Supply-side Responses to Public Quality Ratings: Evidence from Medicare Advantage

Ian McCarthy and Michael Darden

Abstract

Since 2009, the Centers for Medicare and Medicaid Services (CMS) has distributed quality star ratings for Medicare Advantage (MA) contracts to better inform Medicare eligible individuals of MA plan quality. This paper examines the extent to which the quality star rating scheme affected insurer behavior in the MA market. We identify the causal effect of 2009 star ratings on 2010 MA premiums using a regression discontinuity design that exploits plausibly random variation around threshold values underlying the star rating calculations. We find that contracts with higher star ratings in 2009 significantly increased their average 2010 monthly premiums relative to contracts just below the respective threshold values, with increases of more than $26 per month among 3.5 and 4-star contracts. We present additional evidence that the estimated premium increases in 2010 are not purely the result of demand shifts in the prior year. Furthermore, the 2009 star ratings have a significant effect on plan mix, where 3-star contracts disproportionately dropped their $0 premium plans while 2.5, 3.5, and 4-star contracts disproportionately expanded plans into new markets, particularly plans with positive monthly premiums. Broadly, our analysis reveals a relatively large supply-side response to published quality ratings, suggesting some level of caution as policymakers extend quality rating systems into new healthcare markets.

JEL Classification: C51; D21; D43; I11

Keywords: Quality Information, Medicare Advantage, Health Insurance, Insurance Premiums
1 Introduction

Imperfect information regarding the quality of a health insurance plan may justify policies that better inform consumers. Policies mandating quality reporting may also generate strong incentives for private health insurance firms to improve (reported) quality. However, it is not obvious that mandating product transparency is welfare improving. For example, Dranove & Satterthwaite (1992) demonstrate that consumer learning with respect to prices may generate lower prices, but, depending on the strength of competition in the market, that the resulting price-cost margin reduction may cause firms to reduce quality. Indeed, in the context of health insurance, the dissemination of information regarding price or quality presents a signal of information to both consumers and producers, which may have unintended consequences with respect to premiums, benefits, or provider networks (Dranove & Jin, 2010; Werner & Asch, 2005). The extent to which public policy and improved plan information can influence health insurance decision-making is a central question in health and public economics.\(^1\) We examine this issue using evidence from the recent star rating system introduced for Medicare Advantage (MA) contracts.

In 2009, the Centers for Medicare and Medicaid Services (CMS) introduced its overall star rating system in which each MA contract received a quality star rating on a five-star scale.\(^2\) The literature on the MA quality rating initiatives has generally focused on the enrollment effects. To this end, Reid et al. (2013) find large effects of increases in star-ratings on enrollment (particularly among new beneficiaries or switchers), but the authors do not distinguish the effects of true quality from reported quality. Darden & McCarthy (2015) attempt to disentangle true versus reported quality and show that the dissemination of 2009 quality star ratings had a relatively small but significant effect on enrollment, with Medicare eligible individuals enrolling in higher quality contract/plans. The estimated effects are positive, but they are smaller than Reid et al. (2013) and heterogeneous across the quality distribution.

Given evidence that Medicare enrollees responded to the 2009 dissemination of MA quality star ratings by enrolling in higher quality contract/plans, it is natural to expect 2010 contract/plan premiums to increase in higher rated contract/plans and decrease in lower rated contract/plans. On the other hand, revealing quality to consumers may further differentiate contract/plans such that the degree of substitution between contract/plans decreases and premiums increase across the quality spectrum. In this paper, we conduct a comprehensive analysis of 2010 premium adjustments to the 2009 publication of MA quality stars. We examine the population of MA contract/plans in 2009 and 2010 using publicly available data on contract/plan market shares, reported contract quality, plan premiums, and other plan characteristics. Using a novel regression discontinuity (RD) design that exploits plausibly

\(^1\)For example, see Pauly et al. (2014) on the effects of the Affordable Care Act on individual insurance premiums.

\(^2\)An MA contract is a private organization that administers potentially many differentiated plans. All plans within a contract received the same MA star rating. Throughout, we use the term “contract” to refer to a private health insurance firm that contracts with CMS to provide Medicare services through Medicare Advantage.
random variation around 2009 star thresholds, we separately identify the effect of reported 2009 quality on observed 2010 premiums from the overall relationship between quality and price. Our contract-level analysis suggests that contracts just above the 2009 thresholds for 2.5, 3, 3.5 and 4 stars increased their 2010 average monthly premiums by $12.82, $16.25, $28.58, and $26.97, respectively, relative to contracts with 2009 ratings just below the respective threshold values. Among 3, 3.5, and 4-star contracts, the premium increases are most concentrated in the upper end of the premium distribution, while premium increases among 2.5-star contracts occur in the lower end of the premium distribution. These results are broadly consistent across a series of robustness checks, including consideration of alternative bandwidths and falsification tests with counter-factual threshold values.

While the star ratings are calculated at the MA contract level, each MA contract administers potentially many differentiated MA plans. This means that all of the plans operating under the same MA contract will receive the same star rating. To better understand our contract-level results, we re-estimate our premium model at the contract/plan-level for just those contract/plans that existed in both 2009 and 2010. We find mean premium increases of $20.09, $35.80, and $30.82 for 3, 3.5, and 4-star contract/plans, respectively. These results suggest that higher rated contracts, on average, raised the premiums of their existing plans for 2010. While increases in premiums for highly rated contract/plans suggest either enrollment or production differentiation explanations, the significant increase in premiums for low-rated contract/plans suggests that the publication of quality stars further differentiated the MA market. Indeed, from a supplemental analysis in which we examine only contracts with relatively small enrollment changes from 2008 to 2009, we find that our estimated premium increases do not appear to be driven entirely by demand shifts following the 2009 quality ratings.

In addition to directly adjusting its plans’ premiums, an MA contract may also adjust the mix of plans it offers within a market (county). For example, in response to the published star ratings, a contract could alter the number of zero-premium plans; adjust the number of plans that include Medicare Part D coverage; change the drug deductible in plans that offer part D coverage; or add/drop plans entirely. Indeed, our data show that nearly all of the regional variation in plan premiums is due to selection of plan offerings by contracts, as opposed to contracts charging different premiums in different areas of the country. We show that contracts just above the 3-star threshold in 2009 were more likely to drop $0 premium plans in 2010. In addition, 3.5 and 4-star contracts were more likely to introduce positive premium plans into new markets. Meanwhile, low quality contracts (those just above the 2.5-star threshold in 2009) maintained or expanded their 2009 plan offerings with relatively small changes to premiums in 2010, while contracts just below the 2.5-star threshold in 2009 were much more likely to exit the market altogether in 2010.

While Dranove & Jin (2010) survey a large literature on the supply-side responses to quality dis-
semination mandates, until recently, the potential for supply-side responses to MA policy has received little attention from researchers. For example, Stockley et al. (2014) examine how MA plan premiums and benefits respond to variation in the benchmark payment rate – the subsidy received by an MA contract for each enrollee. Those authors find that contracts do not adjust premiums directly as a result of changes in benchmark payment rates, instead adjusting the generosity of plan benefits. Perhaps more closely related to our paper, Hirth & Huang (2016) show that the publication of quality star ratings of nursing home facilities caused highly rated nursing homes to raise their prices by over $3. Finally, as an examination of differential enrollment effects from 2009 to 2010, Darden & McCarthy (2015) present initial evidence that contract/plans in 2010 raise premiums in response to higher 2009 contract-level quality star ratings. However, the supply-side analysis in that paper was largely preliminary and based on a specific subset of contracts in 2010. Furthermore, that paper focused on direct premium increases at the plan level, ignoring the overall effect at the contract level as well as the possibility of indirect premium adjustments such as changing the number of zero-premium plans or expanding plans into new counties.

The welfare implications of our findings depend on a complicated and dynamic structural model of MA plan choice. As discussed above, an MA contract has many choices regarding premiums and plan mix, and assessing a Medicare eligible individual’s willingness-to-pay for quality, prescription drug coverage, and provider network is beyond the scope of our data. Furthermore, we would not expect to see large premium adjustments between 2009 and 2010 if the publication of star ratings in 2009 did not provide new information on quality; however, gauging an MA contract’s prior beliefs regarding own and competing contract quality ratings is difficult.3 For these reasons, our central goal in this paper is to provide credible estimates of the casual effect of the 2009 star ratings on 2010 premiums.

While the welfare aspects of our findings are beyond the scope of this paper, our contract-level results are sizable compared to the observed average premium increases of $9 to $15 from 2009 and 2010. Our analysis therefore reveals a relatively large supply-side response to published quality ratings, suggesting some level of caution as policymakers extend quality rating systems into new healthcare markets. For example, the Affordable Care Act mandates that Department of Health and Human Services devise a quality rating system for qualified health plans that operate in the health exchanges, and both premium and plan stability are central open questions regarding health reform.

The paper proceeds as follows. We discuss the institutional details of Medicare Advantage and the recent star rating program in Section 2, along with a conceptual framework for supply-side responses in Section 3. The data and methods are discussed in Section 4, with results in Section 5. Section

3Recall that 2009 was the first year that the quality star rating metric was published. From the enrollee perspective, Dafny (2009) show a 2001 publication of Medicare + Choice (the precursor to MA) quality information shifted enrollment but not substantially relative to empirical proxies for an individual’s information set.
6 examines the potential mechanisms underlying our estimated premium adjustments, Section ?? examines the role of competition in premium responses, and the final section concludes.

2 Institutional Background

Medicare beneficiaries have historically had the option to receive Medicare benefits through traditional Medicare fee-for-service (FFS) or instead through private health insurance plans (sometimes referred to as Medicare replacement plans). The Balanced Budget Act of 1997 (BBA) classified all private Medicare health insurance plans as Medicare Part C plans (Medicare + Choice), and it allowed for additional types of business models including Preferred Provider Organizations (PPOs), Provider-Sponsored Organizations (PSOs), Private fee-for-service (PFFS) plans, and Medical Savings Accounts (MSAs). Later, in addition to the beneficiary entitlement to prescription drug coverage, the Medicare Modernization Act of 2003 renamed the Medicare + Choice program with Medicare Advantage (MA) and introduced several additional reforms into the Medicare private health insurance market.

Medicare beneficiaries can now choose to enroll in traditional Medicare FFS or a private MA plan during an open enrollment period from November 1st through December 31st. By enrolling in an MA plan, enrollees must pay Medicare Part B premiums in addition to any additional premium charged by the plan. In exchange, MA plans provide at least (often more than) the services covered by traditional FFS Medicare. In 2009, 38% of MA plans charged no additional premium, while 77% of plans also offered prescription drug coverage. Given the generosity of plan coverage at possibly no additional cost relative to traditional Medicare FFS, the MA program has grown dramatically in recent years with the share of Medicare eligible individuals in an MA plan increasing from 13.7% in 2003 to 30% in 2014.4

Broadly, an MA contract is an agreement between a private insurance company and CMS whereby the company agrees to insure Medicare beneficiaries in exchange for reimbursement. A contract is approved by CMS to operate in specific counties, and an approved contract typically offers a menu of MA plans that are differentiated by premium, prescription drug coverage, and, if covered, the prescription drug deductible. Most MA contracts are required to offer at least one plan that includes prescription drug coverage. For the 2015 enrollment year, 78% of all Medicare beneficiaries lived in a county with access to at least one plan that offers prescription drug coverage (MA-PD) and charges no additional premium (above the Part B premium).5

Newhouse et al. (2015) neatly summarizes the relationship between MA contract/plans and CMS.

---

Since 2006, each MA insurer submits a bid that represents the per enrollee reimbursement from CMS. The actual amount reimbursed to the insurer depends on the relationship of the bid to the benchmark rate, a measure of the average cost per enrollee in traditional medicare adjusted for area level characteristics. If the bid is less than the benchmark, as nearly all bids are, then CMS and the insurer split the difference, and the additional reimbursement that goes to the insurer is required to be used in for additional services and benefits to the enrollee. In the case that the bid exceeds the benchmark rate, enrollees pay the difference in plan premiums. In exchange for reimbursement, the insurer is required to provide the same (or more) coverage that would be available in traditional FFS Medicare. From the perspective of a Medicare eligible individual, MA plans often offer more complete coverage (eliminating the need to purchase supplemental Medigap coverage), lower cost sharing at the time of care, and a greater variety of covered services. On the other hand, MA plans restrict covered services to providers within their network, whereas the vast majority of providers accept traditional FFS Medicare.

Since the 1990s, CMS has been concerned with informing the Medicare public of MA quality. Staring in the 2007 enrollment year, CMS began collecting and distributing a one to five-star quality rating in each of five quality domains (e.g., “Helping You Stay Healthy”). Each domain was itself an aggregation of many individual quality metrics such as the percentage of enrollees with access to an annual flu vaccine. These individual quality metrics are calculated based on data from a variety of sources, including the Healthcare Effectiveness Data and Information Set (HEDIS), the Consumer Assessment of Healthcare Providers and Systems (CAHPS), the Health Outcomes Survey (HOS), the Independent Review Entity (IRE), the Complaints Tracking Module (CTM), and CMS administrative data. Starting in enrollment year 2009, CMS began aggregating the domain level quality stars to an overall contract rating of between one and five stars (in half-star increments). And since 2011, CMS constructs the contract-specific quality ratings as a function of Part D coverage, when relevant. Our focus is on the 2009 and 2010 enrollment years – the first two years of the overall contract star rating program and the years in which all contracts, including those offering prescription drug coverage, were rated based on the same underlying quality metrics.

Dranove & Jin (2010) summarize the large literature on supply-side responses to quality disclosure (both voluntary or mandatory), much of which focuses on the extent to which quality disclosure causes firms to improve quality. For example, Chen (2008) finds that the Nursing Home Quality Initiative, which allows consumers to compare the quality of different nursing homes via mandated

---

6 For a complete discussion of the star rating program, see Appendices A and B along with the discussion in Darden & McCarthy (2015).

7 A large literature exists on the economics of voluntary quality disclosure. Under the “unraveling” hypothesis, voluntary disclosure occurs because the highest quality firm has an incentive to promote their quality. This incentive exists for all firms that have yet to disclose, and it results in all firms voluntarily disclosing. In our context, disclosure is mandatory because CMS calculates and publishes MA contract quality star ratings.
quality disclosure, caused low quality nursing homes to improve quality. On the other hand, Werner & Asch (2005) argue that quality reporting may actually reduce quality if “quality” targets do not align with patient outcomes, and Bar-Isaac et al. (2012) show that when some dimensions of quality are reported and not others, firms may end up reducing “overall” quality by investing exclusively in improving reported quality. A systematic review led Fung et al. (2008) to conclude that this literature sends many “mixed signals” as to whether quality disclosure actually improves quality.

Our goal is to empirically identify the extent to which the 2009 publication of quality star ratings caused MA contracts to change premium and product offering decisions. Hirth & Huang (2016) ask a similar question with respect to the December 2008 CMS publication of nursing home quality stars. While they find that nursing home prices increased overall, prices increased more for highly rated nursing homes, and most of the price increases are concentrated in highly competitive nursing home markets. In our setting in 2009, the mean number of MA plans available to beneficiaries was roughly 11 plans per county. However, there exists considerable regional variation in the availability of MA plans, and enrollments in MA plans are concentrated in a few national contracts. Indeed, according to the Kaiser Family Foundation (KFF), 60% of all plans offered in 2015 are affiliated with just seven health insurance companies. In what follows, we present a conceptual framework for the exogenous revelation of quality information in a relatively competitive market.

### 3 Conceptual Framework for Premium Changes

In this paper, we analyze the effect of quality star ratings on MA firm behavior, specifically the extent to which the star ratings caused firms to change their pricing behavior. Following Dranove & Satterthwaite (1992), we lay out a simple economic framework that sheds light on pricing decisions when consumer beliefs regarding quality become more precise. Finally, we discuss one other plausible mechanism that may explain any observed increase in premiums.

Consider a monopolistically competitive market in which each MA contract offers a set of differentiated health insurance products and, because of free entry/exit, each contract makes zero economic profit. Consumers in this market may choose any of the plans offered by any contract, or they may select traditional fee-for-service Medicare, which represents an outside option. For ease of exposition, assume that contract/plans are only differentiated on the basis of premium and quality. Consumers have prior beliefs regarding contract/plan quality, which could be a function of past experience, word of mouth, or other private sources of quality information (e.g., U.S. News and World Report). The

---

8Author’s calculation. See Section 4 for a presentation of our data.
10Dranove & Satterthwaite (1992) also assume that price is uncertain from the consumer perspective. In their model,
exogenous revelation of quality (through the star ratings) serves to reduce the posterior variance of quality beliefs of consumers.

The reduction in the posterior variance of consumer quality beliefs has two main effects on the market. First, two contract/plans that were previously undifferentiated may become differentiated depending on the granularity of both the prior beliefs and the signal of quality information. For example, two contract/plans that were previously viewed as “good” may now be viewed as 4 and 3.5 star contract/plans. Thus, the products become increasingly differentiated. Importantly, increasing the degree of product differentiation decreases the degree of substitution between plans, and premiums for all contract/plans increase. Second, Dranove & Satterthwaite (1992) show that a decrease in consumers’ quality belief variance increases consumer sensitivity to quality, and thus increases the elasticity of demand with respect to quality. The increased elasticity suggests that quality disclosure will cause consumers to forgo low-quality contract/plans for high quality contract/plans. The increase in demand for high quality contract/plans would increase high-quality premiums while lowering the demand for low-quality contract/plans and the associated premiums. Thus, in both cases above, high-quality premiums would be expected to rise, whereas low-quality premiums may increase or decrease.

Health insurance makes for an interesting application of the above framework because, as enrollments in contract/plans change, the resulting premium changes depend on the marginal cost curve, whose relevant shape depends upon who is enrolling. Indeed, Newhouse et al. (2015) explain that, because Medicare requires that contract/plans accept all beneficiaries, if, for example, sicker individuals select into higher quality plans because of the decrease in the variance of quality beliefs, then the marginal cost curves of high rated plans may be steeper than low rated plans. In this case, we would expect to see larger increases in contract/plan premiums for higher rated contract/plans. Selection on the basis of health may also generate incentives for both types of firms to increase premiums to discourage high cost enrollees.

Another explanation for an enrollment effect due to the star ratings may stem from the fact that most of the quality star rating is based on provider characteristics and quality. Indeed, from the consumer perspective, the main drawback of MA contract/plans is that MA restricts enrollees to a given provider network. If the quality star rating disclosure signals that a given provider network is

---

11 The main instruments with which a plan can improve its risk pool is through their choice of provider networks, covered benefits, and choice of geographic areas in which to operate.

12 Of course, it possible that healthier individuals could select into higher rates contract/plans, but, as with the standard adverse selection argument, individuals who anticipate more interaction with the medical system are likely to seek out higher quality care. In the case that healthier individuals select into higher rated contract/plans, an increase in enrollment could be associated with a decrease in premiums if the marginal cost curve were steeply downward sloping.
high quality, then the high quality contract/plan enrollment increase may come from Medicare eligible
individuals previously in traditional FFS Medicare, thus limiting the enrollment effect on low quality
contract/plans.

Finally, as an alternative to the enrollment and product differentiation mechanisms above, if quality
stars mainly reflect provider quality, the disclosure of quality stars may increase in bargaining power
of providers when negotiating reimbursement from the MA contract/plan. For example, hospitals and
physicians may use their high quality scores (and subsequent high star rating among MA contracts) as
a justification for higher prices, which are then passed on to the Medicare beneficiary. This mechanism
would operate through the marginal cost curve.

We study the premium response in 2010 to the 2009 publication of MA quality stars. That is, we
model the premium decisions of MA contract/plans in 2010 as a function of 2009 star ratings. Darden
& McCarthy (2015) and Reid et al. (2013) demonstrate that 2009 enrollment shifted in response to
2009 quality star rankings. In 2010, to summarize our conceptual framework, we expect to observe
increases in premiums for highly rated contract/plans because of both enrollment gains and product
differentiation arguments.

4 Data and Methods

We collect data on market shares, contract/plan characteristics, and market area characteristics from
several publicly available sources for calendar years 2009 and 2010.\footnote{See Appendix B for a detailed discussion of our dataset and specific links.} We use the Medicare Service Area
files to form a census of MA contracts that were approved to operate in each county in the United States
in 2009 and 2010. To these contract/county/year observations, we merge contract/plan/county/year
data on enrollment and other contract characteristics.\footnote{The Service Area files are needed because the enrollment files do not account for migration. For example, it is possible for the enrollment file to contain a positive enrollment record for a contract/plan in a county even if that contract is not approved to operate in the county. See Appendix B for further details.} To our market share data, we merge further
information on MA contract quality ratings, contract/plan premiums, county-level MA market share,
and census data. The CMS quality information includes an overall summary star measure; star ratings
for different domains of quality (e.g., helping you stay healthy); as well as star ratings and continuous
summary scores for each individual metric (e.g., percentage of women receiving breast cancer screening
and an associated star rating). Data are not available for the overall continuous summary score
(i.e., the score rounded to generate an overall star rating), but we are able to replicate this variable
by aggregating the specific quality measures following CMS instructions. We explain this process
thoroughly in Appendix B.
Our enrollment data are available monthly; however, there is little variation in enrollments across months due to the nature of the open enrollment process at the end of each calendar year. Furthermore, plan premiums are specific to a calendar year. We therefore measure enrollments based on the average enrollment of each plan across months in a given year. The resulting unit of observation is the contract/plan/county/year. Our analysis focuses only on health maintenance organizations (HMO), local and regional preferred provider organizations (PPO), and private fee-for-service (PFFS) contracts. We exclude all special needs plans and employer/union-specific plans (also known as 800-series plans), and we drop all observations that pertain to United States Territories and Outlying Areas. Our final sample includes 247,978 contract/plan/county/years.

Table 1 provides summary statistics for our final dataset at the plan, county, and contract level. The data consist of 51,442 and 34,642 plan/county observations in 2009 and 2010, respectively, with an increase in average MA enrollment per plan from 292 in 2009 to 361 in 2010.15 The county-level summary statistics also reveal an increasing penetration of MA in the overall Medicare market, from 15.6% in 2009 to 16.5% in 2010, alongside a decrease in the number of plans offered per county, an increase of just over $15 in average premiums, an increase in the percentage of plans offering prescription drug coverage, and an increase in the proportion of HMO and PPO plans relative to PFFS plans. Finally, the bottom panel of Table 1 illustrates a slight rightward shift in the distribution of star ratings from 2009 to 2010, with 1.5-star contracts either improving in rating in 2010 or exiting the market, and with a relative increase in the percentage of 4.5 and 5-star contracts in 2010.

Since star ratings are assigned to contracts rather than specific plans, our initial analysis follows Town & Liu (2003), Cawley et al. (2005), Dafny & Dranove (2008), Frakt et al. (2012) and others in aggregating plan characteristics to the contract/county level. We consider several different outcomes, including the change in average premium across plans within a contract, the change in the median premium, the change in the 25th and 75th percentile premiums, as well as the change in the percentage of $0-premium plans in a contract.16 We then examine the relationship between a contract’s quality star rating in 2009 and the contract/county outcomes in 2010.

The CMS quality rating system relies on a continuous summary score between 1 and 5 which is rounded to the nearest half. A contract with a 2.24 summary score is therefore rounded down to a

15As indicated in Table 1, enrollment data are not available for all plans as CMS does not provide enrollment counts for plans with 10 or fewer enrollments. As such, the mean enrollment figures presented are higher than the true mean as they exclude a large number of plans with missing enrollment data.

16CMS suppresses enrollment counts for contract/plans with 10 or fewer enrollees. Consistent with CMS reporting criteria, we exclude plans with 10 or fewer enrollees when examining contract and plan premiums, as the premiums for such small enrollment plans are not generally representative of the premium for an average enrollee.
2-star rating, while a contract with a 2.26 summary score is rounded up to a 2.5-star rating. Intuitively, these two contracts are essentially identical in quality but received different quality ratings. We propose to exploit the nature of this rating system using a regression discontinuity (RD) design.\textsuperscript{17} Following Imbens & Lemieux (2008) and Lee & Lemieux (2010), we present initial evidence supporting our RD design in Figure 1, where we graph the average change in premiums for different values of the continuous summary score as well as a 4th order polynomial regression line estimated separately above and below the relevant threshold value.\textsuperscript{18} The observed discontinuities in Figure 1 tend to support our proposed RD design.

\textbf{Figure 1}

To formally estimate the effect of reported quality in 2009 on premiums in 2010, we first denote by $R_c$ the underlying summary score, by $\hat{R}$ the threshold summary score at which a new star rating is achieved (e.g., $\hat{R} = 2.25$ when considering the 2.5 star rating), and by $\tilde{R}_c = R_c - \hat{R}$ the difference between the underlying summary score and the threshold value. Denoting the outcome of interest for contract $c$ in market $m$ by $\bar{y}_{cm}$, we are interested in estimating the following pooled regression equation:

$$\bar{y}_{cm} = \gamma_1 + \gamma_2 \times I(R_c > \hat{R}) + \gamma_3 \times \tilde{R}_c + \gamma_4 \times I(R_c > \hat{R}) \times \tilde{R}_c + \varepsilon_{cm}, \text{ for } R_c \in [\hat{R} - h, \hat{R} + h],$$  

where $\gamma_2$ is the main parameter of interest. We estimate this pooled regression model using kernel-weighted local linear regression with a bandwidth of $h = 0.125$ and a triangular kernel (Imbens & Lemieux, 2008; Lee & Lemieux, 2010; Nichols, 2011).\textsuperscript{19} We examine the sensitivity of our results to the choice of bandwidth, along with other robustness checks, in subsection 5.3 and Appendix B.

Note that changes in aggregate contract characteristics such as average premiums can arise in several ways, most directly via changes to specific plans operating within a contract. To investigate this possibility, we also estimate a regression of changes in plan premiums from 2009 to 2010, again based on our RD design. This analysis is akin to estimating equation 1 but where our analysis is at the plan level rather than aggregating to the contract level. For this analysis, we examine only plans that were available in the same county in both 2009 and 2010.

\textsuperscript{17}See Imbens & Lemieux (2008) and Lee & Lemieux (2010) for a detailed discussion of the RD design and its application in economics.

\textsuperscript{18}Figure 1 focuses on the average premium across plans in the same contract/county (excluding plans with 10 or fewer enrollments). Graphs based on the median premium, the 25th percentile, and the 75th percentile are similar but excluded for brevity. All graphs are based on a bin width of 0.02.

\textsuperscript{19}Here, “pooled” refers to pooling observations on either side of the relevant threshold. Due to likely heterogeneities across the quality rating distribution, we do not pool the observations across different thresholds, instead estimating each regression separately for different threshold values.
5 Results

5.1 Contract-level Premiums

Table 2 presents the results of our nonparametric estimation of equation 1 with triangular kernel and bandwidth $h = 0.125$. Each panel in the table reflects a different outcome, with panel 1 presenting results for the average contract premium.\textsuperscript{20} To examine the broader distribution of premiums within a contract, panels 2, 3, and 4 present results for the 25th, 50th, and 75th percentiles of premiums. Here, our outcome is calculated as the change in the $p^{th}$ percentile from 2009 to 2010 among all plans offered within a given contract/county/year. Finally, panel 5 presents results for the percentage of plans charging a $0$ premium. In addition to the point estimates in Table 2, the results for average premiums are presented graphically in Figure 2. Figures for other outcomes are available in Appendix B.

Table 2 and Figure 2

The results reveal large average premium increases across all star ratings, particularly among contracts receiving a 3.5 or 4-star rating, with these contracts increasing average premiums by more than $26$ per month from their 2009 levels relative to contracts with one-half star lower ratings.\textsuperscript{21} The results also show a differential change across the premium distribution. For example, 2.5-star contracts just above the 2.25 threshold tend to increase premiums among the lower premium plans, with the 25\textsuperscript{th} and 50\textsuperscript{th} percentile premiums increasing by $11$ and $20$, respectively, compared to 2-star contracts just below the 2.25 threshold. These 2.5-star contracts similarly reduce the proportion of $0$ premium plans. Meanwhile, price increases among 3-star contracts relative to 2.5-star contracts are primarily concentrated among the higher priced plans as indicated by the estimated $38$ increase in the 75\textsuperscript{th} percentile of premiums. Finally, 3.5 and 4-star contracts exhibit increases across the premium distribution, with the largest increases at the 75\textsuperscript{th} percentile.\textsuperscript{22}

\textsuperscript{20}Since enrollments are also influenced by contract quality ratings, we do not present results for average premiums weighted by enrollments. Using 2009 enrollments as weights, the results are broadly similar to those based on the unweighted averages presented in panel 1 of Table 2; however, 2009 enrollments necessarily ignore new plans introduced from 2009 to 2010, which is a potentially important source of premium variation (discussed in Section 6). We therefore focus our results on the simple average premium.

\textsuperscript{21}Although the majority of counties have at least one PFFS plan available, there are only 8 PFFS contracts in our final data, the majority of which received a 2 or 2.5-star rating in 2009. Our results are therefore largely identified from the HMO and PPO contracts. If we re-estimate our effects focusing only on managed care contracts, we see no significant changes at the 2.5, 3.5, and 4-star rating thresholds; however, the effect among 3-star contracts more than doubles from an increase of $16$ to an increase of $38$. This suggests a differential premium response among managed care contracts versus PFFS, but due to the small number of PFFS contracts, we cannot examine these differences in sufficient detail. We thank an anonymous reviewer for highlighting this point.

\textsuperscript{22}Estimates around the 3.75 threshold are less precisely estimated due to the relatively small sample of 4-star contracts.
By virtue of the RD design and the nature of the CMS star rating program, we argue that these estimates can be interpreted as the causal effect of a one-half star increase in quality ratings separate from the quality of the contract itself. For example, 3.5-star contracts of comparable “true” quality to 3-star contracts were able to increase their premiums on average $29 per month. Looking purely at sample averages, all other contracts receiving a 3.5-star rating in 2009 increased their premiums by an average of $12, while 3-star contracts falling just below the 3.25 threshold increased their premiums by just over $3. The estimated increase of $29 is therefore more than 2 times the average increase in premiums among existing 3.5-star contracts and nearly 10 times the increase in average premiums among existing 3-star contracts.

5.2 Plan-level Premiums

Table 3 and Figure 3 present the results for changes in individual plan premiums from 2009 to 2010 as a function of the 2009 quality rating as in equation 1. This analysis therefore estimates premium changes at the plan level (for the same plans offered in both 2009 and 2010), rather than analyzing aggregate characteristics at the contract level as in Table 2. As with our analysis of average contract premiums, the plan-level results are based on a pooled regression model estimated using kernel-weighted local linear regression with a bandwidth of $h = 0.125$ and a triangular kernel.

Table 3 and Figure 3

For the same plan/county/contract, the results again show a large and statistically significant increase in premiums for plans operating under 3, 3.5, and 4-star contracts, with premiums increasing by between $20 and $36 per month for the same plans. Perhaps surprisingly, the results also show a small but statistically significant decrease in premiums among plans operating under 2.5-star contracts relative to 2-star contracts. Since this analysis is limited to plans offered in both 2009 and 2010, these results suggest that contracts are adjusting their average premiums in ways other than direct changes to existing plans (i.e., through adjustments in their plan mix). In particular, it may be that 2.5-star contracts must lower premiums following a disproportionate exit among 2-star contracts. We investigate this issue in more detail in Sections 6 and 7.

To better interpret these results, we calculate the total amount of additional premiums paid based on the estimates in Table 3. In 2010, there were nearly 396,000 MA beneficiaries enrolled in a plan that received a 3-star rating in 2009 and with a continuous summary score of between 2.75 and 2.875. All such beneficiaries therefore paid an estimated $20 in premiums for essentially no improvement in quality (relative to plans with a summary score of between 2.625 and 2.75). Similarly, nearly 96,000
beneficiaries were enrolled in a plan receiving a 3.5-star rating in 2009 and with a continuous summary score of between 3.25 and 3.375, while over 587,000 beneficiaries were enrolled in a plan receiving a 4-star rating in 2009 and with a continuous summary score of between 3.75 and 3.875. In total, this amounts to over $353 million in additional premium payments for plans with higher reported star ratings but similar underlying quality (based on the continuous summary scores).\textsuperscript{23}

5.3 Robustness and Sensitivity

In adopting the RD design, we have implicitly assumed that contracts are largely identical in quality just above versus below the star-rating thresholds (within a specific bandwidth of $h = 0.125$). We have also assumed that contracts cannot sufficiently manipulate their summary scores so as to obtain higher quality ratings. In Appendix B, we examine the sensitivity of our initial results to these assumptions. We also re-estimate our models with county-level covariates and rectangular (uniform) kernels.

These extensions generally support the validity of our RD design, with no evidence that quality metrics differ above versus below the relevant thresholds and no evidence that contracts sufficiently manipulate their continuous quality scores. Our results are also largely consistent across alternative bandwidths, inclusion of covariates, and alternative kernels. Finally, a falsification test of discontinuities around counter-factual threshold values reveals the largest premium increases to have occurred at or near the true threshold values underlying each star rating.

6 Mechanisms for Premium Adjustment

Comparing our contract-level (Table 2) and plan-level (Table 3) analysis, we see clear differences in premium changes at the plan level versus at the contract level. These results suggest that increases in average premiums at the contract level do not arise solely from increases in premiums of the same plans from 2009 to 2010. Rather, contracts may also alter their plan mix from one year to the next (e.g., dropping plans within a contract, introducing new plans under the same contract, or expanding plans to new counties).

Differential changes in plan mix by star rating is supported by the overall trends in plan entry and exit from 2009 to 2010. For example, over 99\% of 1.5-star plans leave the market from 2009 to 2010, with less than 10\% exit among 4 and 4.5-star plans.\textsuperscript{24} Regarding entry, 37\% of 1.5-star plans (in 2009) expanded into a new county in 2010, compared to 55\% of 2-star plans, while higher rated contracts were relatively less likely to enter into new markets. Collectively, the exit and entry figures reflect

\textsuperscript{23}This total still exceeds $350$ million when incorporating the premium reduction among 2.5-star plans.

\textsuperscript{24}The 1.5-star contracts that stayed in the market from 2009 to 2010 also had a marginally higher star rating in 2010. As such, there are no 1.5-star contracts remaining in 2010 (see Table 1).
larger turnover in plan offerings among lower rated contracts relative to higher rated contracts. This is perhaps expected as higher rated contracts may be more deliberate in their market entry/exit decisions and less likely to quickly cycle through new plans from one year to the next. In the remainder of this section, we examine more closely how contracts may change their plan mix in response to CMS star ratings, and we discuss how such plan mix adjustments may ultimately influence a contract’s average premium at the county level.

6.1 Analysis of Plan Exit

To examine plan exit more directly, we follow Bresnahan & Reiss (1991), Cawley et al. (2005), Abraham et al. (2007), and others in assuming that an insurance company will only offer a plan in a given county if the plan positively contributes to the contract’s profit. Assuming profit is additively separable across geographic markets (counties), our observed plan choice indicator becomes:

\[ y_{c(j)m} = \begin{cases} 
1 & \text{if } \pi_{c(j)m} \geq 0 \\
0 & \text{if } \pi_{c(j)m} < 0 
\end{cases}. \] (2)

Again exploiting the RD design, we analyze plan exit with a kernel-weighted local linear regression model using a triangular kernel and a bandwidth of 0.125. In this case, \( y_{c(j)m} = 1 \) indicates that the contract continued to offer the plan in 2010 and \( y_{c(j)m} = 0 \) indicates the plan was dropped. By definition, this analysis is based on existing plans as of 2009.

The results of our analysis of plan exit are summarized in the top section of Table 4. The first panel presents results for all plans, while the remaining panels present results for plans with $0 premiums and plans with positive premiums, respectively. Overall, we see that 2.5-star contracts are significantly less likely to exit markets than 2-star contracts of similar overall quality. Relative to 2.5-star contracts, 3-star contracts are less likely to exit overall (although small in magnitude), but they are significantly more likely to drop their $0 premium plans and less likely to drop positive premium plans. Somewhat surprisingly, contracts receiving a 3.5-star rating are more likely to drop plans overall; however, from the middle panel of Table 4, we see that this result is largely driven by 3.5-star contracts dropping their $0 premium plans. Finally, 4-star contracts are significantly less likely to exit overall.\(^{25}\)

\(^{25}\)The robustness of our plan exit results to bandwidth selection is presented in Appendix B. For sufficiently wide bandwidths, the results for 2.5, 3.5, and 4-star contracts are consistent across a range of bandwidths. Meanwhile, the overall results (top panel of Table 4) at the 2.75 threshold appear relatively sensitive to bandwidth selection, with the statistical significance, magnitude, and sign of the point estimates changing within bandwidths from 0.1 to 0.2. In terms of hypothesis testing, we interpret this as evidence in favor of the null that the star rating has no effect on plan exit at the 2.75 threshold. Given the small magnitude of the overall effect for these contracts in Table 4, the sensitivity to bandwidth selection is not surprising in this case.
6.2 Analysis of Plan Entry

An important and relatively unique aspect of the MA market concerns the distinction between plan and contract-level decisions. Specifically, contracts must obtain CMS approval in order to be offered in a given county; however, conditional on receiving CMS approval, the decision of which plan(s) to offer in a county is relatively less regulated. As a result, we argue that the fixed costs of entry are primarily incurred at the contract level while the plan-level entry/exit decisions are based on the variable profits per enrollee (i.e., regardless of market share).

The full set of plans available to a contract in a given market \( m \) is identified by taking all plans offered under that contract across the entire state in the same year. All such plans are therefore considered “eligible” to be operated in any given county, and the contract must choose which of those plans to offer in each county, where \( y_{c(j)m} = 1 \) indicates that the plan was added to the county (under that contract) in 2010, and \( y_{c(j)m} = 0 \) indicates that the plan was not offered. As with our analysis of plan exit, we estimate the entry-equivalent to equation 2 using kernel-weighted local linear regression, with entry considered as a function of 2009 contract quality as in equation 1.

The bottom section of Table 4 summarizes the results of our analysis of plan entry based on the RD design. Note that these results only apply to markets in which the contracts previously operated (i.e., we do not consider the contract-level entry decisions and instead focus specifically on the plan-level entry of pre-existing contracts). The RD results indicate that a one-half star improvement for 3.5 or 4-star contracts makes them significantly more likely to expand their plans into new markets. The bottom panel of Table 4 further reveals that the increase in probability of plan entry among 3.5 versus 3-star contracts occurs among positive premium plans. We also estimate that 2.5-star contracts are more likely to expand plan offerings relative to 2-star contracts, again with the largest effect among positive premium plans. Finally, 3-star contracts are less likely to expand their plan offerings relative to 2.5-star contracts of similar overall quality.\(^{26}\)

6.3 Role of Competition

An insurer’s change in premiums intuitively depends on the competitiveness of the local MA market; however, the direction of this relationship in a differentiated product market is not clear \textit{ex ante}. For

\(^{26}\)The robustness of our plan entry results to bandwidth selection is summarized in Appendix B. Similar to our analysis of plan exit, the results at the 2.25, 3.25, and 3.75 threshold are generally consistent across a range of alternative bandwidths. The results at the 2.75 threshold are more sensitive to bandwidth selection, with insignificant estimates for most bandwidth choices.
example, firms operating in more competitive markets may benefit more from quality reporting as this may soften price competition and allow for premium increases, whereas firms operating in less competitive markets are less influenced by quality reporting. In this section, we examine potential differential responses in premium changes across markets with different levels of competitiveness.

To do so, we extend our RD design to sub-samples of markets with different levels of competition as measured by the Herfindahl-Hirshman Index (HHI) at the contract-level. Results are presented in Figure 4, where the HHI percentile is presented along the $x$-axis in each sub-figure. The left-most portion of each sub-figure therefore reflects the estimated effect of quality ratings on plan premiums among the most competitive markets, and moving along the $x$-axis incorporates increasingly concentrated markets into the estimation.

**Figure 4**

The results reveal two interesting findings: 1) premium changes in response to quality reporting depends in-part on the level of competition in the market; and 2) these effects vary by quality rating. For example, 3-star contracts just above the 2.75 threshold do not significantly increase premiums in sufficiently competitive markets. It is only when these contracts operate in relatively less competitive markets where they can adjust premiums in response to their 3-star rating (relative to 2.5-star plans with similar underlying quality). However, the effects of competition are largely unchanged beyond the 50th percentile of the HHI distribution.

Meanwhile, 4-star contracts increase premiums the most in more competitive markets, with decreasing effects of star ratings on premiums as we include more concentrated markets in the analysis. This result is consistent with a differentiated product market in which the highest quality firms only benefit from such a high quality if they are faced with some level of competition from other insurers.\(^\text{27}\) If instead competition is relatively weak, then there is less of an effect from being reported as higher quality. The results for 3.5-star contracts then reflect a blend of the results among 4-star and 3-star contracts. We also find evidence that the effects of competition are largest among the smallest market share plans, with little variation in effects among plans with market shares above the median share. Graphical results for this analysis are provided in Appendix B.

\(^\text{27}\)Indeed, receiving a 4-star rating in 2009 was essentially equivalent to being labeled as the highest quality in your market, where 87% of the 4-star plans operating in 2009 were also the highest quality plans in their market.
7 Discussion

The potential supply-side response of MA contracts to the CMS quality rating system is critical both from a policy perspective as well as a consumer welfare perspective. If contracts can take advantage of improved quality scores by increasing premiums (holding the contract’s true quality constant), then this suggests a lack of competitiveness in the MA market with contracts raising prices without any true improvement in quality. Building on the preliminary results of Darden & McCarthy (2015), the current paper finds strong evidence of such premium increases among average to above average star-rated contracts.

Our analysis suggests that the estimated increases in premiums are not due to chance but are instead reflective of a true increase in premiums following an increase in reported quality. When we further compare the mean contract-level premium changes to the plan-level analysis, we see that 2.5-star contracts tend to increase average premiums by introducing new plans in existing markets, particularly new plans with positive monthly premiums. Conversely, changes in average premiums among higher rated contracts are achieved more through direct premium adjustments rather than changes in plan mix.

At least some portion of the estimated premium increases can be explained by a change in enrollments induced by the rating system (Reid et al., 2013; Darden & McCarthy, 2015); however, there is evidence that enrollment changes do not explain all of the premium response. For example, even among plans with relatively small enrollment changes in 2009, we still estimate a $47 increase in premiums among 3-star relative to 2.5-star plans and a $30 increase in premiums among 3.5-star relative to 3-star plans.28 An increase in product differentiation may explain the observed increase in premiums when enrollment does not change.29 Our analysis of plan entry and exit further suggests that insurers are not responding purely to enrollment changes. Specifically, outward demand shifts would only be expected to increase the premium of the same plan. Evidence that insurers drop and add some plans in response to their quality ratings (particularly since quality ratings are specific to the contract) is therefore suggestive of a response driven by something other than a demand shift. Nonetheless, attributing a specific percentage of price changes to enrollment effects versus other supply-side responses would require a general equilibrium framework beyond the scope of the current paper. Parsing out demand-side versus supply-side responses to quality ratings is an important area of future research.

28See Appendix B for details of this analysis. This evidence is suggestive of other factors besides an enrollment effect, but not definitive. For example, plans with small enrollment changes may nonetheless still experience an “enrollment effect” due to a change in the make-up of the plan’s customers. Alternatively, these plans may already have more inelastic enrollees. This may allow for an increase in premiums that is still explained by changes in enrollments or by the pre-existing mix of customers. Without detailed data on plan enrollees, we cannot separately identify these effects. We thank an anonymous referee for clarification on this analysis.

29An important line of future research is to examine enrollment patterns with changes MA quality star ratings.
both specific to the MA market and in the broader economics literature.
References


8 Tables and Figures

Table 1: Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>2009</th>
<th>2010</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Plan-level Data, n=51,442 and 34,642</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enrollment</td>
<td>291.55 (1,413)</td>
<td>361.17 (1,600)</td>
</tr>
<tr>
<td>Overall Share %</td>
<td>1.18</td>
<td>1.26</td>
</tr>
<tr>
<td>Premium</td>
<td>37.69 (42.23)</td>
<td>53.27 (52.97)</td>
</tr>
<tr>
<td>Drug Coverage, %</td>
<td>58.58</td>
<td>64.39</td>
</tr>
<tr>
<td>HMO, %</td>
<td>16.32</td>
<td>24.12</td>
</tr>
<tr>
<td>PPO, %</td>
<td>18.53</td>
<td>33.71</td>
</tr>
<tr>
<td><strong>Market Characteristics, n=3,139 and 3,094</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MA Penetration</td>
<td>15.59 (11.03)</td>
<td>16.50 (12.12)</td>
</tr>
<tr>
<td>Mean Number of Plans</td>
<td>37.38 (22.31)</td>
<td>26.61 (17.58)</td>
</tr>
<tr>
<td>Population &gt; 65 in 1,000s</td>
<td>12.22 (34.90)</td>
<td>12.59 (35.74)</td>
</tr>
<tr>
<td>Population &gt; 85 in 1,000s</td>
<td>1.72 (5.11)</td>
<td>1.79 (5.34)</td>
</tr>
<tr>
<td>Unemployed, %</td>
<td>5.79</td>
<td>9.01</td>
</tr>
<tr>
<td>White, %</td>
<td>86.30</td>
<td>86.41</td>
</tr>
<tr>
<td>Black, %</td>
<td>9.11</td>
<td>9.18</td>
</tr>
<tr>
<td>Female, %</td>
<td>50.16</td>
<td>50.17</td>
</tr>
<tr>
<td>College Graduates, %</td>
<td>18.68</td>
<td>18.62</td>
</tr>
<tr>
<td>South, %</td>
<td>42.02</td>
<td>42.63</td>
</tr>
<tr>
<td><strong>Contract-level Star Ratings %, n=252 and 295</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td>1.98</td>
<td>0.00</td>
</tr>
<tr>
<td>2.0</td>
<td>9.92</td>
<td>4.07</td>
</tr>
<tr>
<td>2.5</td>
<td>24.21</td>
<td>24.41</td>
</tr>
<tr>
<td>3.0</td>
<td>28.97</td>
<td>29.83</td>
</tr>
<tr>
<td>3.5</td>
<td>21.43</td>
<td>20.67</td>
</tr>
<tr>
<td>4.0</td>
<td>11.11</td>
<td>12.20</td>
</tr>
<tr>
<td>4.5</td>
<td>2.38</td>
<td>7.78</td>
</tr>
<tr>
<td>5.0</td>
<td>0.00</td>
<td>1.02</td>
</tr>
</tbody>
</table>

*Enrollment data available for 20,768 plans in 2009 and 17,334 plans in 2010. Remaining plans have 10 or fewer enrollments and specific enrollments are therefore not provided by CMS.*
Table 2: RD Results for Contract Characteristics$^a$

<table>
<thead>
<tr>
<th>Star Threshold</th>
<th>2.25</th>
<th>2.75</th>
<th>3.25</th>
<th>3.75</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma_2$ Average Premium</td>
<td>12.82***</td>
<td>16.25***</td>
<td>28.58***</td>
<td>26.97***</td>
</tr>
<tr>
<td>$(\hat{\gamma}_2)$</td>
<td>(3.42)</td>
<td>(4.85)</td>
<td>(5.06)</td>
<td>(12.79)</td>
</tr>
<tr>
<td>$\gamma_2$ Change in 25th Percentile</td>
<td>11.30***</td>
<td>-0.42</td>
<td>23.68***</td>
<td>23.22*</td>
</tr>
<tr>
<td>$(\hat{\gamma}_2)$</td>
<td>(3.60)</td>
<td>(4.21)</td>
<td>(4.93)</td>
<td>(12.22)</td>
</tr>
<tr>
<td>$\gamma_2$ Change in 50th Percentile</td>
<td>20.28***</td>
<td>8.94**</td>
<td>27.20***</td>
<td>19.93*</td>
</tr>
<tr>
<td>$(\hat{\gamma}_2)$</td>
<td>(3.35)</td>
<td>(4.20)</td>
<td>(4.80)</td>
<td>(12.09)</td>
</tr>
<tr>
<td>$\gamma_2$ Change in 75th Percentile</td>
<td>9.22*</td>
<td>38.11***</td>
<td>35.91***</td>
<td>36.52***</td>
</tr>
<tr>
<td>$(\hat{\gamma}_2)$</td>
<td>(5.04)</td>
<td>(7.21)</td>
<td>(5.61)</td>
<td>(12.01)</td>
</tr>
<tr>
<td>$\gamma_2$ Change in Proportion of $0$ Premium Plans</td>
<td>-0.37***</td>
<td>0.00</td>
<td>0.20***</td>
<td>-0.33***</td>
</tr>
<tr>
<td>$(\hat{\gamma}_2)$</td>
<td>(0.07)</td>
<td>(0.07)</td>
<td>(0.05)</td>
<td>(0.07)</td>
</tr>
<tr>
<td>N</td>
<td>2,029</td>
<td>982</td>
<td>432</td>
<td>310</td>
</tr>
</tbody>
</table>

$^a$Results based on kernel-weighted local linear regressions with triangular kernel and a bandwidth of $h = 0.125$. Outcomes are aggregated to the contract/county level, excluding plans with 10 or fewer enrollees. Results were excluded for the 1.5 and 4.5 star ratings due to an insufficient number of contracts on the lower and upper ends of the 1.75 and 4.25 thresholds, respectively. Bootstrapped standard errors in parenthesis with 200 bootstrap replications. * $p<0.1$. ** $p<0.05$. *** $p<0.01$.

Table 3: RD Results for Plan-level Characteristics$^a$

<table>
<thead>
<tr>
<th>Star Threshold</th>
<th>2.25</th>
<th>2.75</th>
<th>3.25</th>
<th>3.75</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma_2$ Change in Plan Premium</td>
<td>-9.86***</td>
<td>20.09***</td>
<td>35.80***</td>
<td>30.82***</td>
</tr>
<tr>
<td>$(\hat{\gamma}_2)$</td>
<td>(2.01)</td>
<td>(3.23)</td>
<td>(3.51)</td>
<td>(7.77)</td>
</tr>
<tr>
<td>$\gamma_2$ Switch to $0$ Premium Plan$^b$</td>
<td>0.03</td>
<td>0.001</td>
<td>0.03***</td>
<td>-0.27***</td>
</tr>
<tr>
<td>$(\hat{\gamma}_2)$</td>
<td>(0.02)</td>
<td>(0.002)</td>
<td>(0.01)</td>
<td>(0.07)</td>
</tr>
<tr>
<td>$\gamma_2$ Switch to &gt;$0$ Premium Plan$^c$</td>
<td>0.12**</td>
<td>0.007</td>
<td>-0.25***</td>
<td>0.03</td>
</tr>
<tr>
<td>$(\hat{\gamma}_2)$</td>
<td>(0.05)</td>
<td>(0.03)</td>
<td>(0.06)</td>
<td>(0.05)</td>
</tr>
<tr>
<td>N</td>
<td>2,244</td>
<td>1,639</td>
<td>616</td>
<td>713</td>
</tr>
</tbody>
</table>

$^a$Results based on kernel-weighted local linear regressions with triangular kernel and a bandwidth of $h = 0.125$. The analysis includes all plans with more than 10 enrollees operating in both 2009 and 2010. Results were excluded for the 1.5 and 4.5 star ratings due to an insufficient number of contracts on the lower and upper ends of the 1.75 and 4.25 thresholds, respectively. Bootstrapped standard errors in parenthesis with 200 bootstrap replications. * $p<0.1$. ** $p<0.05$. *** $p<0.01$.

$^b$Indicator set to 1 if plan went from charging a positive premium to 2009 to a $0$ premium in 2010.

$^c$Indicator set to 1 if plan went from $0$ premium in 2009 to a positive premium in 2010.
Table 4: RD Results for Plan Exit & Entry\(^a\)

<table>
<thead>
<tr>
<th>Star Threshold</th>
<th>2.25</th>
<th>2.75</th>
<th>3.25</th>
<th>3.75</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Plan Exit</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall Results</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\hat{\gamma}_2)</td>
<td>-0.73***</td>
<td>-0.20**</td>
<td>0.14**</td>
<td>-0.12**</td>
</tr>
<tr>
<td>(0.03)</td>
<td>(0.09)</td>
<td>(0.06)</td>
<td>(0.06)</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>11,334</td>
<td>10,121</td>
<td>1,213</td>
<td>1,438</td>
</tr>
<tr>
<td>Among Plans with Premiums = $0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\hat{\gamma}_2)</td>
<td>-0.65***</td>
<td>0.38***</td>
<td>0.21</td>
<td>-0.14</td>
</tr>
<tr>
<td>(0.03)</td>
<td>(0.09)</td>
<td>(0.38)</td>
<td>(0.17)</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>9,601</td>
<td>613</td>
<td>140</td>
<td>282</td>
</tr>
<tr>
<td>Among Plans with Premiums &gt; $0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\hat{\gamma}_2)</td>
<td>-1.42***</td>
<td>-1.36***</td>
<td>0.08</td>
<td>-0.10*</td>
</tr>
<tr>
<td>(0.17)</td>
<td>(0.17)</td>
<td>(0.06)</td>
<td>(0.06)</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>1,733</td>
<td>9,508</td>
<td>1,073</td>
<td>1,156</td>
</tr>
<tr>
<td><strong>Plan Entry</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall Results</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\hat{\gamma}_2)</td>
<td>1.21***</td>
<td>-0.17***</td>
<td>0.18***</td>
<td>0.25***</td>
</tr>
<tr>
<td>(0.12)</td>
<td>(0.06)</td>
<td>(0.06)</td>
<td>(0.04)</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>6,546</td>
<td>2,457</td>
<td>1,257</td>
<td>855</td>
</tr>
<tr>
<td>Among Plans with $0 Premium</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\hat{\gamma}_2)</td>
<td>0.85***</td>
<td>-0.17**</td>
<td>-0.48</td>
<td>0.34***</td>
</tr>
<tr>
<td>(0.21)</td>
<td>(0.08)</td>
<td>(0.60)</td>
<td>(0.05)</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>3,468</td>
<td>793</td>
<td>172</td>
<td>332</td>
</tr>
<tr>
<td>Among Plans with $&gt;0 Premium</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\hat{\gamma}_2)</td>
<td>1.53***</td>
<td>-0.37**</td>
<td>0.17***</td>
<td>0.15***</td>
</tr>
<tr>
<td>(0.09)</td>
<td>(0.19)</td>
<td>(0.06)</td>
<td>(0.06)</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>3,078</td>
<td>1,664</td>
<td>1,085</td>
<td>523</td>
</tr>
</tbody>
</table>

\(^a\)Results based on kernel-weighted local linear regressions with triangular kernel and a bandwidth of \(h = 0.125\).

The analysis for plan exit includes all plans operating in 2009, and the analysis for plan entry includes all plans operating in both 2009 and 2010. Results were excluded for the 1.5 and 4.5 star ratings due to an insufficient number of contracts on the lower and upper ends of the 1.75 and 4.25 thresholds, respectively. Bootstrapped standard errors in parenthesis with 200 bootstrap replications. * \(p<0.1\). ** \(p<0.05\). *** \(p<0.01\).
Figure 1: Binned Averages for Mean Contract Premiums with Fitted Regression Lines

a. 2.25  
b. 2.75  
c. 3.25  
d. 3.75  

*Binned averages calculated with a bin width of 0.02. Fitted regression lines based on a 4th order polynomial regression.*
Figure 2: RD Graphs for Mean Contract Premiums

\*Results based on kernel-weighted local linear regressions with triangular kernel and a bandwidth of $h = 0.125$.\*
Figure 3: RD Graphs for Plan-level Premiums

\*Results based on kernel-weighted local linear regressions with triangular kernel and a bandwidth of $h = 0.125$.\*

\[\text{Change in Premium} \begin{cases} \text{Distance from Threshold} \\ \text{a. 2.25} \\ \text{b. 2.75} \\ \text{c. 3.25} \\ \text{b. 3.75} \end{cases}\]
Figure 4: Differential Effects by Market HHI\textsuperscript{a}

\begin{itemize}
\item a. 2.25
\item b. 2.75
\item c. 3.25
\item d. 3.75
\end{itemize}

\textsuperscript{a}Results based on kernel-weighted local linear regressions with triangular kernel and a bandwidth of $h = 0.125$. 