benefit (θ)

Marginal
Benefit (θ* > θ)

Gap

**Net marginal
cost**

Bunching

q_L q_H

q₁

q₂

q
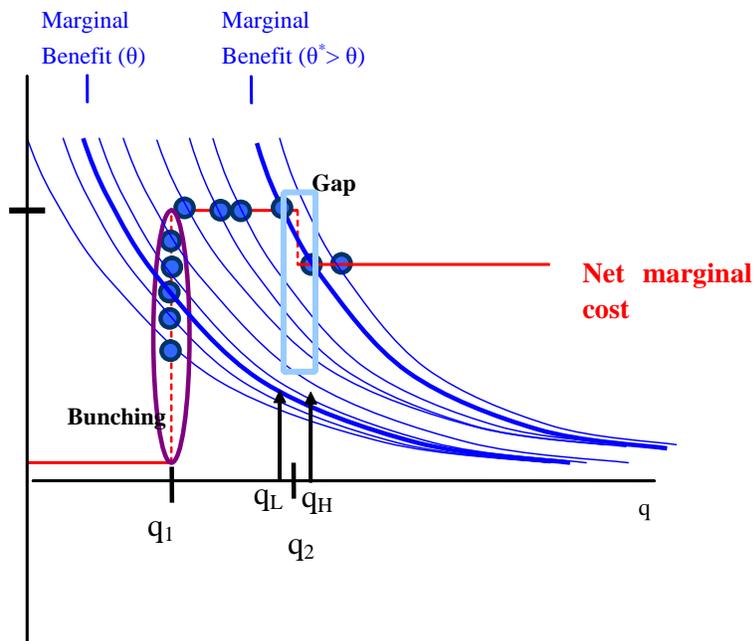
Figure 2: Optimal Decision Rule for Low γ^a Agent
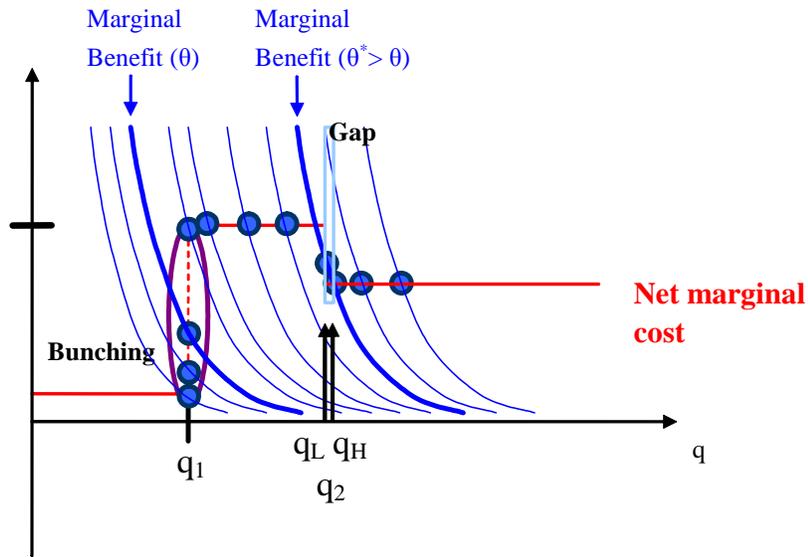
Figure 3: Optimal Decision Rule for High γ^a Agent
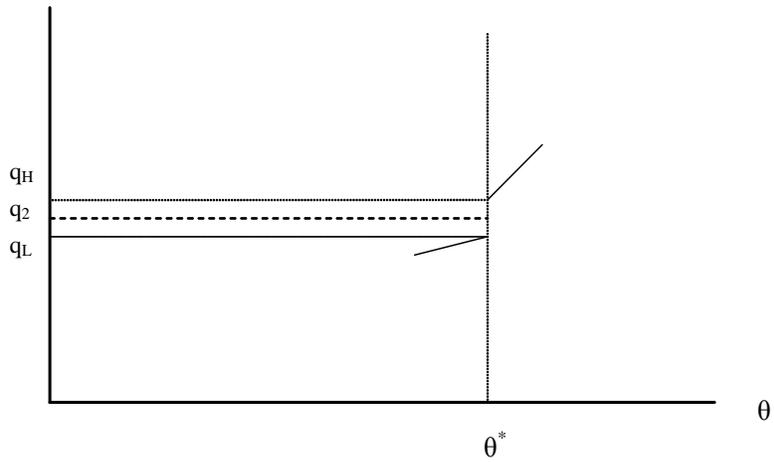
Figure 4: The slope of quantile function $q(\theta)$ changes at θ^*

Table 1: Summary Statistics

	BMT	Kidney	Liver
Total # Patients	511	265	344
Total # Hospitals	7	6	7
Avg. Charge per Patient (in $1000)	168.56 (114.87)	140.74 (84)	319.63 (246.4)
Avg. Reimbursement per Patient (in $1000)	100.27 (67.98)	76.5 (42.12)	196.12 (145.61)
Avg. # Patients per Hospital	73 (36.82)	44.17 (11.57)	49.14 (30.39)
Avg. q_1 across Hospitals (in $1000)	121.97 (13.65)	80.09 (9.42)	181.57 (17.34)
Avg. q_2 across Hospitals (in $1000)	163.10 (22.44)	128.24 (12.63)	248.37 (28.57)
Avg. δ_1 across Hospitals	0.75 (0.08)	0.82 (0.08)	0.79 (0.07)
Avg. δ_2 across Hospitals	0.56 (0.06)	0.51 (0.05)	0.58 (0.06)
Avg. % Patients with $q < q_1$	37.39 (16.22)	8.92 (5.62)	10.33 (6.58)
Avg. % Patients with $q_1 \leq q \leq q_2$	24.76 (11.45)	42.20 (15.76)	33.98 (15.55)
Avg. % Patients with $q > q_2$	37.85 (20.33)	48.88 (11.02)	55.69 (12.21)

Inside the parentheses are standard deviations.

Table 2: Maximum Likelihood Estimates, Bone Marrow Transplant (q measured in $1,000)

	Obs	q_1	q_2	δ_1	δ_2	$\hat{\phi}_{GAP}$	Slope L	Slope R	$\hat{\phi}_{SLOPE}$
H1	28	135.71	190	0.7	0.5	8.61 (78.46)	25.21	49.45	24.24 (14.93)
H2	73	133.33	181.82	0.75	0.55	5.13 (4.05)	22.40	58.00	35.59 (10.38) ***
H3	74	107.33	123.85	0.75	0.65	4.70 (8.42)	11.28	146.29	135.01 (21.35) ***
H4	148	120	150	0.75	0.6	7.10 (2.63) **	15.42	95.99	80.57 (7.46) ***
H5	64	126.67	172.73	0.75	0.55	35.22 (16.23) *	20.31	66.21	45.91 (7.45) ***
H6	71	130.77	170	0.65	0.5	3.44 (3.62)	20.31	66.19	45.87 (14.04) ***
H7	53	100	153.33	0.92	0.6	10.27 (16.63)	15.98	91.44	75.46 (7.71) ***

Inside the parentheses are bootstrapped standard errors.

* Significant at 10% ** Significant at 5% *** Significant at 1%

Table 3: Maximum Likelihood Estimates, Kidney Transplant (q measured in $1,000)

	Obs	q_1	q_2	δ_1	δ_2	$\hat{\phi}_{GAP}$	Slope L	Slope R	$\hat{\phi}_{SLOPE}$
H1	29	97.14	123.64	0.7	0.55	0.78 (6.96)	11.30	51.53	40.23 (5.27) ***
H2	37	74.67	124.44	0.75	0.45	6.69 (25.01)	14.25	74.07	59.82 (24.72) **
H3	48	75.29	130.61	0.85	0.49	4.67 (3.98)	14.66	57.15	42.49 (14.41) ***
H4	47	79.41	137.78	0.85	0.49	3.80 (4.13)	16.75	50.41	33.65 (6.37) ***
H5	41	83.38	144.64	0.85	0.49	1.95 (4.97)	19.05	44.66	25.60 (8.13) ***
H6	63	70.65	108.33	0.92	0.6	1.44 (3.32)	7.62	55.86	48.24 (8.76) ***

Inside the parentheses are bootstrapped standard errors.

* Significant at 10% ** Significant at 5% *** Significant at 1%

Table 4: Maximum Likelihood Estimates, Liver Transplant (q measured in $1,000)

	Obs	q_1	q_2	δ_1	δ_2	$\hat{\phi}_{GAP}$	Slope L	Slope R	$\hat{\phi}_{SLOPE}$
H1	95	178.53	206	0.75	0.65	2.39 (5.76)	13.05	181.72	168.67 (22.39) ***
H2	31	166.47	288.78	0.85	0.49	13.99 (17.21)	42.14	141.11	98.97 (25.58) ***
H3	23	160	218.18	0.75	0.55	10.01 (258.73)	14.77	155.15	140.37 (125.67)
H4	38	200	254.55	0.7	0.55	7.77 (12.08)	25.85	155.15	129.30 (17.6) ***
H5	42	198.8	271.09	0.75	0.55	26.02 (13.51) *	33.34	155.15	121.81 (16.72) ***
H6	90	198.72	250	0.78	0.62	6.54 (5.73)	25.19	173.30	148.11 (13.45) ***
H7	25	168.48	250	0.92	0.62	33.30 (29.81)	25.19	173.30	148.11 (13.45) ***

Inside the parentheses are bootstrapped standard errors.

* Significant at 10% ** Significant at 5% *** Significant at 1%

Table 5: Kernel Estimates (q measured in \$1,000)

	Obs	Bandwidth	Slope L	Slope R	$\hat{\phi}_{SLOPE}$
H2 (BMT)	73	Silverman's Plug-In	34.70 (0.91)	51.87 (1.69)	17.17 (2.6) ***
H4 (BMT)	148	Silverman's Plug-In	21.36 (0.48)	186.29 (9.01)	164.93 (9.48) ***
H3 (Kidney)	48	Silverman's Plug-In	46.90 (3.59)	71.6 (2.03)	24.7 (5.63) ***
H5 (Kidney)	41	Silverman's Plug-In	34.08 (0.79)	72.53 (5.7)	38.45 (6.48) ***
H6 (Kidney)	63	Silverman's Plug-In	16.35 (0.26)	88.06 (2.75)	71.71 (3.0) ***
H2 (Liver)	31	Silverman's Plug-In	130.71 (50.71)	109.98 (10.95)	-20.73 (61.67)
H6 (Liver)	90	Silverman's Plug-In	54.12 (4.81)	140.87 (5.88)	86.75 (10.7) ***
H2 (BMT)	73	$1/2 \times$ Silverman's	37.04 (2.21)	42.72 (1.88)	5.67 (4.09)
H4 (BMT)	148	$1/2 \times$ Silverman's	19.61 (0.74)	211.55 (26.38)	191.93 (27.11) ***
H3 (Kidney)	48	$1/2 \times$ Silverman's	48.39 (7.89)	49.22 (1.32)	0.83 (9.21)
H5 (Kidney)	41	$1/2 \times$ Silverman's	29.54 (1.02)	78.6 (14.5)	49.06 (15.52) ***
H6 (Kidney)	63	$1/2 \times$ Silverman's	17.13 (0.59)	68.5 (2.59)	51.37 (3.18) ***
H2 (Liver)	31	$1/2 \times$ Silverman's	323.97 (1544)	80.65 (8.64)	-243.32 (1553)
H6 (Liver)	90	$1/2 \times$ Silverman's	57.35 (11.46)	109.71 (5.56)	52.36 (17.02) ***
H2 (BMT)	73	$2 \times$ Silverman's	39.54 (0.67)	71.25 (2.19)	31.7 (2.86) ***
H4 (BMT)	148	$2 \times$ Silverman's	27.21 (0.49)	200.7 (5.63)	173.5 (6.12) ***
H3 (Kidney)	48	$2 \times$ Silverman's	43.22 (1.41)	116.31 (4.36)	73.09 (5.76) ***
H5 (Kidney)	41	$2 \times$ Silverman's	44.5 (0.87)	87.1 (4.93)	42.6 (5.81) ***
H6 (Kidney)	63	$2 \times$ Silverman's	17.5 (0.16)	125.74 (4)	108.24 (4.16) ***
H2 (Liver)	31	$2 \times$ Silverman's	113.66 (16.67)	170.64 (20.47)	56.99 (37.14)
H6 (Liver)	90	$2 \times$ Silverman's	54.81 (2.5)	211.4 (9.94)	156.59 (12.44) ***

Inside the parentheses are standard errors.

* Significant at 10%   ** Significant at 5%   *** Significant at 1%

Table 6: Kernel Estimates, "Control" Group (q measured in $1,000)

	Obs	Bandwidth	Slope L	Slope R	$\hat{\phi}_{SLOPE}$
H1 (BMT)	30	Silverman's Plug-In	65 (6.55)	58.91 (3.27)	-6.09 (9.81)
H1 (Kidney)	36	Silverman's Plug-In	97.17 (32.18)	26.68 (1.12)	-70.49 (33.3) **
H2 (Kidney)	46	Silverman's Plug-In	27.92 (0.98)	69.59 (3.56)	41.67 (4.54) ***
H1 (Liver)	60	Silverman's Plug-In	770.889 (5452)	422.43 (166.13)	-348.46 (5618)
H1 (BMT)	30	$1/2 \times$ Silverman's	73.62 (19.02)	43.34 (2.6)	-30.29 (21.62)
H1 (Kidney)	36	$1/2 \times$ Silverman's	295.37 (1807)	27.94 (2.57)	-267.43 (1810)
H2 (Kidney)	46	$1/2 \times$ Silverman's	29.62 (2.33)	55.24 (3.56)	25.62 (5.89) ***
H1 (BMT)	30	$2 \times$ Silverman's	72.51 (4.54)	83.4 (4.63)	10.89 (9.18)
H1 (Kidney)	36	$2 \times$ Silverman's	55.78 (3.04)	37.16 (1.51)	-18.62 (4.56) ***
H2 (Kidney)	46	$2 \times$ Silverman's	28.62 (0.53)	102 (5.61)	73.38 (6.14) ***
H1 (Liver)	60	$2 \times$ Silverman's	290.81 (146.34)	581.06 (216.18)	290.25 (362.52)

Inside the parentheses are standard errors.

* Significant at 10% ** Significant at 5% *** Significant at 1%