

# Information Disclosure and Firm Performance: Evidence from the Dialysis Industry

Subramaniam Ramanarayanan  
Anderson School of Management, UCLA  
Los Angeles, CA 90095  
[subbu@anderson.ucla.edu](mailto:subbu@anderson.ucla.edu)

Jason Snyder  
Anderson School of Management, UCLA  
Los Angeles, CA 90095  
[jason.snyder@anderson.ucla.edu](mailto:jason.snyder@anderson.ucla.edu)

September 2012

## Abstract:

We study the impact of information disclosure policies on firm performance by exploiting a policy change that provides plausibly exogenous “shocks” to firms’ reputations based on their allocation to coarse performance categories. Medicare grades dialysis firms using three coarse performance categories based on patient survival rates: *worse than expected*, *as expected*, and *better than expected*. We exploit the underlying continuous performance measures used to create these categories to implement a regression discontinuity design. We find firms that just barely fall into the *worse than expected* category subsequently experience a reduction in patient mortality rates. We provide suggestive evidence that this improvement is driven largely by strategic patient selection. There is no impact of ratings on overall patient volumes, but facilities receiving poor grades treat fewer well-informed patients post-disclosure. We do not find comparable supply-side or demand-side effects for firms that just barely fall into the *better than expected* category. The overall evidence is consistent with disappointing information being a significant motivator of firm behavior.

---

We are grateful for helpful comments by Simon Board, Vanessa Burbano, Christopher Carpenter, Barton Hamilton, Thomas Koch, Phillip Leslie, Steve Lippman, Moritz Meyer-Ter-Vehn, Lamar Pierce, Ashley Swanson, and seminar and conference participants at the American Society of Health Economists Meeting, International Industrial Organization Conference, NBER HealthCare program meeting, University of Texas Austin, UCLA, UC Irvine, UC Berkeley and the Olin School of Business at Washington University at St. Louis. We are thankful to Scott Klein and Jennifer LaFleur from ProPublica for their help with accessing the Dialysis Facility Compare data and to Leemore Dafny and Chris Ody for sharing the Medicare Cost Reports data with us. Nicholas Ross provided outstanding research assistance. Special thanks to Al Roth’s Market Design blog for making us aware of the issue and the data.

## I. Introduction

Over the past 50 years, an extensive literature in economics has highlighted the role of information in decision making. A central insight from this work is that improving the transparency of information usually leads to desirable social outcomes. This finding has spurred the development of disclosure programs aimed at increasing the transparency of firm and product quality in numerous settings. Prior research on the impact of disclosure programs has found that, all else equal, consumers with more information sort towards higher-quality firms and, less conclusively, that information disclosure causes firms to invest in quality improvements.<sup>1</sup> However, relatively little attention has been paid to the impact that the framing of the disclosed information has on firm behavior and consumer choice. In practice, this is a crucial input into the design of quality disclosure programs, given their focus on identifying sellers at both extremes of the quality distribution and on eliciting a strong response.<sup>2</sup> There is little causal evidence in the literature that compares the relative effectiveness of positive and negative information disclosure.<sup>3</sup>

In this paper, we study firm and consumer responses to positive and negative information disclosure in the dialysis industry.<sup>4</sup> Quality ratings of dialysis facilities are assigned by the Centers for Medicare and Medicaid Services (CMS) and are designed to expose the best- and worst-performing facilities. Since 2001, CMS has made available, through the Dialysis Facility Compare (DFC) tool available on its website, quality information on dialysis facilities in the form of three coarse categories based on a facility's

---

<sup>1</sup> For example, Jin and Sorenson (2006), Dafny and Dranove (2008), and Hastings and Weinstein (2008) all find evidence of consumer sorting on quality and Jin and Leslie (2003) find evidence of disclosure leading to improvements in product quality. See Dranove and Jin (2010) for a comprehensive overview.

<sup>2</sup>As an illustrative example of the type of program that targets the extremes of the quality distribution, consider the Blue Ribbon Schools Program administered by the Department of Education, which has recognized approximately 6,000 truly exceptional schools out of over 100,000 eligible schools. In contrast, the Environmental Working Group recently attempted to shame cereal makers responsible for the so-called “cookie breakfast” by compiling a list of the “Ten Most Sugary Cereals.”

<sup>3</sup> Presenting this paper in academic seminars, we have often asked the audience: “Would you rather have a program that publically discloses the names of faculty in the top 10 percent of teachers or the bottom 10 percent?”

<sup>4</sup> Dialysis is one of the primary means of treating kidney failure. Patients on dialysis have mortality rates as high as 20 percent in the first year of treatment. Additionally, there is considerable dispersion in patient survival rates across facilities (Fields 2010).

continuous risk-adjusted patient survival rate (averaged over the previous four years): *worse than expected*, *as expected*, and *better than expected*.

Estimating the causal impact of a disclosure program designed to highlight the firms at the tails of the quality distribution is challenging. Beyond endogeneity and omitted-variables-bias concerns, mean reversion undermines most attempts to study such programs. For example, if a firm receives a negative rating (resulting from either poor underlying quality or bad luck),<sup>5</sup> any subsequent improvement in its performance could simply be driven by mean reversion, as opposed to being attributable to disclosure.<sup>6</sup>

To overcome these empirical challenges, we use a regression discontinuity design that exploits the fact that the coarse rankings were discontinuously assigned. Our identification strategy relies on two unique institutional features. First, facilities with similar expected patient survival rates could receive drastically different quality ratings due to the discontinuous assignment process.<sup>7</sup> Second, in 2008, CMS updated the statistical method used to classify facilities into performance categories. This revision, intended to better identify the best and worst performers, resulted in the percentage of facilities being classified as performing *as expected* dropping from 96 percent in previous years to 80 percent in 2008.<sup>8</sup> Correspondingly, the number of facilities rated superior or inferior increased sharply. This change provides the statistical power to implement the regression discontinuity design. The underlying continuous scores that are used to construct the categorical rankings were made available to the public only in 2010. We observe the continuous scores and the coarse ratings after the fact, but at the time covered by our study, only the coarse ratings were visible to the market participants (both dialysis facilities and patients).

Our findings from the regression discontinuity design show that a facility receiving a *worse than expected* performance grade in 2008 (that is, based on patient survival data for 2004-2007) experiences an improvement in patient survival compared to a facility that is

---

<sup>5</sup> This is especially important when there is a random component to the measurement of product quality measurement.

<sup>6</sup> The initial negative rating might have been due to an atypical negative shock to product quality.

<sup>7</sup> For example, in 2008, a firm in the bottom 11th percentile of patient survival rates would receive an *as expected* rating while a firm in the bottom 10th percentile would receive a *worse than expected* rating.

<sup>8</sup> This information was released on the DFC website towards the end of 2008. In Section IV.B, we expand on the timeline of events relevant to our analysis.

graded *as expected*. The magnitude of this improvement is economically meaningful, translating into a nearly 15-point improvement in the national risk-adjusted mortality ranking for the average facility rated *worse than expected*. Given an unconditional average death rate of 25 patients per 100 patient-years at such a facility, the quality disclosure would result in six fewer patient deaths per 100 patient-years on average. In contrast, we find that distinguishing between facilities that performed *better than expected* and *as expected* has little impact on patient survival rates. Further analysis suggests that the performance improvement of *worse than expected* facilities is largely a matter of strategic patient selection, with the greatest degree of improvement being registered among facilities prone to engaging in cherry-picking healthier patients. The improvement does not vary according to a facility's ownership status—for-profit versus not-for-profit—but we find evidence suggesting that performance improvement is more likely in facilities operating in more competitive settings. It is also found to be persistent only amongst firms with the strongest incentives to improve.<sup>9</sup>

We document demand-side responses to disclosure as well; facilities rated *worse than expected* subsequently treat fewer well-informed patients. Patients with prior dialysis experience are less likely to transfer into these facilities and patients that have seen a kidney specialist prior to dialysis are less likely to begin treatment at these facilities. All told, we find strong causal evidence that disclosure policies aimed at the extremes of the quality distribution impact both firm performance and consumer choice, but this impact is heavily skewed towards influencing the worst performers.

This study makes a number of contributions with important policy implications beyond the dialysis industry. First, we make use of a regression discontinuity design that allows for sharp identification of the causal impact of information disclosure while mitigating concerns about mean reversion.<sup>10</sup> Second, the setting of the study enables us to examine the effect of disclosure programs that identify firms on both ends of the quality distribution. Our

---

<sup>9</sup> We identify such firms as those that had better performance in 2004 than in any other year in our 2004-2007 timeframe. Given that the performance measure for 2009 is computed using patient survival rates averaged across 2005-2008 (and excludes 2004), such firms will experience a mechanical decline in performance, conditional on their performance in 2008, when compared to firms for which 2004 was not the best year. We surmise that such firms face relatively stronger incentives to improve performance as a result.

<sup>10</sup> Luca (2011) uses a similar methodology to estimate the impact of consumer reviews on restaurant revenue. However, he does not focus on the extremes of the quality distribution.

results thus shed light on the optimal design of disclosure programs, suggesting that negative information is stronger than positive information as a motivator of firm behavior and consumer choice.

## II. Information Disclosure and Its Impact on Consumers and Sellers

This study builds on an extensive body of research on the impact of quality disclosure programs, such as rankings and report cards, across a variety of settings, including healthcare. We focus on two of the primary findings most relevant to this study. First, at the consumer level, there is substantial evidence that quality disclosure leads to vertical sorting. In other words, consumers are more likely—all else being equal—to choose higher-quality products after disclosure. Hastings and Weinstein (2008) provide evidence of such sorting among parents choosing public schools for their children. Jin and Sorenson (2006), Dafny and Dranove (2008), and Scanlon et al. (2002) show similar findings among enrollees choosing health plans. Consumer response to information disclosure has been shown to be moderated by a variety of factors. For example, Wedig and Tai-Seale (2002) find a strong response to quality disclosure only among first-time enrollees in health plans.

Recent studies have also pointed to the role of consumer attention in determining the overall response to information about quality. To the extent that consumers do not take all available information into account when they make decisions, the salience of the information presented becomes a driver of consumption choices. For example, DellaVigna and Pollet (2009) find that earnings announcements made on Fridays lead to much weaker stock price responses—due to consumer inattention—while Pope (2006) finds that consumers respond to changes in *U.S. News & World Report's* hospital rankings, even when the underlying continuous quality is controlled for. Further, consumer response has been shown to be asymmetric to the information disclosed, with negative information being a much stronger driver of demand. In a study focusing on the introduction of hospital report cards in New York State in the early 1990s, Dranove and Sfekas (2008) find that hospitals that received

negative news experienced a drop in demand, but positive reports brought no benefits. Scanlon et al. (2002) find a similar asymmetric consumer response to health plan ratings. Both studies imply that choosing an appropriate design for quality disclosure programs is central to achieving desirable social outcomes with respect to consumer choice, yet neither study uses a research design that addresses the possibility of mean reversion.

The second major finding from the literature on the impact of quality disclosure programs is that disclosure leads sellers to invest in product quality improvements. For example, Jin and Leslie (2003) find a significant decline in hospital admissions for food-borne illnesses after the adoption of restaurant hygiene grade cards in Los Angeles County. Analogous to the variation in consumer response to quality, seller response has also been shown to be heterogeneous with respect to dimensions such as the competitiveness of the local market. For example, Chen (2008) finds that nursing homes in more competitive markets experienced larger quality improvements after disclosure. But there are other, less admirable ways to improve performance. Dranove et al. (2003) show that, upon the introduction of physician report cards in Pennsylvania, doctors were more likely to cherry-pick the healthiest patients.<sup>11</sup> Forbes, Lederman and Tombe (2012) examine the disclosure of on-time performance by airlines and find significant evidence of gaming of the disclosure thresholds through misreporting, especially among airlines that report delays manually. This gaming behavior is also found to be more prevalent among airlines that provide incentives to employees tied to on-time performance.

While the asymmetric effect of disclosure programs has been investigated from a demand perspective, the literature is silent on whether good and bad ratings are equally or differentially effective at evoking the desired seller response. On the one hand, tournament theory (Lazear and Rosen 1981) suggests that disclosing the identity of the very best firms can create a powerful incentive for all firms to improve product quality. Conversely, the behavioral economics literature (Kahneman and Tversky 1979) provides compelling evidence suggesting that evaluations that fall short of expectations can be a markedly more powerful

---

<sup>11</sup> Kolstad (2012) shows that disclosure can also lead to intrinsically motivated improvements among surgeons.

motivator than evaluations that exceed expectations.<sup>12</sup> For example, Mas (2006) finds that arrest rates fall when police officers receive pay raises that fall short of their expectations, but the corresponding effect is much smaller when they receive raises that exceed expectations. This suggests that singling out the worst performers can provide a meaningful way to encourage improvements in product quality. The setting of our study enables us to disentangle the impact of positive and negative disclosure on the behavior of both firms and consumers.

### III. Institutional Setting and Data

#### *A. End-Stage Renal Disease*

End-stage renal disease (ESRD) refers to a stage of chronic kidney disease when the kidneys completely fail in their function of removing waste from the body.<sup>13</sup> At this point, the only treatments available are dialysis and organ transplantation.<sup>14</sup> Most physicians prefer transplantation primarily because ESRD patients that undergo an organ transplant live longer and healthier lives in comparison to those treated with dialysis. However, there are not nearly enough healthy kidneys available for the number of ESRD patients.<sup>15</sup> As a result, nearly 70 percent of ESRD patients in the United States (approximately 400,000 patients) undergo dialysis every year.<sup>16</sup>

In 1972, the Social Security Act extended Medicare Part A and Part B benefits to people with ESRD regardless of age (Nissenson and Rettig 1999). This entitlement currently covers over 90 percent of all ESRD patients in the U.S. Medicare covers both inpatient and outpatient dialysis (under Parts A and B, respectively) and typically pays 80 percent of the

---

<sup>12</sup> Our paper is suggestive of—but not a direct test of—prospect theory. We focus on disclosure at the extremes of the quality distribution while prospect theory posits a kink in the utility function at a reference point.

<sup>13</sup> Common causes of kidney failure are diabetes and hypertension.

<sup>14</sup> These treatments are not exclusive. Many patients undergo dialysis while on the waitlist for a kidney transplant.

<sup>15</sup> According to the National Kidney Foundation, over 80,000 patients were on the waitlist to receive a kidney transplant in the U.S. in 2009, but only 16,500 kidney transplants were performed in 2008.

<sup>16</sup> <http://kidney.niddk.nih.gov/kudiseases/pubs/kustats/>.

approved amount. Patients pay the rest, either out-of-pocket or through supplemental insurance policies such as Medicaid or Medigap.

Since the program's inception, Medicare has reimbursed dialysis facilities a fixed fee for each treatment. A base rate is intended to cover the entire bundle of services, tests, and certain drugs for up to three dialysis sessions per week.<sup>17</sup> This base rate is then adjusted to account for differences in patient case-mix across facilities based on the patient age's and body mass index.<sup>18</sup> Finally, differences in local input prices—that is, wages—across facilities are also incorporated into the final payment.<sup>19</sup> An important implication of this near-universal coverage is that, with prices mostly fixed, dialysis facilities compete on service quality and location. ESRD accounted for approximately six percent of the Medicare budget in 2010 (Fields 2010). Given that Medicare expenses accounted for approximately 13 percent of the 2010 federal budget, *ESRD expenditures made up almost one percent of the entire federal budget in 2010.*<sup>20</sup>

## ***B. The Dialysis Industry***

Dialysis is designed to replicate the cleaning function that was performed by the kidneys before they failed. It helps ESRD patients live longer but is not intended as a permanent cure for kidney failure. There are two major categories of dialysis, based on the approach to removing waste from the bloodstream. Hemodialysis uses a special membrane to filter the blood and is usually performed at a dialysis facility. Peritoneal dialysis uses the peritoneum, the lining of the abdominal cavity, to filter the blood and is usually performed at the patient's residence. Patients may choose to switch from one mode to the other as their treatment progresses.

---

<sup>17</sup> For 2012, the base rate is \$234 for freestanding and hospital-based facilities.

<sup>18</sup> Starting in 2011, the patient-level adjustment will also account for six other comorbidities, according to the Medicare Payment Advisory Commission (<http://www.medpac.gov>).

<sup>19</sup> Starting in 2011, Medicare phased in a new Prospective Payment System, which bundles together all dialysis services and items that were previously billed separately. This change occurs outside the timeframe of our data and hence does not affect our analyses.

<sup>20</sup> Expenditures incurred by patients with a diagnosis of kidney disease made up 31 percent of Medicare expenditures in 2009 (<http://www.usdrs.org>).

The vast majority of dialysis patients in the U.S. are treated in-unit at one of approximately 5,000 dialysis centers three times a week, with each treatment lasting three to five hours. Over 90 percent of these centers are freestanding facilities and, in addition to dialysis services, may provide lab testing and drug infusion services. A typical center provides approximately 50 treatments a day using 15-20 dialysis stations. Each center is required to have a medical director who must be board-certified in internal medicine or pediatrics and have experience in dealing with ESRD patients (Lawler et al. 2003). In addition, the Centers for Medicare and Medicaid Services (CMS) mandates the presence of at least one licensed registered nurse, a social worker, and a dietitian. Centers may employ additional patient care technicians, but at least one licensed healthcare provider (such as a doctor or registered nurse) must be present when a patient undergoes dialysis. Staffing ratios vary by state and few states have regulations regarding these numbers (Wolfe 2011).<sup>21</sup> Mortality rates on dialysis are grim. Approximately 20 percent of patients die within their first year on dialysis and 65 percent die within five years.<sup>22</sup>

The market structure of the dialysis industry has changed dramatically over the last decade. While the number of facilities has grown from around 2,000 in 1991 to over 5,000 in 2009, the industry has also become increasingly concentrated; the two largest dialysis providers, DaVita and Fresenius, together accounted for over 60 percent of market share in 2009 (United States Renal Data System 2011).<sup>23</sup> Nearly 80 percent of dialysis facilities are designated as being for-profit (Fields 2010).

---

<sup>21</sup> As an example, Georgia mandates a staffing ratio of registered nurses to dialysis patients of 1:10 while Texas requires a ratio of 1:12. The National Kidney Foundation releases recommended staffing ratios in terms of the number of dietitians and social workers per patient undergoing dialysis (1:100 and 1:75, respectively). See Wolfe (2011) for more details.

<sup>22</sup> Dialysis patients have mortality rates similar to those of patients with stage III colon cancer.

<sup>23</sup> Independently owned facilities accounted for 15 percent, with hospital-based facilities and some smaller chains accounting for the rest.

### *C. The Dialysis Facility Compare Data*

There is considerable dispersion in patient survival rates, as measured by the standardized mortality ratio (SMR) across dialysis facilities. The SMR compares the observed mortality rate in a particular facility to the death rate that would be expected based on national death rates for patients with characteristics similar to those treated at the facility. The SMR is adjusted for patient demographics such as age, sex, race, and ethnicity; for comorbidities such as diabetes, body mass index (BMI), and duration of ESRD; and for regional variables such as overall state death rates.

**Figure 1** shows the distribution of SMRs for dialysis facilities in the U.S. using data from the years 2004-2007 (we discuss the data source in detail below). The wide dispersion in patient survival rates referenced earlier can be readily seen here. In the bottom 10th percentile of the distribution, a facility has a 30-percent-lower-than-expected mortality rate. In the top 90th percentile, a facility has a 33-percent-higher-than-expected mortality rate.<sup>24</sup> Prior research has investigated some of the determinants of these differences. Garg et al. (1999) find that for-profit ownership is associated with higher mortality rates. Powe et al. (2002) report, based on interviews with dialysis administrators, that patient education, staffing ratios, and wage levels are crucial determinants of facility quality. In a cross-sectional study of 90 facilities, Spiegel et al. (2010) find that those categorized as performing *better than expected* were associated with more engaged patients and with better communication and coordination between physicians and staff.

In response to this variation in quality, CMS developed and released the Dialysis Facility Compare (DFC) tool in 2000. The primary impetus for this program came from the Balanced Budget Act of 1997, which required CMS to “develop and implement (by not later than January 1, 2000) a method to measure and report the quality of renal dialysis services provided under the Medicare program” (Frederick et al. 2002). DFC was first introduced on the [www.medicare.gov](http://www.medicare.gov) website on January 19, 2001 and provided consumers with information on the location, hours, and quality (as measured by the SMR) of almost all of the nation’s dialysis facilities.

---

<sup>24</sup> Because approximately 20 percent of patients die each year, these differences are large.

At first, CMS decided—based on recommendations from a Consumer Information Workgroup—not to report a facility’s actual SMR. Instead, it put each facility into one of three coarse performance categories—*as expected*, *better than expected*, and *worse than expected*—based on where that facility’s SMR fell in the nationwide distribution.<sup>25</sup> The data used for this categorization come from the National Kidney Foundation, which uses a four-year estimation window in order to calculate facility SMRs. For example, performance categories reported by the DFC in 2008 are based on facilities’ patient survival rates (SMRs) for 2004-2007. CMS uploads the new performance categories to the DFC website each November.<sup>26</sup>

In December 2010, ProPublica, an independent nonprofit investigative news outlet, made available to the public on its website the precise underlying mortality data for all four-year windows between 2002 and 2009.<sup>27</sup> Robin Fields, an investigative reporter and senior editor with ProPublica, obtained this information by filing multiple Freedom of Information Act requests over the course of two years. ProPublica posted this information on its website out of concern that the coarse SMR categories reported on the CMS website did not enable patients to adequately compare facilities on the basis of quality.<sup>28</sup>

Using the data uploaded on ProPublica.org, we constructed a dataset that includes facility-level SMRs (both the continuous scores and the coarse categories reported by CMS) and a number of facility characteristics for the years 2004-2009. In order to exploit the

---

<sup>25</sup> This workgroup included representatives of physicians, nurses, patients, social workers, and facility administrators (Frederick et al. 2002).

<sup>26</sup> In section IV (B), we provide more information on the exact timeline of events relevant to the empirical analysis.

<sup>27</sup> Prior to that year, the National Kidney Foundation was extremely reluctant to disclose the underlying risk-adjusted mortality data and consumers were privy only to the information conveyed by the coarse performance categories.

<sup>28</sup>Fields (2010) describes the inadequacy of the DFC ratings using an example: “Innovative Renal Care and Midtown Kidney Center, clinics about two miles apart in Houston, had similar stats on Dialysis Facility Compare in 2007, including “as expected” survival rates. But the full data show that Innovative Renal’s average annual death rate—after factoring in patient demographics and complicating conditions—was 34 percent higher than expected. Midtown’s average rate was 15 percent lower than expected. Dialysis Facility Compare has since changed Innovative’s survival rating to “worse than expected,” but how much worse? The unpublished 2009 data reveal that the clinic performed more poorly, versus expectations, than 92 percent of all facilities nationwide.”

policy change instituted by CMS, we use the time period 2008-2009 for our analysis.<sup>29</sup> Given that mortality rates are computed using four-year windows, the analysis sample contains information on dialysis patient survival rates starting with the 2004-2007 timeframe. The sample includes information on all 4,665 firms that received performance evaluations on the DFC website in December of 2008. For each of these firms, we obtained information on its past and future patient survival rates, organizational form, ownership, and patient characteristics on a yearly basis from 2004 through 2009. All of our analyses are performed at the facility-year level.

The top panel of **Table 1** presents summary statistics of organizational characteristics for the 4500+ dialysis facilities in our sample in 2007. Eighty percent of the facilities are associated with for-profit ownership and are affiliated with a national or regional chain. Each facility treats over 110 patients and has, on average, around 18 dialysis stations. For two reasons, there is some discrepancy in the number of observations across variables. First, 153 of the 4,665 facilities reported on the DFC website in December 2008 closed in 2009. This could introduce survivorship bias into our results, a concern we address in Section V. Second, some of the reported variables are not uniformly recorded across all firms. These missing observations are a relatively small portion of the overall data and do not appear to have any systematic correlation with firm characteristics or performance.

#### **IV. Estimating the Causal Impact of Information Disclosure on Performance**

Establishing the causal impact of quality disclosure on firm performance is difficult. In the absence of a natural experiment, the prevalent approach in the literature is a differences-in-differences design to assess outcomes before and after the implementation of the disclosure

---

<sup>29</sup> These data are available on the ProPublica website in the form of reports (in Adobe Acrobat PDF format) for each facility for each year. In order to construct a dataset, we downloaded the entire set of reports (there are approximately 4,500 facilities in the U.S. and the data span nine years from 2002 to 2010) and then read the relevant variables (such as annual facility performance, facility characteristics, patient characteristics, and patient volumes) into Microsoft Excel and later into Stata.

program relative to a control group.<sup>30</sup> This approach is limited by the fact that it cannot disentangle the effect of the program from mean reversion. For example, facilities receiving a *worse than expected* rating may have just experienced a bad draw of patients and will naturally revert back to their underlying average quality. This problem is exacerbated when looking at the extremes of the quality distribution, where differences-in-differences designs can substantially overstate the impact of disclosure programs. As Chay et. al (2005) show, regression discontinuity designs can overcome the problem of mean reversion endemic to differences-in-differences approaches, especially for programs that operate at the extremes of the quality distribution.

Our identification strategy relies on an exogenous change in CMS’s methodology in 2008 that led to markedly fewer facilities receiving the *as expected* rating. The following subsections describe the main parts of our empirical approach in greater detail: the change in methodology used by CMS to assign performance categories, the timeline underlying the analysis, and the intuition behind the regression discontinuity design.

### *A. The Change in CMS Methodology*

When the DFC program was introduced in 2001, facilities were categorized into one of three performance categories based on patient survival rates: *better than expected*, *as expected*, and *worse than expected*. Specifically, a facility was categorized as having a patient survival rate that was *better (worse) than expected* if the upper (lower) confidence limit for the facility’s SMR was lower (greater) than 0.8 (1.2). This method led to the vast majority of facilities (96 percent) being put into the *as expected* category prior to 2008.

In 2008, CMS updated the category thresholds in order to “help consumers make better distinctions among facilities’ survival rates.”<sup>31</sup> In particular, a facility was now classified

---

<sup>30</sup> Pope (2009) is a notable exception. He uses an instrumental variables approach to assess the impact of hospital rankings on future admissions.

<sup>31</sup> In addition to the change in patient survival categorization, DFC was also modified to report two anemia measures—the percentage of patients whose hemoglobin was considered too low (below 10g/dL) or too high (above 12 g/dL)—whereas earlier versions of the tool had only reported the proportion of patients with high hemoglobin levels. See the CMS press release “Medicare Publishes New Information on Quality of Care at

as performing *better than expected* only if its SMR was less than 1.00 and the difference was statistically significant ( $p < 0.05$ ). If the facility's SMR was greater than 1.00 and the difference was statistically significant ( $p < 0.05$ ), the facility was now classified as performing *worse than expected*. All other facilities were classified as performing *as expected*.<sup>32</sup>

As a direct result of this change, a far greater number of firms were now classified as performing *better* or *worse than expected*; on examining the SMRs computed using the 2004-2007 four-year window, approximately 80 percent of facilities now received an *as expected* grade. **Figures 2** and **3** use data from the 2004-2007 timeframe to illustrate the impact of the change in thresholds on the classification of dialysis facilities into performance categories. Many facilities received an unexpected "shock" to their reputations, providing an ideal setting to examine how sellers respond to quality disclosure. In addition, focusing on the impact of the performance grades released in 2008 (based on the 2004-2007 window) ensures that we have sufficient statistical power to use the proposed regression discontinuity design.

### *B. The Timeline of Information Disclosure*

The facility reports using data from the 2004-2007 window were the first to use the new thresholds. As shown in **Figure 4**, this information was revealed at different times to the facilities, consumers, and researchers:

- *June 2008*. Each facility receives a Dialysis Facility Report (DFR) from CMS that contains information on the coarse performance category to which it has been assigned, based on its SMR computed with the 2004-2007 window. Given that the DFR contains the actual value of the SMR, each facility also learns how close it is to the threshold.<sup>33</sup> 2008 is the first year in which facilities' responses to the release of

---

Dialysis Facilities," dated November 20, 2008, for more details (available online at <http://www.cms.gov/apps/media/press/release.asp?Counter=3370>)

<sup>32</sup> Facilities with three or fewer deaths are not included in the classification.

<sup>33</sup> Note that the SMR is based on the comparison of each facility's performance to the performance of firms nationwide, so no facility can gauge how close it is to the performance thresholds based solely on its own unadjusted patient survival rate.

this information are reflected in the data.<sup>34</sup> The middle panel of Table 1 presents summary statistics for facility performance measured in 2008.

- *December 2008.* The DFC website is updated with information on each facility's performance using the 2004-2007 window. Patients looking to undergo dialysis treatments (or currently being treated) can learn the coarse performance category of each facility they are considering. Facilities can now learn where their competitors stand. 2009 is the first year in which facility-level consumer choice data reflect a substantial response to this disclosed information. The bottom panel of Table 1 presents summary statistics for patient flows measured in 2009.
- *December 2010.* The precise mortality data from all four-year windows between 2002 and 2009 is made available on ProPublica.org and made accessible to researchers and to the general public. Before this date, the only facility performance information available to the public was the coarse performance categories on the DFC website (Fields 2010).

### *C. The Regression Discontinuity Design*

As articulated by Imbens and Lemieux (2007), the idea behind using the regression discontinuity design for evaluating the causal effects of interventions is that “assignment to a treatment is determined at least partly by the value of an observed covariate lying on either side of a fixed threshold.”<sup>35</sup> The fundamental identifying assumption is that, close to a threshold of interest, all other characteristics and choices that could influence an outcome will be orthogonal to the treatment being studied. In our empirical specifications, we aim to estimate the magnitude of the impact on various outcome measures when firms receive information about their performance relative to that of their competition. We separately

---

<sup>34</sup> Technically, this effect would be reflected in facility performance only for the latter half of the year, since the reports are distributed only in June. But since we only observe SMRs annually, all our analyses are at the facility-year level.

<sup>35</sup> Note that the covariate may itself be associated with the outcome, but the key assumption is that this association is smooth. Therefore, any discontinuity of the outcome as a function of the covariate at the threshold is taken as evidence for a causal effect of the treatment. See Imbens and Lemieux (2007) for further discussion and for examples of various settings in which the regression discontinuity approach has been used to estimate the causal effects of treatments.

estimate the difference between firms being assigned to the *as expected* and *worse than expected* categories from the difference between being assigned to the *as expected* and *better than expected* categories. Our principal specification is the following equation, estimated on facility-level data, where  $i$  indexes a firm:

$$(1) \quad \text{Outcome}_i = \alpha + \beta * \text{Threshold}_i + \gamma * \text{polynomial of CI}_i + [\lambda * \text{polynomial of CI}_i * \text{Threshold}_i] + [\delta * X_i] + \varepsilon_i$$

We estimate this equation as a pooled regression on a sample window that extends on both sides of each of the thresholds; that is, one specification for the threshold between *worse than expected* and *as expected* and one specification for the threshold between *as expected* and *better than expected*. The primary predictor in these specifications, *Threshold*, is an indicator variable set to 1 if the facility is classified as performing *better than expected* (or *worse than expected* in the corresponding set of regressions), based on mortality data from 2004-2007.<sup>36</sup> In our analysis examining the impact of disclosure on seller quality, we use the national SMR percentile rank of the facility in 2008 as the dependent variable, *Outcome*.<sup>37</sup> We include polynomial functions of the upper (or lower, as the case may be) confidence interval of the SMR in order to control for the underlying trend and allow *Threshold* to identify the discontinuous break in the data. As per Lee and Lemieux (2010), we estimate separate regressions on each side of the threshold by allowing for interactions between *Threshold* and the polynomial terms. We cluster our results at the state level to account for auto-correlation across co-located firms. The results are unchanged without clustering or when clustering by the name of the chain.

We also include control variables in some specifications (denoted by  $X_i$  in brackets); these include the facility's SMR as computed using the 2004-2007 data (to control for underlying quality), the age of the facility (to account for experience-related effects on outcomes), and other organizational characteristics measured in 2007 (indicators for chain affiliation and for-profit ownership, number of total patients, facility size as measured by the number of stations, and competitive intensity of the market as measured by the number of

---

<sup>36</sup> In none of our specifications will *worse than expected* and *better than expected* be included in the same regression equation.

<sup>37</sup> We use the percentile rank instead of the continuous risk-adjusted patient survival rate for ease of exposition. Our conclusions are unchanged if we use the 2008 SMR as the dependent variable.

facilities within a one-mile radius). These control variables are useful in establishing the validity of the regression discontinuity design, as discussed in the next section.

**Figure 5** provides a graphical representation of Equation (1) and illustrates the application of the regression discontinuity design in this context. The figure provides compelling evidence of a discontinuity precisely at the point where a firm is classified either as performing *as expected* or as performing *worse than expected*. A firm that just barely falls into the *worse than expected* category is much more likely to exhibit a substantial improvement in performance in 2008 than a firm that just barely falls into the *as expected* category. In the next section, we discuss our empirical specifications that aim to estimate the magnitude of this improvement.

## V. How Do Firms React to Quality Disclosure?

### A. Establishing the Validity of the Regression Discontinuity Design

In order to establish the validity of the regression discontinuity design, we demonstrate that *Threshold* is not a significant predictor of the baseline control variables measured in 2007. This helps confirm that the assignment of facilities into performance categories is indeed randomized around the cutoffs. We implement this test and report the distribution of the t-statistics of the estimates of *Threshold* in **Appendix Figure A1**. In other words, we estimate separate regressions, analogous to Equation (1), in which we use each of the controls as the main dependent variable. The figure lists the distribution of t-statistics on the *worse than expected* and *better than expected* indicators from each of these regressions with the normal distribution overlaid on top. As can be seen, the distribution of the t-statistics is quite similar to that of the normal distribution and does not exhibit any excess dispersion, thereby providing support for the validity of the regression discontinuity design. Additionally, visual inspection of Figures 2 and 3 reveals that the distribution of firms is continuous around the cutpoints. This suggests that there is no strategic manipulation.

One concern is that our results from the regression discontinuity design could be affected by survivorship bias in the sample. Among the facilities rated *worse than expected* in 2008, if the worst-performing facilities end up exiting the sample (that is, shutting down), one would expect the quality of the surviving facilities to be mechanically higher and thus bias our results upward. In **Appendix Table A1**, we estimate the impact of the 2008 ratings on the probability that a facility in the sample in 2008 exits the sample by 2010.<sup>38</sup> We find that the 2008 quality ratings do not have any impact on the likelihood of exit, thus mitigating concerns about survivorship bias.<sup>39</sup>

### *B. The Impact of Information Disclosure on Performance*

**Table 2** presents coefficient estimates of our primary specification examining the impact of being assigned to the *worse (better) than expected* performance categories on subsequent risk-adjusted survival rates. The estimates in Columns 1-4 represent the effect of being assigned to the *worse than expected* category, while Columns 5-8 present results from specifications that include the *better than expected* indicator. As indicated in Equation (1), we include a third-order polynomial of the lower (upper) confidence interval as a control variable in all specifications examining the impact of being classified *worse (better) than expected*. In our base specifications, the estimates are calculated using a sample window of width 0.25 around the threshold.<sup>40</sup>

The coefficient in Column 1 provides strong evidence that a negative report (a *worse than expected* grade) causes a firm to substantially improve its relative performance ranking in 2008. The magnitude of the effect is quite striking: A facility that is just barely categorized as *worse than expected* experiences an 11-point percentile improvement in the 2008 SMR percentile ranking. The magnitude of the effect is even larger (a 19-point percentile improvement) when we include interaction terms between *Threshold* and the polynomial

---

<sup>38</sup> Acquisitions remain in the sample.

<sup>39</sup> In any case, only two percent of the firms exited the sample.

<sup>40</sup> This corresponds to using observations with lower confidence intervals of the SMR lying between 0.875 and 1.15. For *better than expected* threshold regressions, a sample window range of 0.25 would imply that the upper confidence interval of the SMR would fall between 0.875 and 1.15.

terms in Column 2. In Columns 3 and 4, we show that this result is robust to two major factors that influence regression discontinuity estimates: the size of the sample window and the addition of control variables. In Column 3, we expand the sample window to include all facilities with a lower confidence interval between 0.75 and 1.25; that is, a sample window of width 0.5. This increases the number of firms in the estimation sample from 789 to 1,808. While the coefficient drops in magnitude, the economic significance is still substantial. In Column 4, we include the entire vector of control variables referenced above. The coefficient on the indicator is effectively unchanged, suggesting that the treatment is randomly assigned.<sup>41</sup>

In Columns 5-8, we present results from similar specifications that examine the 2008 SMR percentile ranking of firms just barely categorized as *better than expected*.<sup>42</sup> In contrast to the earlier results, the coefficients are much smaller in magnitude and none are statistically significant. Furthermore, the absolute values of the coefficients are statistically different from each other at the five-percent level for Columns 4 and 8.<sup>43</sup> Taken together, the results in Table 2 suggest that firms respond to disclosure of negative information by improving quality. The response of exceptional firms to positive disclosure is limited and difficult to distinguish from zero.

What exactly do these improvements in percentile ranking imply in terms of patient mortality? As an illustrative example from our data, consider the facilities DCI Boston and Satellite Dialysis Round Rock, which treated 138 and 130 patients, respectively, in 2008. There were 14 recorded patient deaths at DCI Boston and 15 at Satellite Dialysis; the expected death rate per 100 patient-years, which forms the basis for computation of the SMR in 2008, was 19.54. Based on these 2008 SMR figures, DCI Boston and Satellite Dialysis Round Rock had percentile ranks of 24 and 45, respectively. This example

---

<sup>41</sup> While it is surprising that the facility SMR (based on 2004-2007 data) is not a significant predictor of facility performance in 2008, we find that there is indeed a strong and significant association between these measures once we exclude the *Threshold* indicator from the model. These results are presented in Appendix Table A2.

<sup>42</sup> In the Table beneath Figure A1 we show that there is no statistically significant difference in capacity between the facilities that were just barely *better than expected* and those that are just barely *as expected*.

<sup>43</sup> The only difference that is not near standard significance levels is that between the coefficients in Columns 1 and 5. This is attributable to both a lack of power from the narrow sample and the relatively rigid functional form imposed. In appendix Table A3 we show these results using the 2008 SMR, not the ranking, as the dependent variable. In those results we find that the differences in the absolute value of the coefficients is statistically different.

demonstrates the wide fluctuation in short-term performance (measured as rankings from a single year of mortality data) that can result from one or two patient deaths.<sup>44</sup> This likely implies that the quality improvement seen in our specifications arises from facilities implementing changes that result in a few less patient deaths. Given that the average center treats just over 100 patients, this quality improvement is economically significant.

### *C. Robustness Checks*

We examine the robustness of our results to various alternate orders of polynomial controls and to various sample window sizes and present the results in graphical form. **Appendix Figures A2** and **A3** demonstrate the robustness of the coefficient estimates in Table 2 (on the *worse than expected* indicator) to varying the order of the control polynomial for sample windows of width 0.25 and 0.5 around the cutoff, respectively. Our results are also robust to varying the width of the sample window from 0.05 to 0.5 for a third- and fifth-order polynomial, as shown in **Appendix Figures A4** and **A5**.<sup>45</sup> In the next subsection, we examine possible mechanisms underlying this improvement.

### *D. Mechanisms Underlying Performance Improvement*

As seen earlier, the improvement in firm performance is economically meaningful and the magnitude of the effect is robust across multiple specifications. In this subsection, we (a) discuss some possible levers that could help firms achieve this improvement and (b) provide suggestive evidence on the prevalence of these mechanisms.

We begin by analyzing whether the performance improvement is driven largely by changes in the facility's actual death rate or in its expected death rate. The latter is based on the national death rates for patients with the characteristics of the facility's patients; as mentioned earlier, the adjustment includes patient demographics as well as regional variables

---

<sup>44</sup> This is also the reason behind the National Kidney Foundation's decision to use four-year averages for computing performance categories.

<sup>45</sup> Complete results are available from the authors upon request.

such as state death rates. We estimate specifications analogous to Equation (1) but replace the percentile mortality rank with the number of patient deaths (observed and expected) for all patients as dependent variables. We present the results in **Table 3**.

The coefficients indicate that the performance improvement shown by firms receiving a *worse than expected* rating in 2008 is driven by reductions in the observed death rate, while the expected death rate remains unaffected. In other words, such facilities experience fewer patient deaths, while the composition of their patient population (in terms of their likelihood of dying, controlling for characteristics present in the SMR adjustment formula) remains largely unchanged. Firms that received a *better than expected* rating do not experience any change in either measure.

What explains this decrease in patient deaths? We start by looking at how incentives mediate the response to the *worse than expected* rating by facilities. To that end, we estimate specifications analogous to Equation (1), in which we include an interaction term between the *worse than expected* indicator and various measures for how likely firms are to respond to incentives to improve their performance. The results are presented in **Table 4**. (We do not report results for firms receiving a *better than expected* indicator, given such firms' lack of response to disclosure.) As before, we present results for various sample windows, both with and without facility-level control variables.

In Columns 1 and 2, we look at the interaction between the facility's ownership status (measured by a for-profit indicator) and *worse than expected*. If for-profits were more responsive to negative information, we should expect to see a negative coefficient on this interaction term. In both specifications, however, we find imprecise estimates for the interaction effect.

Next, we examine how a firm's response to disclosure is moderated by its competitive environment, as measured by the number of competing firms within a certain radius. If competition is an incentive to improve quality, then one might expect firms in more competitive environments to be more likely to respond to disclosure; that is, the interaction term between *worse than expected* and the number of neighboring facilities should be negative. In Columns 3-6 of Table 4, we present results from specifications testing this hypothesis. We find little support, with the exception of Column 6, in which the interaction

is negative and statistically significant; that is, facilities with more competitors within a five-mile radius exhibit a stronger response to a negative rating.<sup>46</sup> In **Appendix Figure A6**, we test for the robustness of this effect to varying the radius used for measuring the degree of competitive intensity. For all but the smallest radii, the interaction effect is negative and statistically significant.

While the improvements attributable to ownership and competition seem limited, it is still important to understand how they are achieved. One explanation is that facilities are able to drive up their quality ratings by cherry-picking healthier patients (that is, selecting patients along dimensions that are not accounted for in the SMR's risk-adjustment formula) and "firing" others. According to a survey conducted by Desai et al. (2009), a majority of dialysis practitioners and staff believe that dialysis facilities cherry-pick patients by having lower thresholds for turning away patients who fail to comply with dietary or medication guidelines. Because privately insured patients offer larger reimbursements than patients insured by Medicare, facilities have a clear incentive to screen patients based on insurance status as well. Since the performance measure (SMR) is risk-adjusted for a variety of patient comorbidities, facilities engaging in favorable patient selection to improve their SMR would have to screen patients based on factors not accounted for in the risk-adjustment formula, making it difficult for us to directly measure the extent of such behavior.

We overcome this difficulty by using DFC data to construct a proxy for cherry-picking. We exploit the fact that the dataset contains information on the proportion of a facility's patients whose deaths were attributed to another facility. A facility trying to improve its quality rating can push its sickest patients onto other facilities. In our data, we can measure what percentage of patient deaths in a given year were attributed to another facility.<sup>47</sup> Using the 2007 data, we estimate a facility-level regression that uses the above

---

<sup>46</sup> The reason that the interaction term becomes statistically significant in this specification is probably the interaction between *worse than expected* and the facility controls that is introduced into the regression. Despite the rating being identified, we do not have exogenous variation in the number of neighboring facilities and therefore the best we can hope for is to control for some of the confounding factors such as the facility control variables. Simply controlling for the variables is insufficient; one must interact them with *worse than expected* to control for omitted factors that might influence the interaction.

<sup>47</sup> When a patient transfers from a facility, she remains assigned to it in the database for 60 days in order to ensure that each facility's treatment actions are appropriately captured. Once a patient has been transferred out

measure as the dependent variable and the 2007 SMR (which controls for the baseline mortality rate expected from the facility) as the main predictor. This plot is presented in **Figure 6**. We estimate the residual from this regression for each facility and use it as our measure of the *cherry-picking index*. For a given SMR, the *cherry-picking index* is based on the assumption that if a larger proportion of a facility's patient deaths were attributed to another facility, that latter facility was more likely to be cherry-picking. The greater the value of the index, the greater the facility's tendency to cherry-pick. Summary statistics for this index are presented in Table 1.

Our goal is to examine whether facilities that do more cherry-picking also show greater improvement. We present results from tests for this hypothesis in Columns 7 and 8 of Table 4, where we include an interaction term between *worse than expected* and the cherry-picking index.

Across the specifications, the interaction term is negative and statistically significant, implying that facilities prone to cherry-picking also exhibited the strongest performance improvement after disclosure. Based on the coefficients from the fully specified model in Column 8, a one-standard-deviation change in the cherry-picking index leads to an additional three-point improvement in the SMR percentile rank for 2008. Given the base improvement of approximately 19 points, this implies that such firms experience a nearly 15-percent-greater improvement.

Another possible explanation is that facilities improve by investing in process improvement or staffing. Unfortunately, the staffing measures included in the DFC data are too crude to be useful in testing this hypothesis. In particular, the data only contains aggregate measures of staffing, such as the number of full-time employees, but not other necessary details such as the skill level/training of these employees, the number of work hours in each shift, and employee tenure and turnover. In addition, much of the staffing-related improvement, if there is any, could be driven by higher motivation, which is intrinsically difficult to measure.

---

her old facility for 61 days and has been treated in her new facility for 60 days, she is attributed to the new facility.

### *E. The Persistence of Performance Improvement*

We now examine the extent to which this quality improvement persists beyond the first year after disclosure. In **Table 5**, we present results from regression models that use the national SMR percentile ranking in 2009 (one year after the initial quality disclosure) as the dependent variable. In Column 4, we estimate a specification analogous to the base specification in Table 2. We find that a firm with a *worse than expected* rating in 2008 (recall that the *worse than expected* and *better than expected* indicators are based on the 2008 ratings) has a national SMR rank in 2009 that is statistically indistinguishable from the ranking of a firm that is designated as performing *as expected*. Taking these results together with those from the preceding analysis (outlined in Section V.B), we see that the average firm that receives a *worse than expected* rating in 2008 sees a quality improvement that does not persist beyond the initial year.

While we see no effect on average across firms receiving a *worse than expected* rating, one might expect heterogeneity in firm response based on each firm's incentive to respond to disclosure. We exploit the underlying method used to construct the SMR to identify facilities that have the strongest incentive to respond to disclosure. To the extent that the incentive might vary significantly across facilities, we might observe stronger responses from facilities that stand to gain the most from improving. In particular, as shown in **Figure 7**, we identify firms for which 2004 was the "best" year in terms of having the lowest annual SMR within the 2004-2007 window. Since the performance categories for each year are assigned based on computing average SMRs over the previous four years, the facilities for which 2004 was the best year would experience a mechanical decline in performance (in terms of the 2009 categories) relative to other facilities (conditional on 2008 SMR) once the time window shifts to 2005-2008. Based on this reasoning, we characterize such facilities as having a stronger incentive to institute performance improvements in 2008 compared to facilities for which the best year was 2005, 2006, or 2007.

We capture this variation in the incentive to respond to disclosure by including an interaction term between *worse than expected* and an indicator for whether 2004 was the best year for that facility; the results from these specifications are in Columns 1-3 and 5-7 of Table 5. In Columns 1-3, we examine the national SMR percentile rank in 2008 and find that the interaction between *worse than expected* and an indicator for whether 2004 was the facility's best year is small but imprecisely estimated. In other words, immediately after receiving the rating, the response is homogenous whether or not there will be a mechanical decline in the near future. Across Columns 4-7, we look at performance in 2009. While, as previously discussed, column 4 indicated limited persistence on average, the interaction term in Columns 5-7 is negative and statistically significant, indicating that, among firms receiving a *worse than expected* rating in 2008, the performance improvement persists the following year only for those that would face a mechanical decline in their quality ratings in 2009; which is to say, those firms with the strongest incentive to respond.

In summary, we find that facilities that receive negative information (a *worse than expected* rating) from quality disclosure display an improvement in short-term performance, as measured by SMR in 2008. We find evidence suggesting that the improvement is in part driven by increased cherry-picking of patients and that it persists only for facilities that have a stronger incentive to bring their SMRs down in 2008.

## VI. How Does Quality Disclosure Affect Patient Flow?

### *A. Effect of disclosure on patient volume and composition*

**Table 6A** presents coefficients from specifications that aim to estimate the impact of information disclosure on the number and type of patients choosing a facility. Because patients could not see the coarse performance data until December 2008, we would expect the impact of the new disclosure policy to be felt on the number or composition of new patients only in 2009. We estimate specifications with various dependent variables that measure the volume and composition of patients seen by the facility in 2009. As before, we

estimate separate regressions that include the *worse than expected* and *better than expected* indicators as predictors. We present specifications in which we use sample ranges of 0.25 and 0.5 around the threshold and we include both polynomial interaction terms and facility controls.

Columns 1 and 2 present the impact of receiving a *worse than expected* rating on the total number of patients treated at the facility in 2009. The estimates are noisy and the large standard errors make it difficult to pinpoint the direction of the effect. The same holds true for firms just passing the *better than expected* threshold (as seen by the coefficients in Columns 3 and 4). While we find little evidence of an overall effect, we also consider the impact of disclosure on the composition of patients at a facility. In particular, we focus on proxies for how well-informed patients might be, based on whether or not they consulted a nephrologist prior to starting dialysis (Nephrologists are kidney specialists who help patients manage the course of ESRD). Our assumption is that a patient who did not see a nephrologist before starting dialysis is less likely to be aware of the facility performance rankings. The coefficients from these specifications, presented in Columns 5 and 6, provide strong evidence that when a facility is just barely within the *worse than expected* category, an additional 13 percent of the patients treated in 2009 will not have seen a nephrologist prior to dialysis. This finding suggests that better-informed patients are shying away from facilities deemed *worse than average*.<sup>48</sup> In Columns 7 and 8, we find no corresponding impact on patient composition for facilities that just barely fall within the *better than expected* category.

We also study the effect of quality disclosure on the profile of the pool of patients at a facility in 2009. In particular, we focus on two measures—the percentage of patients admitted in 2009 who transferred from other facilities and the percentage of patients in 2009 who were continuing treatment from 2008. Because a patient transferring from another facility is likely to be more familiar with the system, this measure acts as a proxy for being

---

<sup>48</sup> Our interpretation of this result is based on the assumption that patients who do not visit a nephrologist prior to starting dialysis are less informed about the dialysis process. If the variable measuring the percentage of patients who visited a nephrologist were instead acting as a proxy for the percentage of patients who are healthier—that is, if patients who are sicker are more likely to visit a nephrologist prior to dialysis—the result in Table 5 could be interpreted to mean that facilities rated *worse than expected* in 2008 were treating healthier patients, on average, in 2009. However, when we examine the standardized mortality rates for new patients in 2008 in Appendix Table A2, we find that patient mortality *increases* in the proportion of patients not visiting a nephrologist prior to starting dialysis, making this alternate hypothesis unlikely.

well-informed.<sup>49</sup> **Table 6B** presents results from using these two measures as dependent variables in specifications analogous to the ones estimated earlier. Facilities receiving a *worse than expected* rating see a decline in the proportion of transfer patients and correspondingly treat more patients who are continuing at that facility. This suggests that patients who have been on dialysis longer are more likely to avoid facilities with a *worse than expected* rating.

## VII. Conclusions

The evidence in this paper is consistent with information disclosure programs being most effective when negative information is highlighted. Further, we provide evidence suggesting that the disclosure-prompted performance improvement seen in dialysis facilities seems to be driven substantially through cherry-picking healthier patients. In addition, our finding on the composition of new patients (less knowledgeable patients attending poorly rated facilities) suggest that the information on the DFC website is not being used by all of the patients seeking dialysis and that, more broadly, information disclosure programs can exacerbate inequality between the informed and the uninformed.

Some caveats to our findings are in order. First, our measure of performance improvement is based on short-term facility responses, so our analyses do not fully capture the entire impact on quality, some of which may only manifest over the long term. Second, because the dialysis industry has a few unique features—such as lack of price competition—our findings may not be equally applicable to all settings. Nevertheless, our results clearly imply that designing quality disclosure programs in ways that make the information more salient to firms and consumers could lead to substantial welfare improvements.

---

<sup>49</sup> Surprisingly, many patients are not familiar with the DFC website. According to an evaluation of the DFC tool carried out in 2006, patients were found to be more cognizant of available options after being on dialysis for a certain period of time. The initial few months of dialysis are often an emotionally trying time for patients, limiting their decision-making abilities. Over time, patients come to learn about the range of dialysis facilities available. See Trisolini et al. (2006) for more details.

## Bibliography

Centers for Medicare and Medicaid Services. 2008. "Medicare Publishes New Information on Quality of Care at Dialysis Facilities [Press release]", retrieved from <http://www.cms.gov/apps/media/press/release.asp?Counter=3370>

Chay, K., P. McEwan, and M. Urquiola. 2005. "The central role of noise in evaluating interventions that use test scores to rank schools", *American Economic Review*, 95: 1237-1258.

Chen, M. 2008. "Minimum Quality Standards and Strategic Vertical Differentiation: An Empirical Study of Nursing Homes", *Northwestern University PhD Dissertation*.

Dafny, L., and D. Dranove. 2008. "Do report cards tell consumers anything they don't already know? The case of Medicare HMOs", *RAND Journal of Economics*, 39(3): 790–821.

DellaVigna, S., and Pollet, J. 2009. "Investor Inattention and Friday Earnings Announcements", *Journal of Finance*, 64(2): 709-749.

Desai, A., R. Bolus, A. Nissenson, S. Bolus, M. Solomon, O. Khawar, M. Gitlin, J. Talley, and B. Spiegel. 2008. "Identifying Best Practices in Dialysis Care: Results of Cognitive Interviews and a National Survey of Dialysis Providers", *Clinical Journal of the American Society of Nephrology*, 3(4): 1066-1076.

Desai, A., R. Bolus, A. Nissenson, G. Chertow, S. Bolus, M. Solomon, O. Khawar, J. Talley, and B. Spiegel. 2009. "Is there "Cherry Picking" in the ESRD Program? Perceptions from a Dialysis Provider Survey", *Clinical Journal of the American Society of Nephrology*, 4: 772-777.

Dranove, D., and G. Jin. 2010. "Quality disclosure and certification: Theory and practice", *Journal of Economic Literature*, 48(4): 935–963.

Dranove, D., D. Kessler, M. McClellan, and M. Satterthwaite. 2003. "Is More Information Better? The Effects of 'Report Cards' on Health Care Providers", *Journal of Political Economy*, 111(3): 555–588.

Dranove, D., and A. Sfekas. 2008. "Start spreading the news: A structural estimate of the effects of New York hospital report cards", *Journal of Health Economics*, 27(1): 1201–1207.

Fields, R. 2010. "God Help You. You're on Dialysis", *Atlantic Monthly*, December: 1–11.

Forbes, S., M. Lederman, and T. Tombe. 2012. "Quality Disclosure Programs with Thresholds: Misreporting, Gaming and Employee Incentives", *Working Paper*.

- Frederick, P., N. Maxey, S. Clauser, and J. Sugarman. 2002. “Developing Dialysis Facility-Specific Performance Measures for Public Reporting”, *Health Care Financing Review*, 23(4): 37-50.
- Garg, P., K. Frick, M. Diener-West, and N. Powe. 1999. “Effect of the ownership of dialysis facilities on patients' survival and referral for transplantation”, *New England Journal of Medicine*, 341(22): 1653–1660.
- Hastings, J., and J. Weinstein. 2008. “Information, School Choice, and Academic Achievement: Evidence from Two Experiments”, *Quarterly Journal of Economics*, 123(4): 1373-1414.
- Imbens, G., and T. Lemieux. 2007. “Regression discontinuity designs: A guide to practice”, *Journal of Econometrics*, 142(2): 615–635.
- Jin, G., and P. Leslie. 2003. “The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards”, *Quarterly Journal of Economics*, 118(2): 409–451.
- Jin, G., and A. Sorenson. 2006. “Information and consumer choice: The value of publicized health plan ratings”, *Journal of Health Economics*, 25(2): 248–275.
- Kahneman, D., and A. Tversky. 1979. “Prospect Theory: An Analysis of Decision under Risk”, *Econometrica*, 47(2): 263–291.
- Kolstad, J. 2012. “Information and Quality when Motivation is Intrinsic: Evidence from Surgeon Report Cards”, *Working Paper*.
- Lawler, M., G. Doeksen, F. Eilrich, and V. Schott. 2003. “A Systems Development Guide for a Kidney Dialysis Centre”, available online at [http://www.ruralhealthworks.org/dl\\_additional.html](http://www.ruralhealthworks.org/dl_additional.html)
- Lazear, E., and S. Rosen. 1981. “Rank Order Tournaments as Optimum Labor Contracts”, *Journal of Political Economy*, 89(5): 841–864.
- Lee, D., and T. Lemieux. 2010. “Regression discontinuity designs in economics”, *Journal of Economic Literature*, 48(2): 281-355.
- Luca, M. 2011. “Reviews, Reputation and Revenue: The Case of Yelp.com”, *Harvard Business School Working Paper* No. 12-016.
- Mas, A. 2006. “Pay, reference points, and police performance”, *Quarterly Journal of Economics*, 121(3): 783–821.

Nissenson, A., and R. Rettig. 1999. "Medicare's end-stage renal disease program: Current status and future prospects", *Health Affairs*, 18(1): 161–179.

Pope, D. 2009. "Reacting to rankings: Evidence from 'America's Best Hospitals'", *Journal of Health Economics*, 28(1): 1154–1165.

Powe, N., M. Thamer, W. Hwang, N. Fink, E. Bass, J. Sadler, and N. Levin. 2002. "Cost-quality trade-offs in dialysis care: A national survey of dialysis facility administrators", *American Journal of Kidney Diseases*, 39(1): 116–126.

Scanlon, D., M. Chernew, and C. McLaughlin. 2002. "The impact of health plan report cards on managed care enrollment", *Journal of Health Economics*, 21(1): 19–41.

Spiegel B., R. Bolus, A. Desai, P. Zagar, T. Parker, J. Moran, M. Solomon, O. Khawar, M. Gitlin, J. Talley, and A. Nissenson. 2010. "Dialysis Practices that Distinguish Facilities with Below- versus Above-Expected Mortality", *Clinical Journal of the American Society of Nephrology*, 5(11): 2024-2033.

Stigler, G. 1961. "The Economics of Information", *Journal of Political Economy*, 69(3): 213-225.

Trisolini, M., E. Zerhusen, K. Bandel, A. Roussel, P. Frederick, D. Schatell, and S. Harris. 2006. "Evaluation of the Dialysis Facility Compare Website Tool on Medicare.gov", *Dialysis and Transplantation*, 35(4): 196-214.

United States Renal Data System. 2011. "United States Renal Data System Annual Report", available online at < <http://www.usrds.org/adr.aspx> >

Wedig, G., and M. Tai-Seale. 2002. "The effect of report cards on consumer choice in the health insurance market", *Journal of Health Economics*, 21(6): 1031-1048.

Wolfe, W. 2011. "Adequacy of Dialysis Clinic Staffing and Quality of Care: A Review of Evidence and Areas of Needed Research", *American Journal of Kidney Disease*, 58(2): 166-176.

Zhang, Y., D. Cotter, and M. Thamer. 2011. "The Effect of Dialysis Chains on Mortality among Patients Receiving Hemodialysis", *Health Services Research*, 46(3): 747-767.

Table I: Summary Statistics

Variables	Observations	Mean	Standard deviation	Minimum	Maximum
<i>Facility Characteristics in 2007</i>					
Standardized mortality ratio: 2004-2007	4512	1.00	.29	0	4.73
Worse than expected	4512	.10	.30	0	1
Better than expected	4512	.10	.31	0	1
Chain-affiliated	4490	.81	.40	0	1
For-profit	4490	.81	.39	0	1
Facility age	4512	12.32	8.61	0	41
Total patients	4512	111.99	67.88	0	661
Total stations	4490	17.98	8.32	0	80
In-unit patient : station ratio	4449	3.76	1.85	0	70
Number of neighboring facilities within 1 mile	4512	.41	.84	0	7
Best year for standardized mortality ratio = 2004	3919	.21	.41	0	1
% of own patient deaths attributed to other facilities	4477	.09	.10	0	1
Cherry-picking index	4466	0	.10	-.13	.87
<i>Facility Performance in 2008</i>					
National standardized mortality percentile rank	4485	50.03	28.10	1.00	99
Observed deaths per 100 patient-years	4509	21.53	10.36	0	139.11
Expected deaths per 100 patient-years	4509	21.18	6.18	4.22	99.15
<i>Patient Composition Characteristics in 2009</i>					
Total patients	4512	113.79	67.56	0	680
% never seen a nephrologist prior to dialysis	4464	.30	0.21	0	1
% of patients continuing	4504	.68	.10	0	1
% of patients transferred in	4504	.12	0.08	0	1

Note: There are no observations for 513 facilities in 2004. Zeros for total stations and total patients indicate the facility was not in operation during 2007. For 121 observations, the year the facility opened and the total patients in 2007 is missing.

Table 2: Does Facility Performance Respond to Information Disclosure?

Independent variables	Dependent variable: National standardized mortality percentile rank (2008)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Worse than expected	-11.03 (4.40)***	-19.17 (6.29)***	-15.22 (5.30)***	-14.89 (5.21)***				
Better than expected					4.40 (3.53)	6.74 (5.30)	2.75 (3.60)	3.07 (3.72)
Standardized mortality ratio: 2004-2007				8.01 (7.65)				8.06 (10.06)
Chain-affiliated				-1.28 (1.31)				-1.17 (1.26)
For-profit				2.84 (2.09)				5.31 (1.37)***
Facility age				-.04 (.09)				.06 (.08)
Total patients				-.00 (.09)				-.00 (.01)
Total stations				.21 (.15)				.22 (.12)*
In-unit patient : station ratio				.94 (.65)				.03 (.70)
Number of neighboring facilities within 1 mile				.43 (.52)				.28 (.81)
Order of SMR confidence interval polynomial	3	3	3	3	3	3	3	3
Worse than expected X Polynomial		Yes	Yes	Yes				Yes
Better than expected X Polynomial						Yes	Yes	Yes
Sample window	.25	.25	.5	.5	.25	.25	.5	.5
Observations	789	789	1808	1808	904	904	1765	1765

Note: \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% confidence levels, respectively. Parentheses contain standard errors clustered at the state level.

**Table 3: Are Facility Performance Changes Driven by Observed Deaths or Expected Deaths?**

Independent variables	Dependent variable: Observed deaths per 100 patient-years			Dependent variable: Expected deaths per 100 patient-years				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Worse than expected	-8.14 (4.65)*	-6.50 (2.90)**			-.84 (1.78)	-1.09 (1.11)		
Better than expected			1.99 (1.71)	.34 (.95)			.70 (1.57)	-.90 (.81)
Order of SMR confidence interval polynomial	3	3	3	3	3	3	3	3
Worse than expected X Polynomial	Yes	Yes			Yes	Yes		
Better than expected X Polynomial			Yes	Yes			Yes	Yes
Facility control variables		Yes		Yes			Yes	Yes
Sample window	.25	.5	.25	.5	.25	.5	.25	.5
Observations	790	1809	970	1759	790	1809	907	1759

Note: \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% confidence levels, respectively. Parentheses contain standard errors clustered at the state level.

Table 4: What Explains Heterogeneity in Response to Information Disclosure?

Independent variables	Dependent variable: National standardized mortality percentile rank (2008)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Worse than expected	-13.04 (9.37)	-20.24 (23.82)	-19.17 (6.64)***	-14.98 (5.53)***	-14.35 (5.47)**	-21.75 (23.69)	-19.06 (6.23)***	-19.59 (23.54)
For-profit	4.40 (5.61)	2.59 (2.37)						
Worse than expected X For-profit	-6.65 (7.99)	.67 (6.28)						
Number of neighboring facilities			.50 (1.18)	.57 (.63)	.27 (.09)***	.32 (.09)***		
Worse than expected X Number of neighboring facilities			-.001 (2.89)	-.64 (1.90)	-.19 (.17)	-.51 (.14)***		
Cherry-picking index							21.99 (10.82)**	10.81 (7.19)
Worse than expected X Cherry-picking index							-56.57 (20.75)***	-33.05 (15.77)**
Order of SMR confidence interval polynomial	3	3	3	3	3	3	3	3
Worse than expected X Polynomial	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Facility control variables		Yes		Yes	Yes	Yes	Yes	Yes
Worse than expected X Facility control variables		Yes		Yes	Yes	Yes	Yes	Yes
Radius for number of neighboring facilities	1 mile	1 mile	1 mile	1 mile	5 miles	5 miles	1 mile	1 mile
Sample window	.25	.5	.25	.5	.5	.5	.25	.5
Observations	786	1808	789	1823	1823	1808	788	1808

Note: \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% confidence levels, respectively. Parentheses contain standard errors clustered at the state level. In specifications 3-8, for-profit status is controlled for but not presented. In specifications 1-2 and 7-8, the number of neighboring facilities is controlled for but not presented. This is done to limit visual clutter in the tables. For the specifications in which *worse than expected* is interacted with the facility controls, the baseline estimate of *worse than expected* is of limited informative value. Because of space constraints we only show two specifications for interactions with for-profit and cherry-picking. The results are robust to specifications similar to columns 2 & 3 of Table 3.

Table 5: When Does the Effect Persist Past the First Year?

Independent variables	Dependent variable: National standardized mortality percentile rank (2008)			Dependent variable: National standardized mortality percentile rank (2009)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Worse than expected	-20.13 (6.19)***	-14.89 (5.46)***	-12.24 (43.87)	7.56 (6.18)	8.79 (6.69)	6.00 (7.59)	-67.32 (45.95)
Best year for standardized mortality ratio = 2004	-1.11 (2.66)	.65 (1.73)	.57 (1.71)		5.91 (2.52)**	5.32 (1.83)***	5.20 (1.83)***
Worse than expected X Best year for standardized mortality ratio = 2004	3.81 (4.64)	.33 (3.80)	.27 (3.71)		-8.37 (3.94)**	-9.05 (2.92)***	-9.26 (3.21)***
Order of SMR confidence interval polynomial	3	3	3	3	3	3	3
Worse than expected X Polynomial	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Facility control variables		Yes	Yes			Yes	Yes
Worse than expected X Facility control variables			Yes			Yes	Yes
Sample window	.25	.5	.5	.25	.25	.5	.5
Sample window	756	1716	1716	790	757	1717	1717

Note: \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% confidence levels, respectively. Parentheses contain standard errors clustered at the state level. The smaller number of observations is due to the fact that the facility must have been active in 2004. For the specifications in which the threshold is interacted with the facility controls, the baseline estimate of the threshold is of limited informative value.

**Table 6A: Does Information Disclosure Predict Patient Characteristics in 2009?**

Independent variables	Dependent variable: Total patients				Dependent variable: % never seen a nephrologist prior to dialysis			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Worse than expected	2.51 (11.44)	5.91 (7.20)			.09 (.03)**	.13 (.04)***		
Better than expected			-7.38 (11.56)	-5.19 (5.95)			.04 (.03)	.03 (.04)
Order of SMR confidence interval polynomial	3	3	3	3	3	3	3	3
Worse than expected X Polynomial	Yes	Yes			Yes	Yes		
Better than expected X Polynomial			Yes	Yes			Yes	Yes
Facility control variables		Yes		Yes		Yes		Yes
Sample window	.25	.5	.25	.5	.25	.5	.25	.5
Observations	791	1810	907	1759	786	1803	897	1742

**Table 6B: Does Information Disclosure Predict the Composition of the Admitted Patient Pool in 2009?**

Independent variables	Dependent variable: % of patients transferred in				Dependent variable: % of patients continuing			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Worse than expected	-0.03 (.01)**	-0.03 (.01)**			.02 (.03)	.04 (.02)**		
Better than expected			-0.01 (.01)	-0.01 (.01)			.02 (.02)	-0.01 (.01)
Order of SMR confidence interval polynomial	3	3	3	3	3	3	3	3
Worse than expected X Polynomial	Yes	Yes			Yes	Yes		
Better than expected X Polynomial			Yes	Yes			Yes	Yes
Facility control variables		Yes		Yes		Yes		Yes
Sample window	.25	.5	.25	.5	.25	.5	.25	.5
Observations	788	1805	907	1757	788	1805	907	1757

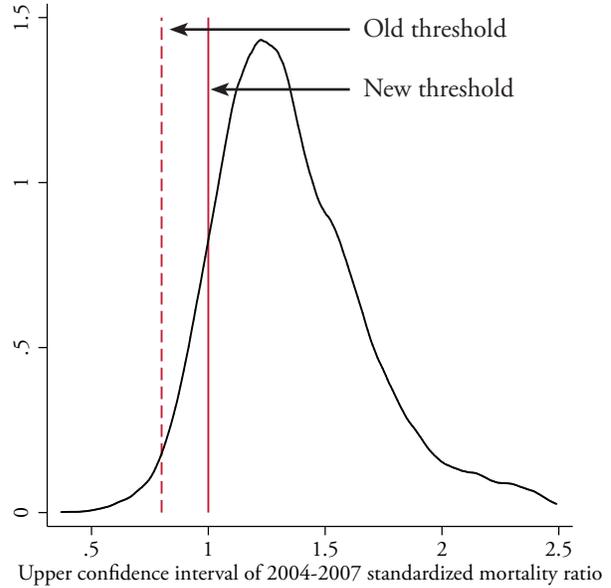
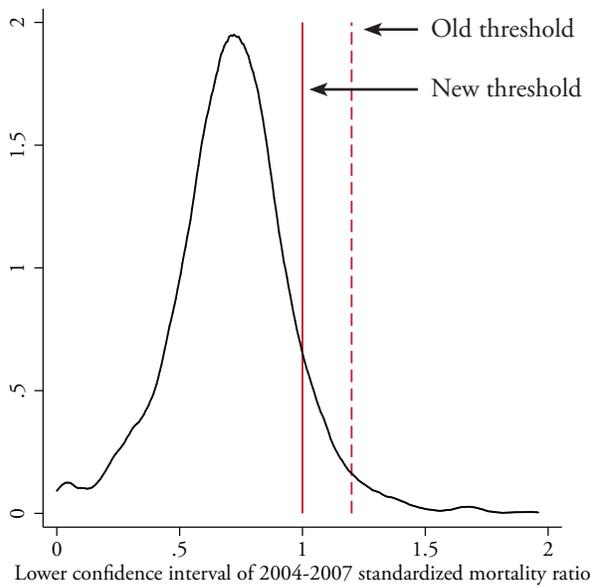
Note: \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% confidence levels, respectively. Parentheses contain standard errors clustered at the state level.

Figure 1: Distribution of Dialysis Facility Performance



Note: This figure presents the Kernel density of the distribution of the standardized mortality ratio reported for each facility in 2008 and computed using mortality data from 2004-2007.

Figures 2 & 3: Change in Quality Reporting Thresholds for Dialysis Centers (2008)



## Figure 4: Timeline of Information Release

The new performance categories (based on 2004-2007 data) are released on Dialysis Facility Compare website

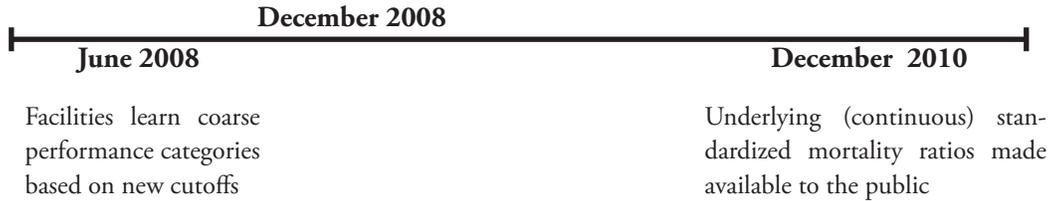
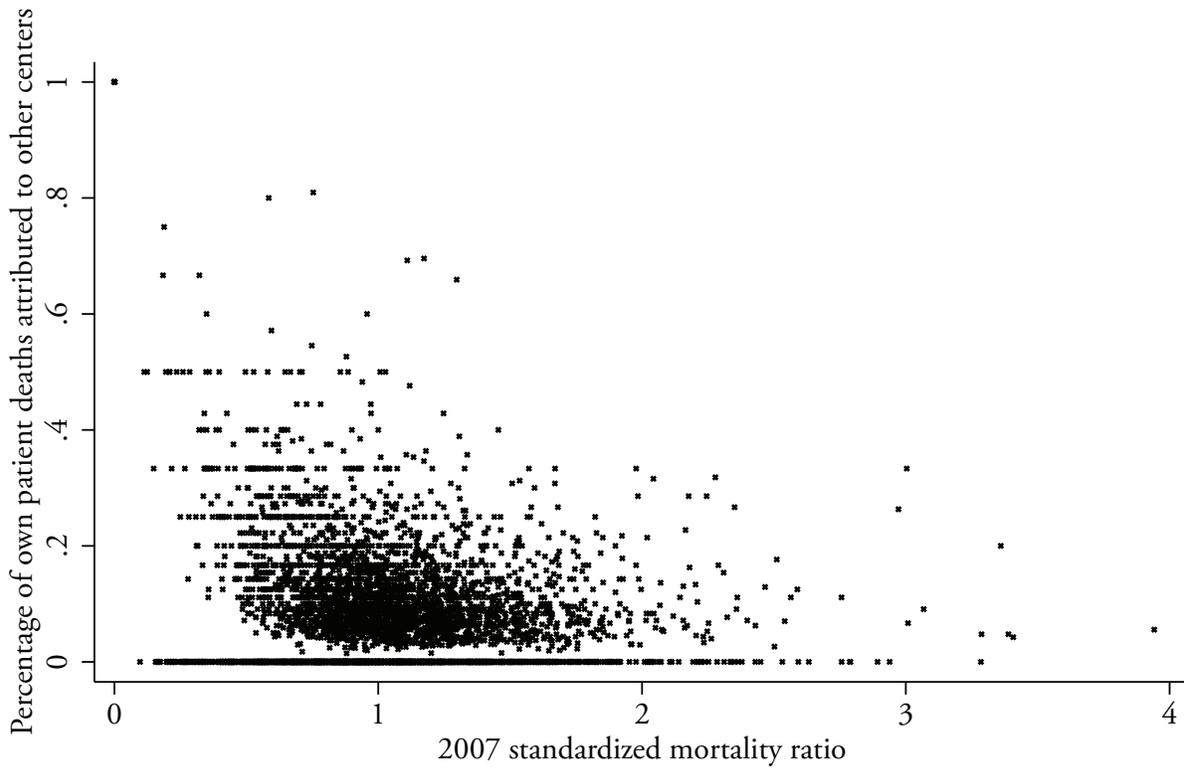


Figure 5: Regression Discontinuity Estimate of the Effect of Disclosure on Subsequent Performance



Note: \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% confidence levels, respectively. Parentheses contain standard errors clustered at the state level.

Figure 6: Transferring Dying Patients and the Standardized Mortality Ratio (2007)



Note: Data for 4,466 facilities come from 2007.

Figures 7: Identifying Facilities with Strong Incentives to Improve Performance: CMS Disclosures for Davita Dialysis Facility - Belair, MD

Measure Name	This Facility
3. Patient survival reported as "as expected," "better than expected," or "worse than expected" for the time period 2004-2007 for this facility	Worse than Expected
Standardized Mortality Ratio (SMR):	1.18
P-value:	0.05

*2004 is the best year. The facility will mechanically decline in future quality.*

Standardized Mortality Ratio (SMR)	2004	2005	2006	2007	2004 - 2007
1j SMR <sup>4</sup>	0.95	1.65	1.17	0.96	1.18
1k P-value <sup>5</sup>	0.89	<.01	0.38	0.91	0.05
1l 95% Confidence interval for SMR <sup>6</sup>					
Upper limit	1.37	2.17	1.60	1.37	1.39
Lower limit	0.64	1.23	0.83	0.65	1.00

Note: These original images are taken directly for the Dialysis Facility Compare website and the 2008 Dialysis Facility Reports. This is to show what the facilities actually see. Information on footnotes and the broader context of these images can be found in these original sources.

Table AI: Information Disclosure Does Not Affect Facility Exit

Independent variables	Dependent variable: Facility drops out of sample: 2010					
	(1)	(2)	(3)	(4)	(5)	(6)
Worse than expected	.01 (.04)	-.002 (.04)	.02 (.04)			
Better than expected				.01 (.03)	.03 (.06)	.03 (.03)
Degree of SMR confidence interval polynomial	3 <sup>rd</sup>	3 <sup>rd</sup>	3 <sup>rd</sup>	3 <sup>rd</sup>	3 <sup>rd</sup>	3 <sup>rd</sup>
Worse than expected X Polynomial		Yes	Yes			
Better than expected X Polynomial					Yes	Yes
Sample window	.25	.25	.5	.25	.25	.5
Observations	810	810	1880	927	927	1810

Note: \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% confidence levels, respectively. Parentheses contain standard errors clustered at the state level. In specification (4), the data perfectly predict the regression because so few firms exit.

Table A2: Predictors of Performance

Independent variables	Dependent variable: Standardized mortality ratio: 2008		Independent variables	Dependent variable: New patient standardized mortality ratio: 2008	
	(1)	(2)		(1)	(2)
Standardized mortality ratio: 2004-2007	.48 (.02)***	.47 (.03)***	Average age of new patients: 2008	.004 (.003)	.004 (.003)
Chain-affiliated		-.02 (.01)	% of patients female: 2008	-.001 (.001)	-.001 (.001)
For-profit		.07 (.02)**	% of patients Asian: 2008	.002 (.005)	.003 (.004)
Facility age		-.001 (.001)	% of patients Black: 2008	.005 (.004)	.005 (.004)
Total patients		.0001 (.0001)	% of patients Native American: 2008	-.000 (.004)	-.001 (.004)
Total stations		.001 (.001)	% of patients White: 2008	.003 (.004)	.003 (.004)
In-unit patient : station ratio		.001 (.003)	% of patients hispanic: 2008	.002 (.001)**	.002 (.001)
Neighboring facilities count within 1 mile		.02 (.01)**	% of patients diabetic: 2008	.000 (.001)	.000 (.001)
Observations	4485	4426	Average female new patient BMI: 2008	.002 (.002)	.002 (.002)
			Average male new patient BMI: 2008	-.006 (.003)**	-.005 (.003)**
			% never seen a nephrologist prior to dialysis: 2008		.004 (.001)***
			Observations	4084	4084

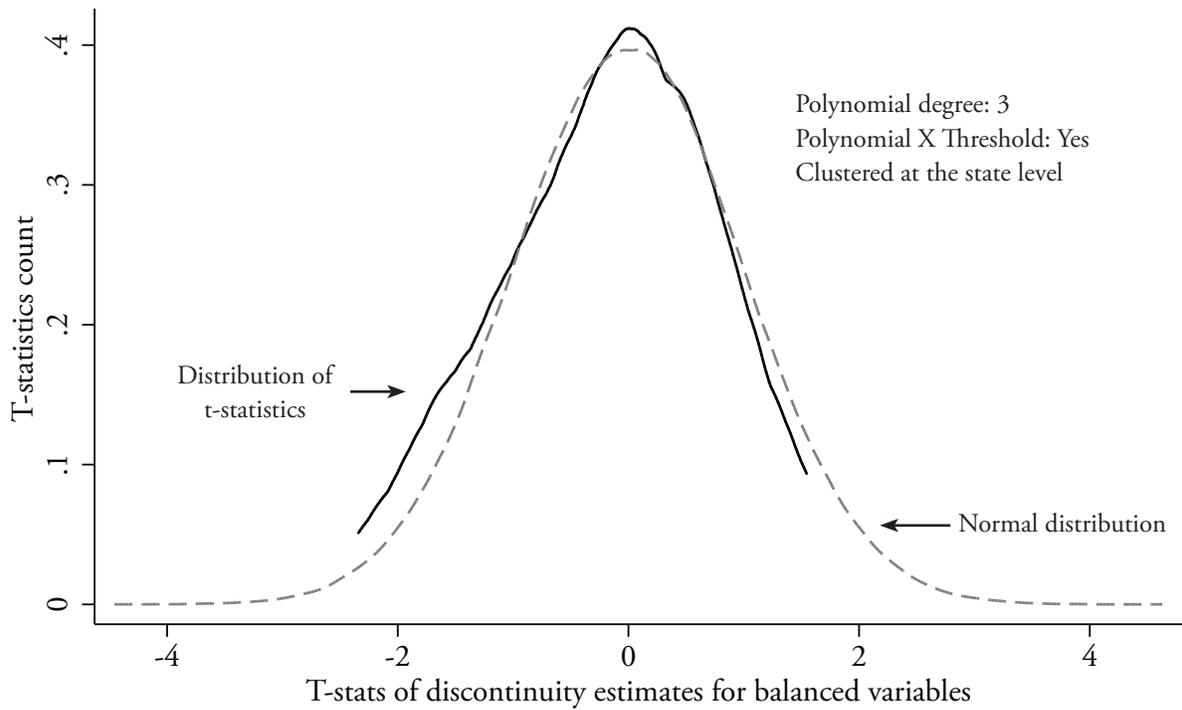
Note: \*, \*\*, \*\*\* indicate significance at the 10%, 5%, and 1% confidence levels, respectively. Parentheses contain standard errors clustered at the state level.

**Table A3: Does Facility Performance Respond to Information Disclosure?**

Independent variables	Dependent variable: Standardized mortality ratio: 2008							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Worse than expected	-.201 (.071)***	-.317 (.099)***	-.249 (.078)***	-.241 (.077)***				
Better than expected					.062 (.041)	.083 (.058)	.039 (.040)	.044 (.042)
Facility control variables				Yes				Yes
Order of SMR confidence interval polynomial	3	3	3	3	3	3	3	3
Worse than expected X Polynomial		Yes	Yes	Yes				
Better than expected X Polynomial						Yes	Yes	Yes
2004 Worst SMR dummy	.25	.25	.5	.5	.25	.25	.5	.5
2004 Worst SMR dummy	789	789	1808	1808	904	904	1765	1765

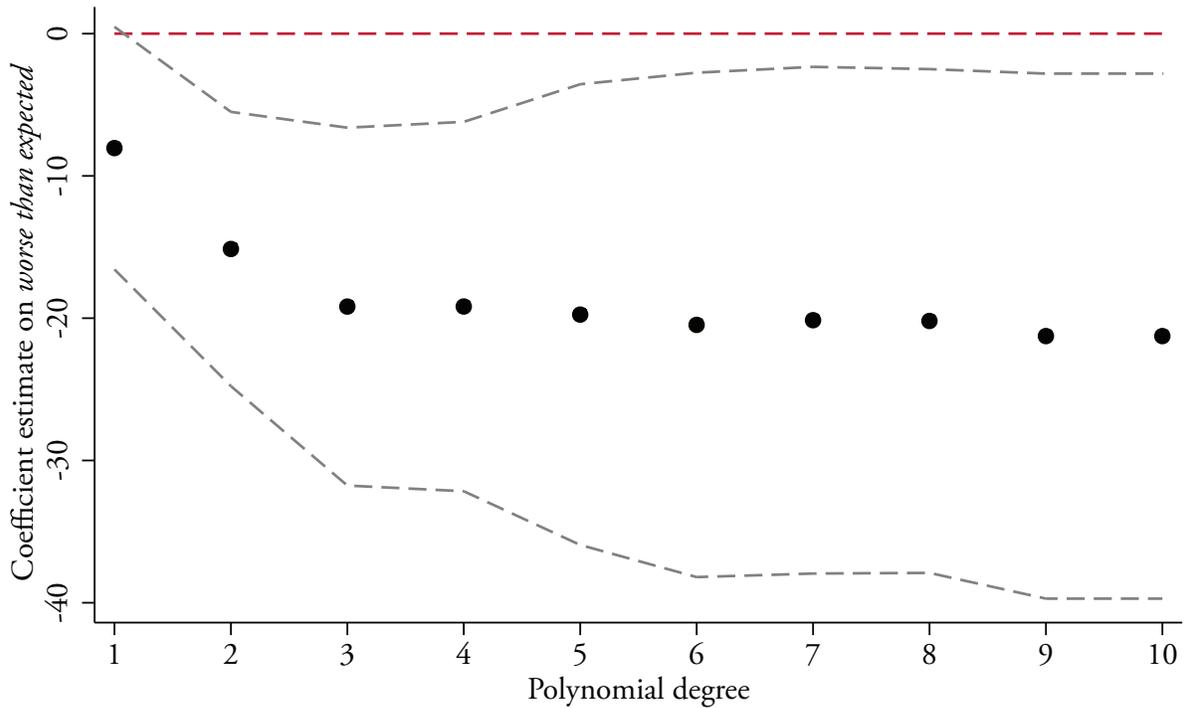
Note: \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% confidence levels, respectively. Parentheses contain standard errors clustered at the state level.

Figure AI: Distribution of T-Stats of Threshold Estimates for Different Outcomes



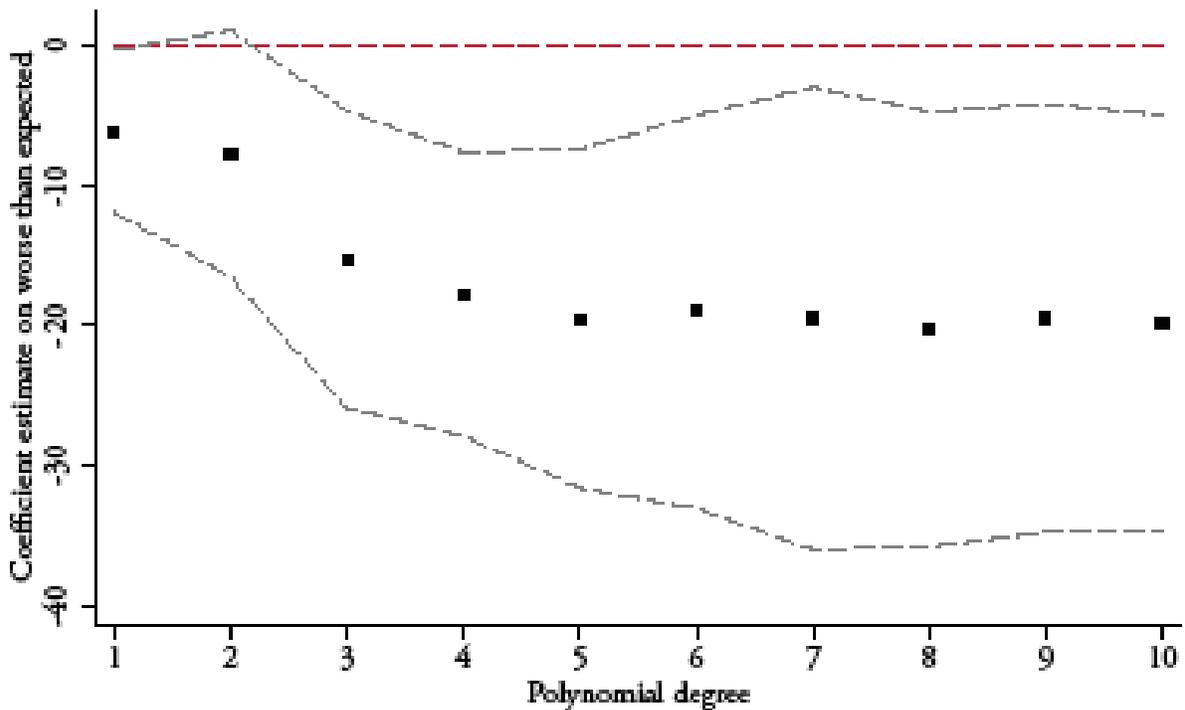
Variable	Worse t-stat	Better t-stat	Variable	Worse t-stat	Better t-stat
Standardized mortality ratio: 2004-2007	-.93	-.66	% of patients diabetic: 2007	.13	.33
Best year for standardized mortality ratio = 2004	-1.51	.95	Average female new patient BMI: 2007	.20	-.01
Chain-affiliated	.02	1.13	Average male new patient BMI: 2007	-1.86	.38
For-profit	.70	.07	Observed deaths per 100 patient-years: 2007	-.74	-.42
Facility age	.35	-1.56	Expected deaths per 100 patient-year: 2007	-.47	.52
Total patients	-.19	-1.23	New patient first year standard mortality ratio: 2007	-1.06	1.17
Total stations	.16	-1.41	New patient observed deaths per 100 patient years: 2007	-.53	.66
In-unit patient : station ratio	-.26	-.07	New patient expected deaths per 100 patient year: 2007	.38	.62
Neighboring facilities count within 1 mile	-.52	-2.09	% of all patients new to dialysis: 2007	-1.16	1.18
Average age of new patients: 2007	-.05	-.17	% of all patients transferred into facility: 2007	.19	-.95
% of patients White: 2007	.39	1.55	% never seen a nephrologist prior to dialysis: 2007	1.18	.87
% of patients female: 2007	-1.25	-.54	% of own patient deaths attributed to other facilities: 2007	-.29	-2.34

Figure A2: Robustness to Polynomial Degree for .25 Window Size



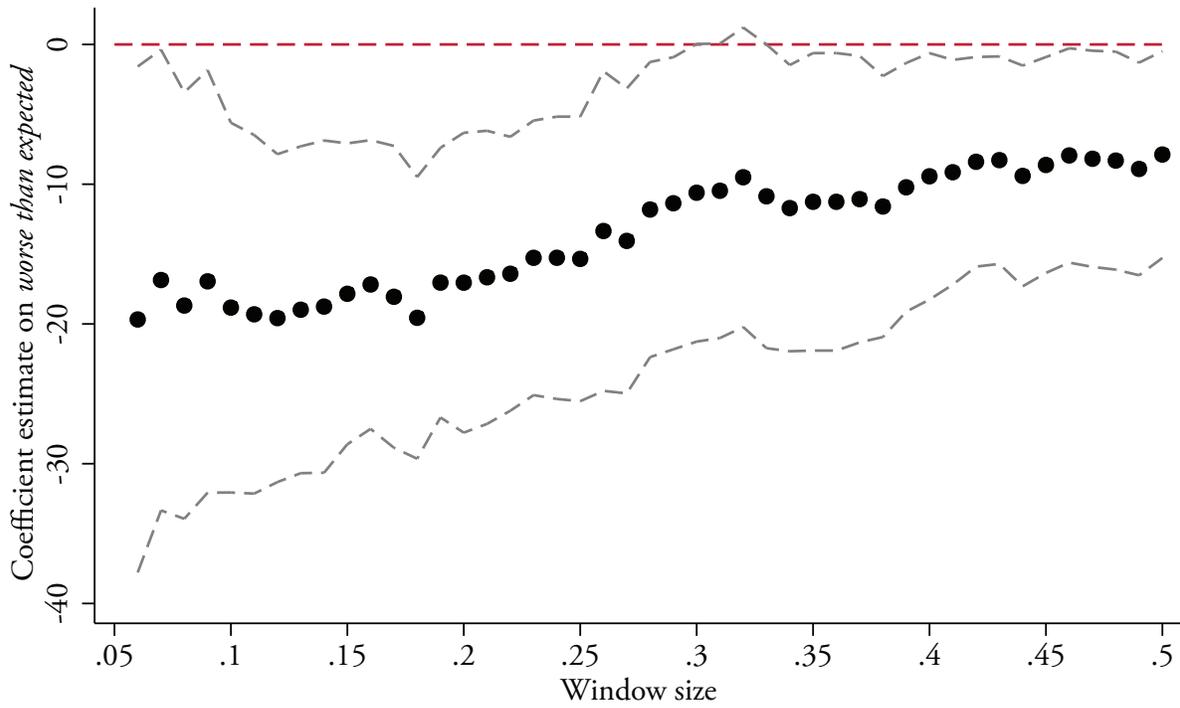
Note: This specification is based on a *worse than expected* dummy variable interacted an  $N^{\text{th}}$ -degree polynomial similar to the specifications in Table 4 with a .25 window size. Dashed lines indicated 95% confidence interval.

Figures A3: Robustness to Polynomial Degree for .5 Window Size



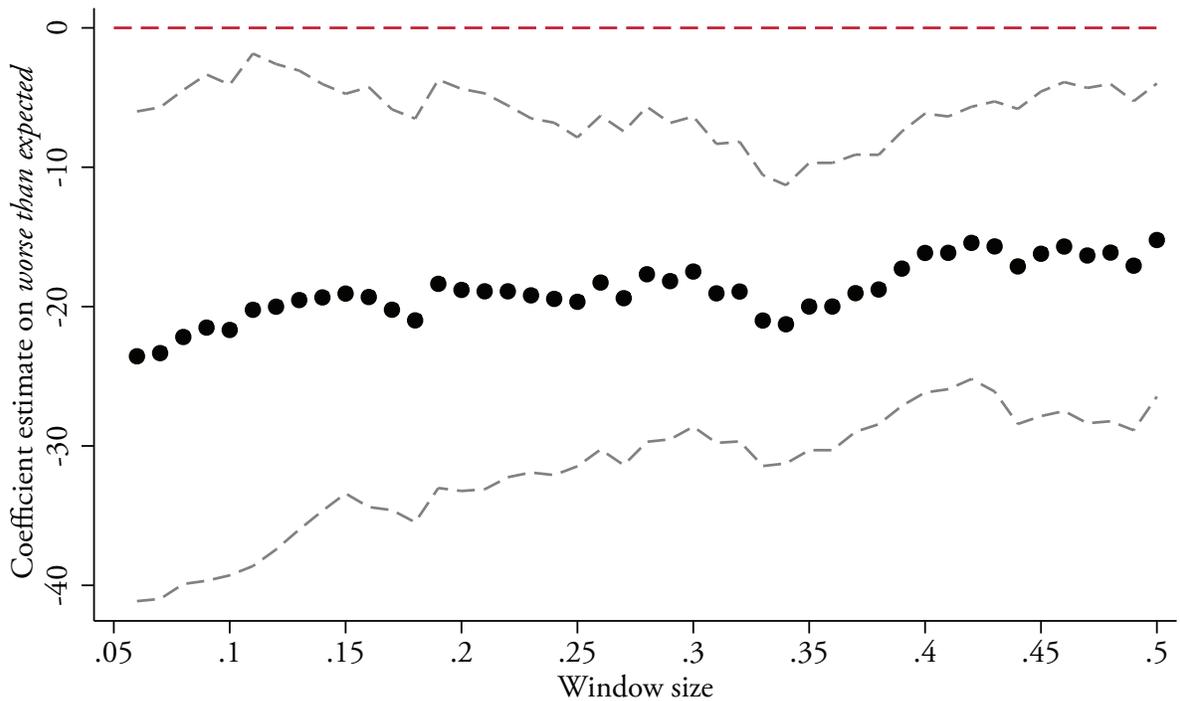
Note: This specification is based on a *worse than expected* dummy variable interacted with an  $N^{\text{th}}$ -degree polynomial similar to the specifications in table 4 with a .5 window size. Dashed lines indicated 95% confidence interval.

Figures A4: Robustness to Window Size for 3<sup>rd</sup>-Degree Polynomial



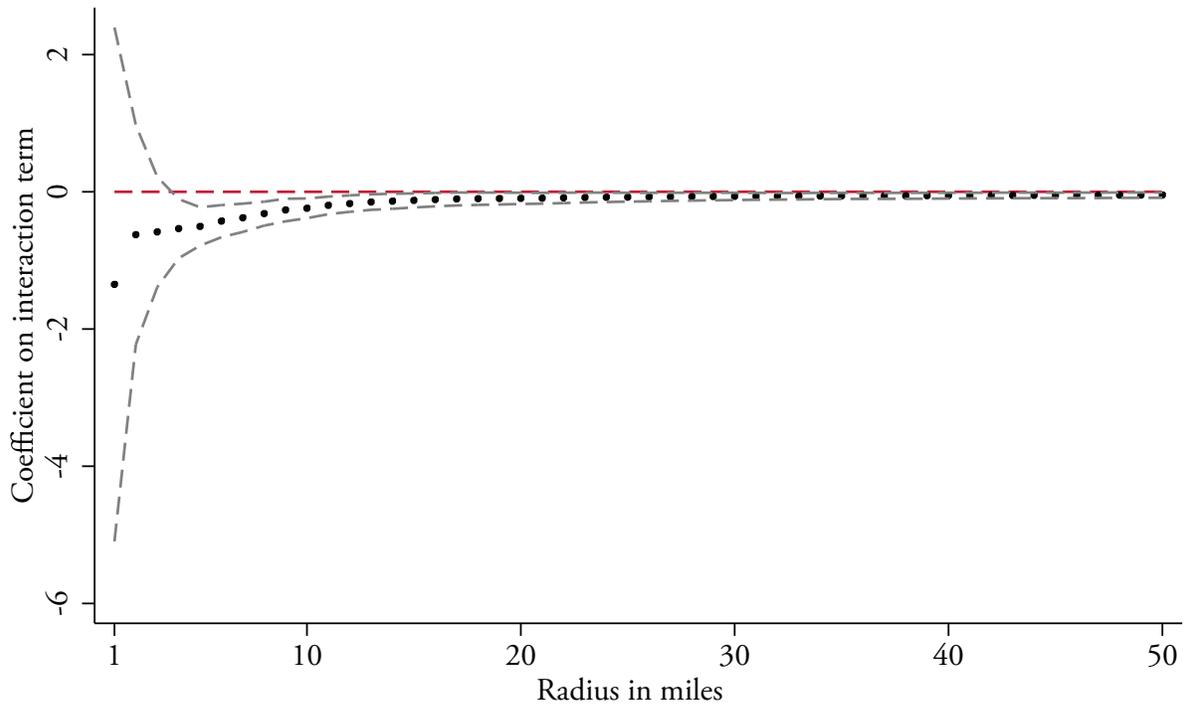
Note: This specification is based on a *worse than expected* dummy variable interacted a 3<sup>rd</sup>-degree polynomial similar to the specifications in Table 4. Dashed lines indicated 95% confidence interval.

Figures A5: Robustness to Window Size for 5<sup>th</sup>-Degree Polynomial



Note: This specification is based on a *worse than expected* dummy variable interacted with a 5<sup>th</sup>-degree polynomial similar to the specifications in Table 4. Dashed lines indicated 95% confidence interval.

Figures A6: Competition Result Robustness to Radius Choice



Note: This specification is based on Column 6 of Table 4. Each data point shows the interaction between the number of neighboring facilities and *worse than expected*. The x-axis represents the size in miles of the radius used to calculate the number of neighboring facilities. Dashed lines indicated 95% confidence interval.