Do Private Insurers Game Medicare Advantage?*

Hyunjin Yun†

November 30, 2025

Abstract

Nearly all Americans aged 65 and older are covered by Medicare, and more than half are now enrolled in Medicare Advantage, a publicly financed, privately administered health insurance program. Although the federal government sets payment schedules for private insurers, the total payments to them have surged in recent years. In this paper, I examine how higher federal payments in Medicare Advantage affect insurers' strategic behavior, plan characteristics, and market entry. Exploiting sharp discontinuities in federal payment rates across counties, I find that insurers retain 88% of the additional payments, ranging from 73% to 97% depending on market competition. The decomposition of payment use shows that most of the additional payments are captured as profit rather than spent on serving enrollees. There is little evidence that higher payments are passed through to enrollees in the form of lower maximum out-of-pocket limits, reduced copayments, or improvements in plan quality. Higher payments also do not induce entry of new insurers but instead increase narrow-network plans that target healthier individuals. Overall, the results raise concerns about whether federal payments to private insurers improve enrollee benefits and promote competition.

Keywords: Medicare Advantage, Private insurers, Benchmark-bid system
JEL Codes: H22, H51, J13

*I am deeply grateful to my advisor, Professor Carlos Dobkin, for his invaluable guidance and support throughout this research project. His thoughtful feedback greatly improved the quality of this research. All remaining errors are my own.

†Ph.D. candidate in Economics, the University of California, Santa Cruz.

1 Introduction

Medicare is one of the largest social insurance programs in the United States, accounting for 3.8% of GDP in 2024 and projected to reach 6.2% by 2049. Originally operated on a fee-for-service basis, the program reimbursed providers directly for each service delivered. To contain rising costs, the federal government introduced Medicare Advantage (MA) in 1997 as a publicly financed but privately delivered alternative. Under MA, the government pays private insurers a predetermined amount per enrollee per month, shifting financial risk from the public to the private sector, and these insurers offer a variety of plan options to beneficiaries. Over the past decade, MA enrollment has more than doubled, with over half of all Medicare beneficiaries enrolled as of 2024. Yet over the same period, government payments to private insurers have more than tripled despite the program's original intent to curb Medicare spending. This rapid growth underscores the importance of understanding how private insurers respond to financial incentives embedded in the MA program.

In this paper, I study the extent to which private insurers retain government payments, how the funds private insurers received are allocated between profit and enrollee-related costs, and how higher payments affect plan characteristics and market entry decisions. Identifying how much of these payments private insurers retain sheds light on whether MA operates as intended, as the program is designed to incentivize insurers to pass through a larger portion of these payments to enrollees, thereby attracting more beneficiaries through competition. Examining changes in plan costs and characteristics helps identify whether retained payments are captured by insurers or passed through to beneficiaries via increased plan generosity. Finally, analyzing how payments influence entry decisions reveals whether the MA program fosters a more competitive market and whether insurers adjust plan offerings in ways that encourage risk selection by targeting healthier beneficiaries.

To identify exogenous variation in government payments to private insurers, I examine the benchmark-bid system, which determines payments, plan design, enrollee premiums, and enrollee benefits. In this system, the federal government sets county-level benchmarks that

cap the maximum payments insurers may receive. Insurers then submit bids reflecting the revenue they require to cover Medicare Advantage enrollees, and these bids determine the payment each insurer ultimately receives. Beginning in 2012, the benchmark formula was reformed to assign different multipliers to four county groups ranked by traditional Medicare (TM) costs. The reform applies lower multipliers to high TM-cost counties to contain federal spending and higher multipliers to low TM-cost counties to encourage insurer entry. These group-specific multipliers create three annual cutoffs in the county ranking of TM costs, providing plausibly exogenous variation in government payments because the county ranking, cutoffs, and multipliers are mechanically determined by federal rules and thus set outside the MA market.

I construct a novel dataset that links county-level TM costs with detailed MA plan data, including plan characteristics, enrollment, and government payments between 2014 and 2019. This dataset provides within-plan, cross-county variation in government payments, enabling identification based on discrete changes induced by the reform in 2012. Because plans operate across multiple counties with different payment rates, the data reveal meaningful within-plan variation in both government payments and bids. By examining how bids respond to these discrete changes, I identify the extent to which private insurers retain additional government payments. I combine these data with insurers' reported allocations of payments across health care costs, administrative expenses, and profits, enabling a decomposition of how the captured payments are used. This feature allows me to assess whether higher payments are passed through to beneficiaries in the form of more generous coverage, or instead captured as insurer profit.

Using a regression discontinuity (RD) design, I find that private insurers capture the vast majority of additional government payments rather than passing them through to consumers. Specifically, by stacking all the cutoffs across years, I find that plans in counties just below the cutoff receive government payments about \$16.8 higher than those in counties just above it. In response, private insurers raise bids by \$14.5, implying that roughly 88% of

additional government payments are absorbed rather than passed through to beneficiaries. The results are consistent across cutoffs and highlight potential flaws in the incentive design of the benchmark-bid system. Instead of lowering premiums or expanding supplemental benefits such as dental services to attract enrollees, insurers choose to retain a larger share of government payments to maximize profits, resulting in lower pass-through to beneficiaries.

These findings align with recent evidence from Duggan et al. (2016), who document a statistically insignificant pass-through rate of roughly one-eighth, and stand in contrast to the much higher pass-through rate of 45% found by Cabral et al. (2018). My results therefore suggest that over time, private insurers have secured a stronger position in the MA market and that the current benchmark system may be delivering lower pass-through to consumers. This pattern is consistent with a highly concentrated MA market in which beneficiaries exhibit substantial inertia. Even in relatively competitive markets, insurers retain a large share of payments, suggesting limited consumer responsiveness to premium or benefit changes.

The extent to which private insurers retain payments varies systematically with degree of market competition. I calculate the Herfindahl–Hirschman Index (HHI) for each county based on plan enrollment and compare the most competitive markets in the bottom 10% of the HHI distribution with the most concentrated markets in the top 10%. Insurers retain 73% of additional payments in competitive counties with an HHI below 1,340, compared with 97% in highly concentrated counties with an HHI above 5,600. Although greater competition modestly improves pass-through, the overall incentive mechanism remains ineffective at translating higher payments into consumer benefits.

Next, I analyze how insurers allocate the additional government payments they secure. Because MA operates as a regulated private market, the amount that insurers actually receive from the government, represented by the plan's bid, is designed to cover the estimated costs of providing Medicare Part A and Part B services, the associated administrative expenses, and a profit component known as the gain/loss margin. The decomposition analysis shows

that nearly 80% of the additional payments are driven by higher gain/loss margin, indicating that private insurers capture most of the increased payments as profit rather than passing them through to expand health care services. Likewise, I find little evidence of reductions in premiums, copayments, or measurable improvements in plan quality. Copayments for primary care, specialist, and urgent care visit remain unchanged. Overall, the evidence indicates that insurers raise bids by inflating their profits rather than improving benefits, suggesting that insurers are gaming the MA system.

Next, I study whether the increase in government payments affects insurers' entry decisions, as the reform that changed the payment formula was designed to encourage market entry in lower TM-cost counties. I find no evidence that insurers increase entry below the cutoff despite receiving higher government payments. A likely explanation is that establishing provider networks involves substantial fixed costs, and the uncertainty associated with the proximity to the cutoff may further discourage entry. Existing insurers could expand by offering additional plans, which is less costly than entering new counties. However, the estimated effect is small and statistically insignificant. When disaggregating by plan type, the only meaningful increase occurs among narrow-network plans that tend to attract healthier beneficiaries, suggesting increased risk selection. I also find an increase in the number of high-quality plans, reflecting insurers' incentives for the higher payments they receive from the government. Yet enrollment in these plans does not rise, ruling out improvements in plan quality through consumer shifting toward higher-quality options. Taken together, the evidence suggests that insurers capture additional payments as profit and engage in risk selection rather than enhancing benefits or improving consumer welfare.

This paper contributes to several strands of literature. First, it contributes to the literature on MA payments by examining the causal effect of government payments on market outcomes under the current policy environment. Several studies identify variation in government payments by constructing a simulated instrument or by using variation within plans over time, but the variation may be associated with unobservables (Song et al., 2013;

Stockley et al., 2014). Two papers identify the causal impact of increased government payments on premiums and plan coverage by exploiting policy-induced payment floors that existed before the 2012 reform, which introduced a new incentive structure based on quartile cutoffs of county-level per capita TM costs. Cabral et al. (2018) exploit a difference-in-differences design using a policy change in 2000 that created payment floors in rural and urban counties, while Duggan et al. (2016) use population thresholds that generate separate rural and urban payment floors. Given that the current benchmark system uses a different incentive structure, new evidence under this regime is both timely and policy relevant. In addition, this study uses plan–county level data, whereas most prior work aggregates to the county level. Because government payment increases are determined at the county level while insurers adjust their decisions given the counties they serve, aggregating only to the county level can blur the relationship between payment changes and insurer behavior. By using plan–county level data, this paper can more precisely capture insurers' micro-level responses to government payments.

Second, this paper contributes to the literature on risk selection by insurers. Specifically, it provides evidence of cream skimming, a strategy in which insurers selectively target low-risk enrollees by adjusting benefits (Frank et al., 2000; Geruso and Layton, 2017), and examines its relationship to market power in MA. Cao and McGuire (2003) and Batata (2004) document selection in MA, showing that average costs are lower among MA enrollees than among beneficiaries in TM. However, evidence on the specific mechanisms insurers use to avoid high-risk individuals is limited. One exception is the study by Decarolis and Guglielmo (2017), who find that because enrollees can switch into five-star plans outside of the regular enrollment period, insurers reduce both premiums and generosity to deter high-risk individuals from enrolling. Similarly, Shepard (2022) finds the private insurers then to exclude top-hospitals to avoid high-cost enrollees. My paper instead shows that private insurers reduce the portion of payments that could be passed through as supplemental benefits such as dental, vision, or hearing services. Limiting these benefits serves as a

mechanism to discourage enrollment by high-risk individuals, who place greater value on them. Furthermore, I document insurers' behavior consistent with risk selection, as they disproportionately expand narrow-network plans that are less attractive to high-cost enrollees.

Last, the paper contributes to the growing literature on overpayments and insurer gaming in health insurance markets, in which private insurers exploit policy design to increase premiums and capture larger government subsidies (Decarolis, 2015; Ho et al., 2017) or to enroll low-cost individuals with high risk scores in order to receive greater subsidies while maintaining low costs (Brown et al., 2014). Whereas the existing literature documents insurers' upcoding behavior to increase risk scores (Dafny, 2010; Dafny et al., 2012; Trish and Herring, 2015; Ho and Lee, 2017), this paper provides more direct evidence that insurers capture additional government payments as profit. This paper is also closely related to Geruso and Layton (2020), who document that private insurers upcode diagnoses to inflate risk scores and capture higher payments, resulting in excess public spending.

The rest of the paper proceeds as follows. Section 2 outlines the institutional background on MA and the benchmark-bid system. Section 3 presents the data, and Section 4 details the identification strategy. Section 5 reports the main results, Section 6 discusses the findings, and Section 7 concludes.

2 Institutional Background

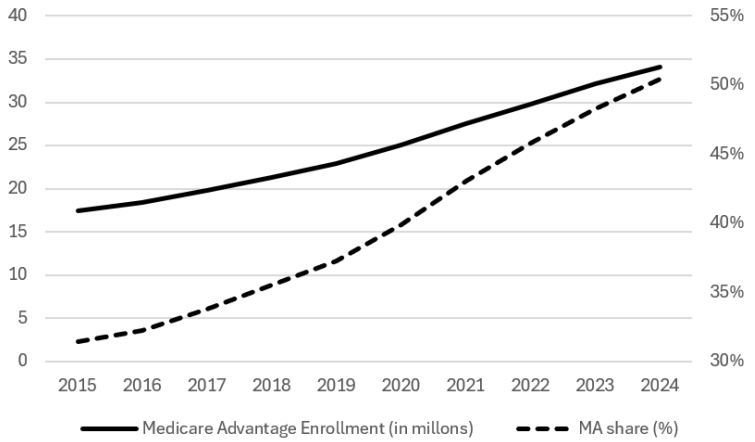
2.1 Medicare Advantage

Fee-for-service TM, the original public health insurance program for individuals aged 65 and older, was introduced in 1965 to provide coverage for those who were often denied enrollment in private health insurance plans due to higher risks associated with aging compared to younger individuals. As the program expanded to cover more costly health care services and the number of enrollees grew, concerns over rising expenditures prompted the government to reduce costs. In 1997, Medicare Advantage (MA) was introduced to address these challenges

by incorporating private insurers into the provision of Medicare benefits.

The rationale behind this shift was as follows. First, it was anticipated that program costs could be reduced, as private insurers have expertise in risk assessment and cost management, thereby potentially lowering overall health care expenditures. This structure also allowed the government to allocate a fixed amount of funds while transferring financial risk to private insurers in case adverse events occurred. Second, in contrast to TM's uniform fee-for-service structure, private insurers provided a range of plan options, allowing beneficiaries to select the plan that best suited their individual needs, thereby enhancing consumer choice. Last, the introduction of private insurers was intended to leverage competition in the market, with the expectation that insurers would compete on premiums and quality.

Figure 1. Medicare Advantage Enrollment and Share

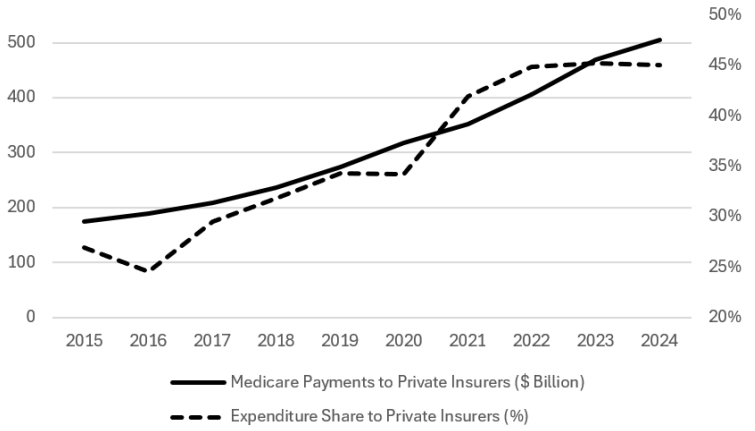


Note: This figure shows how enrollment in Medicare Advantage and its share of total Medicare enrollment have evolved. The other Medicare program is fee-for-service traditional Medicare. The figure is drawn from 2025 Medicare Trustee Report.

Due to these benefits, MA has become increasingly popular among beneficiaries. Figure 1 shows how enrollment in MA and its share to total Medicare have evolved for the last 10 years. Between 2015 and 2024, enrollment increased from about 17 million to 34 million, and the share of MA enrollment rose from 32% to 50%. This share is projected to reach 58% in 2034.[1] Meanwhile, government payments to private insurers increased from approximately \$170 billion to \$500 billion over the same period, as shown in Figure 2. Given that existing

[1]See the 2025 Medicare Trustee report.

Figure 2. Payments to Private Insurers in Medicare Advantage



Note: This figure shows how Medicare Advantage payments from the government to private insurers and its expenditure share of total Medicare expenditure have changed. The figure is drawn from 2025 Medicare Trustee Report.

MA beneficiaries are healthier than TM beneficiaries (Brown et al., 2014), the program costs may have risen as individuals transitioned from TM. Considering the sustained growth in the market share of the top five insurers from 66% in 2010 to 80% in 2024, there is a need to investigate whether these companies are gaming the MA program.[2]

2.2 Benchmark-Bid System

In 2006, the government introduced the benchmark-bid system to foster competition among private insurers. Before this system, payments to private insurers were based solely on the cost of providing TM benefits in a given area, adjusted by an area factor to further capture variation in local health care costs. In contrast, under the benchmark-bid system, the government aimed to facilitate competition in the market by creating incentives for private insurers to reduce plan costs and compete for enrollees through expanded plan coverage or reduced premiums.

Because MA is a publicly funded but privately delivered health insurance program, the determination of premiums, payments, and benefits reflects a blend of market-based

[2]See the Kaiser Family Foundation's 2024 annual report.

incentives and government regulation. Each county functions as a distinct market in which insurers decide whether to enter and, if so, how to set premiums and design plan characteristics under the benchmark-bid system. Under this system, two key elements are crucial for understanding the incentive mechanism of private insurers: the benchmark and the bid.[3] The benchmark represents the maximum payment from the government to a private insurer in a given county, while the bid reflects the revenue that insurers require to provide TM benefits, including health care costs, administrative expenses, and an allowable profit margin to their enrollees.

The bid includes projected utilization and pricing of various health-care services such as inpatient care, outpatient visits, and skilled-nursing-facility stays, together with demographic changes and anticipated intensity of care for the contract year. This estimate is calculated prior to the contract year, and because it determines the insurer's revenue and the plan's design, the bid amount is subject to government evaluation and plans may face sanctions if the bid is materially overstated.[4] Taken together, the ex ante nature of bid determination, its direct linkage to government revenue, and the cases of sanctions for misreporting imply that insurers use bids not only to reflect expected costs but also as an instrument to maximize profit.

I now formally describe the benchmark-bid system, detailing how the benchmarks and bids are determined and how they map into government payments, enrollee premiums, and covered benefits. First, before the contract year t, the government sets the benchmark for MA beneficiaries based on the costs of providing care to TM beneficiaries in the same county, denoted by TMC_{ct} in county c for contract year t. TMC_{ct} is calculated as the average TM costs in county c over year $t-8$ through $t-4$. The government then ranks counties into

[3]The term "bid" used in this context should not be confused with the auction concept of bidding. It is simply the terminology adopted in the MA payment system, and I continue to use the term throughout the paper. Unlike in auctions, nearly all plans that submit a bid are actually offered for sale in the subsequent contract year.

[4]For instance, if a plan's medical loss ratio falls below 0.85 for three consecutive years, the insurer can be prohibited from enrollment or even terminated. For the 2023 contract year, approximately 500 contracts were audited, of which around 90 underwent review and 19 were found to be in violation. Penalties ranged from $10,000 to $30,000.

four groups based on TMC_{ct}, forming quartiles. A multiplier α_{it} is applied to each quartile: 1.15 for the first quartile with the lowest TM-cost counties, 1.075 for the second quartile, 1 for the third quartile, and 0.95 for the fourth quartile with the highest TM-cost counties.[5] These multipliers are incorporated into the benchmark formula,[6] though the final benchmark amount also reflects several policy adjustments.[7]

For plan i serving only a single county c in year t, a benchmark is calculated as follows:

$$B_{ict} = f(TMC_{ct} \times \alpha_{i,t}).$$

Most MA plans, however, operate across multiple counties.

Let plan i serve counties $c \in \lambda_{it}$, where λ_{it} denotes the set of counties in which plan i is available for purchase in year t. Consequently, plan i serving counties $c \in \lambda_{it}$ faces a weighted benchmark based on the number of enrollees and their risk scores for each county. Specifically, a benchmark that plan i faces serving counties $c \in \lambda_{it}$ in year t is as follows:

$$B_{it} = \sum_{c \in \lambda_{it}} r_{ict} w_{ict} B_{ict}.$$

Next, after observing the benchmark, a private insurer submits a bid, which summarizes the estimated cost plus the profit margin for covering its MA enrollees in its plan i for each county c, denoted as b_{ict}. Like the benchmark, the plan-level bid b_{it} is calculated as the weighted average of b_{ict}, based on the set of counties a plan serves, the number of enrollees, and risk scores:

$$b_{it} = \sum_{c \in \lambda_{it}} r_{ict} w_{ict} b_{ict}.$$

The relative amount of the bid to the benchmark determines the amount of payments to private insurers and plan characteristics such as plan premium and benefits. If $b_{it} > B_{it}$,

[5] In the benchmark formula, α_{it} denotes the average multiplier across the previous and current year assigned to a county based on its TM cost ranking. Specifically, $\alpha_{it} = \frac{\alpha^*_{i,t-1} + \alpha^*_{i,t}}{2}$.

[6] It is further modified by the plan's quality rating. A 0.05 increase is added to $\alpha_{i,t}$ for plans with a star rating of 4 or higher, and 0 otherwise. The formula is designed to incentivize plans to improve quality in order to receive a higher benchmark.

[7] For example, it is capped at the level that would have been determined under the pre-ACA policy. In addition, Indirect Medical Education, Graduate Medical Education, and the Geographic Adjustment Factor affect the benchmark. A full description is provided in Appendix A.

the government pays B_{it} to the plan, and the plan must charge premium P_{it}, determined as follows:

$$P_{it} = TMP_t + b_{it} - B_{it},$$

where TMP_t is the baseline premium paid by TM beneficiaries in year t.[8] Per-enrollee revenue is the sum of premium revenue and the government payment, $TMP_t + b_{it}$.

Although the government payment is canceled out from the plan's revenue, it still influences the premium level and thereby affects the demand for the plan. Thus, the profit function of plan i is given by

$$\Pi_{it} = (TMP_t + b_{it} - c_{it}) \cdot Q(TMP_t + b_{it} - B_{it}, b_{it}),$$

where c_{it} denotes the constant marginal cost, assuming full risk adjustment by the government,[9] and $Q(\cdot)$ denotes demand as a function of the plan premium and the bid amount if the plan revenue from the government is used to expand plan benefits. Holding the bid constant, an increase in the benchmark lowers premiums for MA beneficiaries, raising demand. In response, private insurers may raise their bids to increase per-enrollee revenue, even though they may lose some customers. If the increase in bids leads to better plan coverage, insurers may still attract customers despite higher premiums. Thus, private insurers have incentives to raise bids to maximize their profits in response to the increase in the benchmark.

As shown in Figure 3, the government payments that a plan receives depend on the relative sizes of the bid and the benchmark. When $b_{it} > B_{it}$, the plan receives the payments only up to the benchmark and must charge beneficiaries a premium for the portion of the bid above the benchmark. Because the bid represents the revenue insurers expect to need to provide TM services, a higher bid implies that the insurer anticipates higher costs for its MA enrollees relative to TM beneficiaries, resulting in higher premiums.

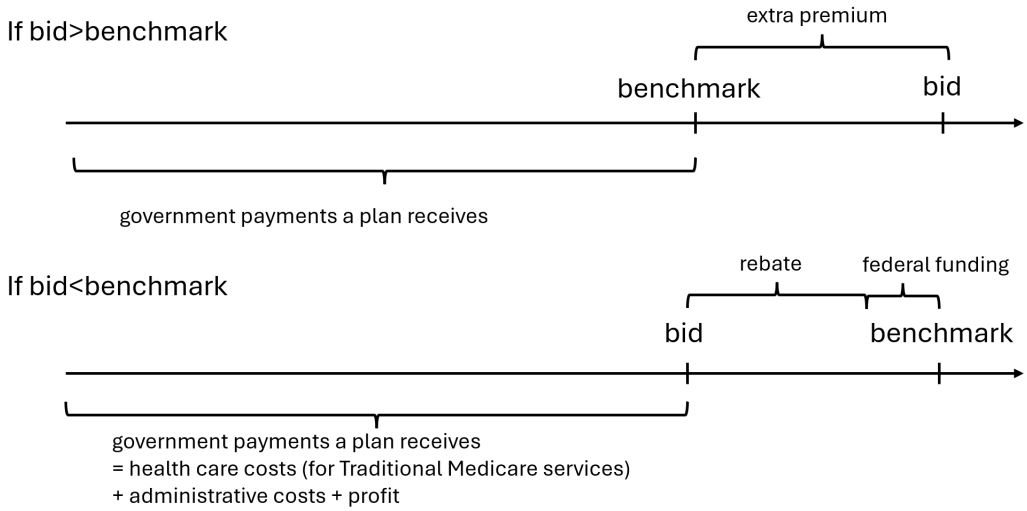
If $b_{it} < B_{it}$, the government pays b_{it} to the plan along with an additional portion of the difference between the benchmark and the bid as a rebate, denoted by $r_{it} = \gamma_{it}(B_{it} - b_{it})$.

[8] In 2024, the Part A premium was free, and the Part B premium was \$178.

[9] See Jaffe and Shepard (2020).

The rate γ_{it} depends on the plan's star rating.[10] This rebate must be used to benefit enrollees by reducing premiums for Part B or Part D, lowering cost sharing, or offering supplemental benefits such as dental services, hearing exams, or eye exams. Note that these are distinct from the TM services incorporated into the bid amount. In this case, the plan's revenue from the government equals its submitted bid.

Figure 3. Benchmark-bid system



Note: This table is based on the CMS documents Advance Notice of Methodological Changes for Calendar Year 2018 and the Medicare Advantage Manual: Payments to Medicare Advantage Organizations. The rebate percentage is determined by the plan's star rating: plans with a 4.5-star rating or higher receive 70%, those rated between 3.5 and 4 receive 65%, and those rated 3 or lower receive 50% of the gap between the benchmark and the bid.

To understand private insurers' incentives under this system, consider the following example. Suppose the benchmark is \$1,000 and the plan's bid is \$900. If the benchmark increases to \$1,100, only the rebate rises, which can be used to lower premiums or enhance supplemental benefits and thereby attract more enrollees. If a private insurer raises its bid, its revenue from the government increases but the rebate falls, meaning that premiums must rise or supplemental benefits must be reduced, potentially causing a loss of enrollees. However, if the increase in per-enrollee revenue from a higher bid outweighs the reduction

[10] $\gamma = 0.7$ for ratings of 4.5 or above, $\gamma = 0.65$ for ratings between 3.5 and 4, and $\gamma = 0.5$ for ratings of 3 or below.

in enrollment, raising the bid will increase profit. Thus, insurers may have an incentive to raise bids in response to benchmark increases.

3 Data

To exploit variation across counties based on the 2012 benchmark formula, I construct a novel dataset by merging administrative data from the Centers for Medicare & Medicaid Services (CMS) covering the universe of MA plans from 2014 to 2019. The unit of observation is the combination of plan and county. The data contain information on the benchmark and bid amounts that each plan faces in the counties it serves, allowing for the observation of sharp variation in benchmarks at the quartile cutoffs.

I exclude several types of plans that are only available to individuals meeting specific eligibility criteria, such as PACE, 1876 Cost plans, Religious Fraternal Benefit plans, Employer/Union-only contracts, and Provider Sponsored Organizations. Consequently, the data include Health Maintenance Organizations (HMO), Health Maintenance Organization Point-of-Service plans (HMO-POS), Local Preferred Provider Organizations (Local PPO), Regional Preferred Provider Organizations (Regional PPO), and Private Fee-for-Service (PFFS). The data also include plan characteristics such as premiums and cost sharing, as well as information on the local insurers offering each plan.

I merge the dataset with county-level information on the average cost of covering TM beneficiaries, which serves as the running variable in the RD design. The data also include each county's assigned quartile and the corresponding multiplier applied in the benchmark calculation. By ranking counties based on their cost levels and identifying their assigned quartile, I recover the TM cost values that determine quartile cutoffs. Since there are four quartiles, there are three cutoffs, which vary from year to year as they are determined by the annual cost distribution.

I also merge the plan–county dataset with a dataset containing the total number of

Medicare beneficiaries, the number enrolled in MA, and the corresponding MA penetration rate. Additionally, I incorporate county-level demographic data from the U.S. Census, including race, gender, and age distributions, as well as provider counts by type from the U.S. Department of Health and Human Services.

Benchmark amounts vary at the plan–county level, as they depend on the cost of covering TM beneficiaries, the applicable multiplier, the plan's star rating, and other county- and plan-specific factors. Plan-level benchmarks are also observed as weighted averages of plan–county benchmarks, with weights based on the share of enrollees that a plan serves in each county. Similarly, bids, reflecting the required revenue to cover MA enrollees, are also observed at the plan level.

Table 1. Summary Statistics.

	(1) All plans	(2) HMO	(3) HMOPOS	(4) LPPO	(5) RPPO	(6) PFFS
Panel 1: Plan–county level						
Cost of Traditional Medicare	763.4	771.4	759.8	745.3	777.8	744.6
	(90.92)	(86.17)	(91.53)	(82.47)	(93.59)	(95.73)
Benchmark	838.4	845.4	842.3	836.2	837.5	814.5
	(71.45)	(72.67)	(68.69)	(69.17)	(71.16)	(69.61)
Bid	757.1	730.8	742.9	768.9	776.2	805.6
	(83.93)	(83.47)	(82.28)	(78.64)	(74.31)	(80.40)
Risk score	1.092	1.156	1.009	1.006	1.132	1.053
	(0.372)	(0.369)	(0.385)	(0.370)	(0.414)	(0.188)
Projected enrollment	347.4	605.6	288.8	231.1	141.8	67.47
	(1663.4)	(2534.8)	(1089.8)	(668.9)	(450.7)	(159.4)
Observations	273,156	104,612	20,872	68,768	51,276	25,840
Panel 2: Plan level						
Premium ($)	21.26	13.86	40.01	33.46	24.16	49.06
	(38.42)	(32.92)	(52.30)	(41.44)	(37.48)	(39.14)
Out of pocket amount ($)	5189.5	5126.8	4664.5	5449.3	6121.0	4700
	(1552.5)	(1596.9)	(1480.9)	(1397.9)	(1055.9)	(328.6)
Copayment for primary care visits ($)	9.236	7.544	11.13	11.44	12.35	15.93
	(8.120)	(7.972)	(8.998)	(7.293)	(6.946)	(6.621)
Copayment for specialist visits ($)	34.06	31.40	35.96	38.47	40.66	42.95
	(13.57)	(14.92)	(10.66)	(9.193)	(9.720)	(7.050)
Copayment for outpatient visits ($)	212.9	201.5	209.2	240.7	213.0	248.7
	(106.1)	(108.5)	(96.28)	(96.48)	(108.2)	(104.6)
Star rating	3.878	3.868	4.051	3.929	3.459	3.559
	(0.559)	(0.581)	(0.584)	(0.476)	(0.383)	(0.396)
Observations	17,025	11,172	1,209	3,871	357	416

Panel 1 presents summary statistics of the plan–county level data. The unit of observation is the year-plan–county combination, indicating which counties each plan serves. Each plan faces a benchmark for every county it serves and calculates a bid for each county. Aggregated plan-level benchmarks and bids are then calculated based on the number of enrollees in each county and risk scores. Private insurers estimate the numbers prior to the contract year. Panel 2 presents plan-level data, where plan characteristics are observed and merged with the benchmark–bid dataset. Although numerous variables describe plan characteristics, only a few provide usable values for the analysis.

Table 1 presents summary statistics for representative MA plans and for each plan type from 2014 to 2019. All the estimates reported in Panels 1 and 2 are expressed on a per-

enrollee, per-month basis. Panel 1 reports statistics at the plan–county level, disaggregated by plan type. The plan types are HMO, HMO-POS, Local PPO, Regional PPO, and PFFS. The cost of TM reflects average spending on individuals enrolled in TM within a given county. This measure is treated as exogenous, as it is determined outside of the MA market. The benchmark is constructed by adjusting TM costs using the formula, resulting in a positive correlation between the benchmark and TM costs. The bid represents the plan's estimated cost with a profit margin of serving MA beneficiaries within a county. Bids for HMO plans tend to be lower relative to other plan types, partly because beneficiaries who select HMOs are typically healthier and opt for plans with narrower provider networks. In contrast, PFFS plans have the highest bids, reflecting both higher costs and greater enrollee flexibility in choosing providers. The risk score indicates the relative health of MA beneficiaries compared to the average TM population, where a score of 1 reflects the average.[11] Across all plan types, risk scores are generally below 1, consistent with evidence that MA enrollees are, on average, healthier than those in TM.

Panel 2 presents plan-level summary statistics, where counties served by a given plan are aggregated to the plan level. The first row reports the plan premium, which is the monthly premium paid by enrollees in addition to the standard Part B premium.[12] Per-enrollee, per-month premiums are lowest among HMO plans, reflecting both narrower provider networks and the relatively healthier composition of enrollees. The subsequent rows report cost-sharing structures across plans. The average out-of-pocket maximum is approximately $5,200, and HMO plans also feature the lowest copayments for primary care visits, specialist visits, and outpatient services.

[11] Although MA beneficiaries are generally healthier than TM beneficiaries, their risk scores are higher largely due to more intensive diagnostic coding and upcoding by private insurers seeking higher payments. See Geruso and Layton (2020).

[12] In 2024, the baseline premium for TM (Part B) is $174. Part A typically has no premium.

4 Identification Strategy

The relationship between the benchmark and private insurers' decisions is endogenous, as the benchmark level is correlated with factors such as demographics, local provider markets, local laws, and regulations. For example, the benchmark is positively correlated with the cost of living, which may also influence bid levels as they reflect estimated health care costs. To overcome this endogeneity, I exploit plausibly exogenous variation in the benchmark across counties induced by the benchmark formula introduced in 2012.

In that year, the CMS implemented a reform designed to increase benchmarks in counties with lower TM cost, thereby encouraging private insurer entry while reducing excessive payments in high-cost counties. Specifically, CMS first calculates the county-level cost of covering TM beneficiaries and ranks counties based on this measure. Then, the CMS divides counties into four groups according to their cost quartiles. For each quartile, CMS assigns a multiplier, α^*_{it}, which is used in benchmark calculations, where $\alpha^*_{i,t} \in \{0.95, 1, 1.075, 1.15\}$. For example, a county in the first quartile, those with the lowest TM costs, is applied by a multiplier of 1.15. Counties in the second quartile are assigned a multiplier of 1.075, and so on.

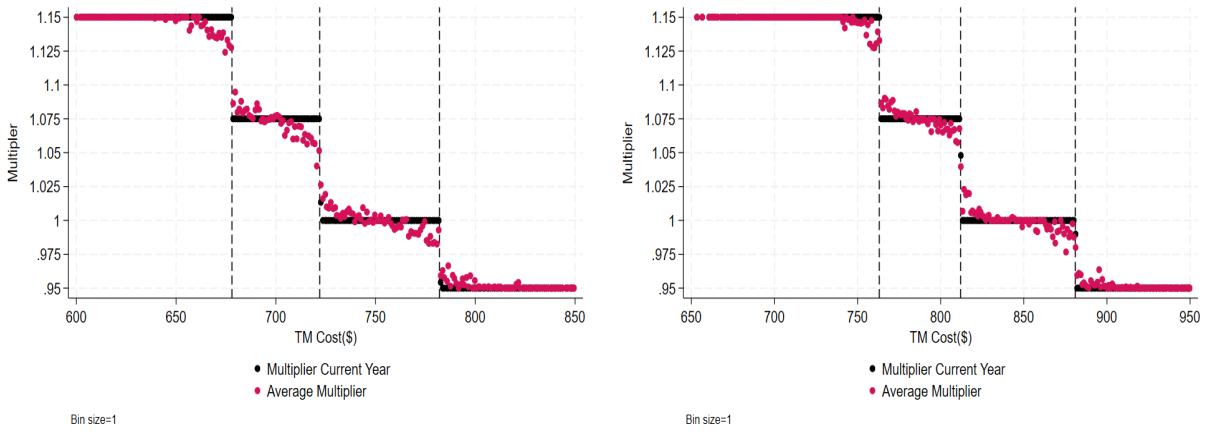
In addition, the multiplier depends not only on the current year's quartile assignment but also on the previous year's. For example, if a county was in the first quartile in the previous year and moves to the second quartile in the current year, it is assigned a multiplier of 1.1125, the average of 1.15 and 1.075. Thus, the actual multiplier applied to each county is given by

$$\alpha_{i,t} = \frac{\alpha^*_{i,t} + \alpha^*_{i,t-1}}{2}.$$

Figure 4 presents the schedule of current-year multipliers $\alpha^*_{i,t}$ and average multipliers $\alpha_{i,t}$ for the years 2016 and 2019. Each dot represents the average value within bins of \$1 in TM costs. In panel (a), counties with TM costs below \$678.11 fall into the first quartile; the second quartile includes counties with costs between \$678.11 and \$722.35; the third quartile

spans $722.35 to $782.38; and the fourth quartile includes counties above $782.38. The black dots depict the current-year multiplier schedule: 1.15 for the first quartile, 1.075 for the second, 1 for the third, and 0.95 for the fourth. Since the actual multiplier assigned to a county is the average of the current and prior year's multipliers, values such as 1.125, 1.0375, or 0.975 may appear around the thresholds.

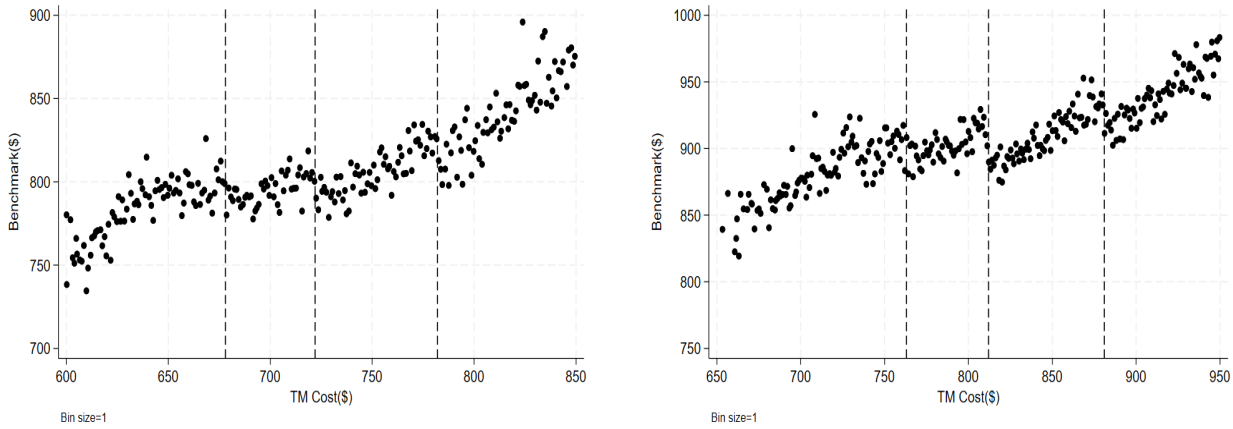
Figure 4. Traditional Medicare Costs and Multipliers in 2016 and 2019



Note: For example, county rank for contract year t is based on traditional Medicare cost in year t-1, calculated as average of traditional Medicare cost in year from t-8 to t-4. For the territories, CMS assign scale factor after they rank counties for the 50 states and the District of Columbia. On the figure, black colored dots represent scale factors for each county for each quartile. The red dots summarize the average of actual scale factor $\alpha'_{i,t}$ for each $1 bin.

In panel (b), while the threshold values shift due to the quartile rankings being redefined annually, the pattern of average multipliers remains qualitatively similar. Although the range of variation in the multipliers narrows after adjustment, the graph displays clear discontinuities at the quartile thresholds. These discontinuities induce exogenous changes in the benchmark when incorporated into the benchmark formula.

Figure 5. Traditional Medicare Costs and Benchmark in 2016 and 2019



Note: The threshold values change annually, as they are determined based on the annual county rank of traditional Medicare costs. Benchmarks are averaged within each $1 bin.

The actual multiplier applied to the benchmark calculation is expressed as follows:

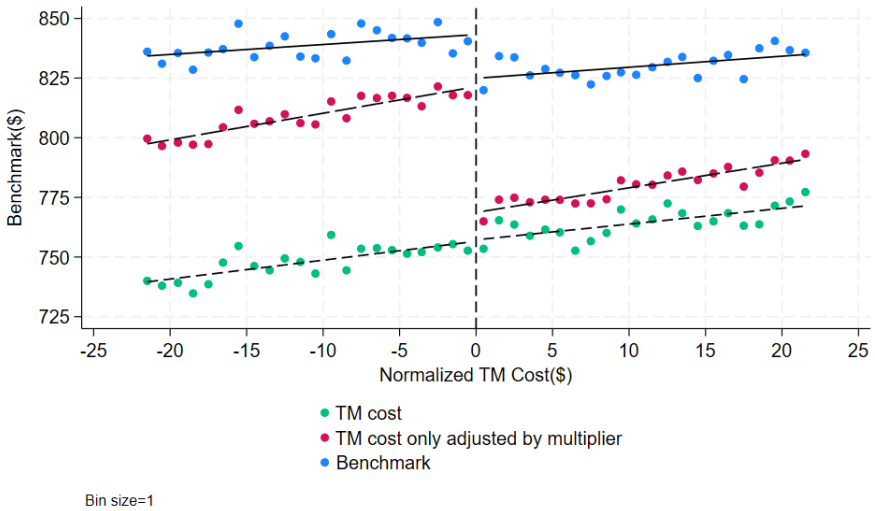
$$B_{ict} = f(TMC_{ct} \times \alpha_{i,t}).$$

At the quartile cutoffs, differences in $\alpha_{i,t}$ generate exogenous variation in the benchmark, though several other factors also influence the benchmark.

Figure 5 illustrates the benchmark schedule by quartile, where each county's benchmark is constructed by applying the multipliers to TM costs in 2016 and 2019. Each panel averages values within $1 bins. Although the benchmark is determined by multiple factors, both panels show clear discontinuities at the quartile cutoffs.

Figure 6 presents how the discontinuity of the benchmark is constructed at the pooled thresholds. First, as the numerical values of each threshold vary across years, I normalize each cutoff to zero and stack them across all years in the dataset. The normalized TM cost is restricted to the range of $[-22, 22]$, as the minimum difference between any two threshold values is $44. Without this constraint, an observation could fall on both sides of a normalized cutoff. For example, if an observation is $30 below an upper-quartile threshold in one year and the corresponding quartile width is only $50, then the same observation would appear at both −$30 and $20 when pooled across thresholds. This duplication would

Figure 6. Benchmark Variation at the pooled cutoff

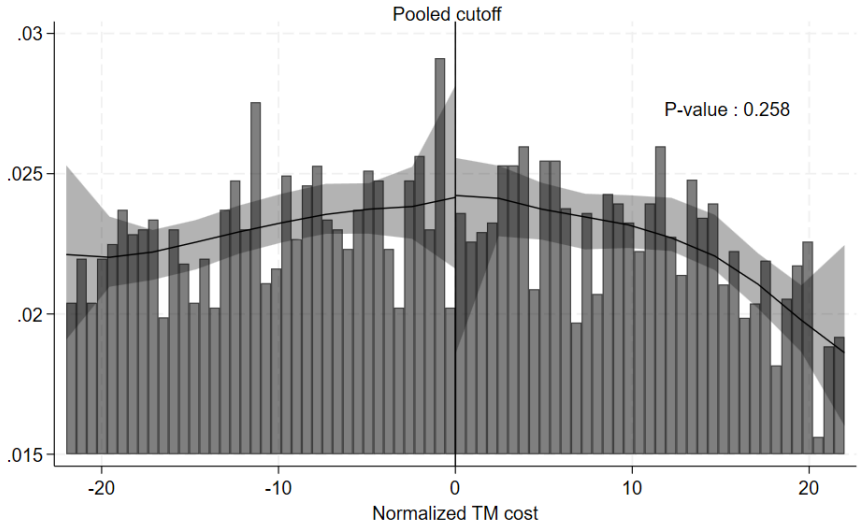


Note: The figure above illustrates the process through which sharp variations in the benchmark arise at the pooled cutoff. To maintain the same bandwidth across all thresholds, I exclude some observations to the left of −$22 from the first cutoff, and to the right of +$22 from the third cutoff. The green dots plot county-level TM costs on both the x-axis and y-axis without applying the benchmark formula. Because the observations are averaged within $1 bins, some noise appears, but overall the points closely follow the 45-degree line. The red dots show how the values change at the cutoffs when the quartile specific multiplier is applied to the TM costs. In practice, the benchmark is determined not only by these two values but also by many other factors. The blue dots show how the benchmark varies at the cutoffs when all of these factors are taken into account. Although the variation is smaller than before, it remains clear, demonstrating that using multipliers is a valid approach.

not introduce meaningful variation and instead inflate the number of observations near the cutoff, attenuating the observed discontinuity in the benchmark.

In the figure, green dots represent the schedule of TM costs and thus lie along the 45-degree line. The dots do not align perfectly with the 45-degree line because the figure is constructed by stacking observations across multiple cutoff values and averaging within each bin. For instance, if more observations fall within the fourth quartile, the average within a bin may lie above the line. Red dots depict the pattern when TM costs are multiplied by the assigned multipliers only, generating visible discontinuities at the pooled cutoff. Although the benchmark is further adjusted by additional factors—reducing the size of the gap at the cutoff—it still exhibits a meaningful discontinuity, as shown by the blue dots.

Figure 7. Density Check



Note: This figure presents a McCrary density test at the pooled cutoff across all years. The absence of a visible jump suggests no evidence of manipulation around the threshold. The estimated p-value is 0.258, and thus we fail to reject the null hypothesis of no discontinuity in the density of the running variable.

Exogenous variation in the benchmark around the threshold relies on the assumption that no other factors are systematically correlated with the benchmark at the cutoff. The benchmark value at the threshold determines the payment level for private insurers in MA, but it is based on the cost of covering TM beneficiaries, a factor determined outside the MA market. Furthermore, because the threshold is determined by the rank ordering of counties, private insurers cannot manipulate their placement into a lower-cost quartile to receive a more favorable multiplier. To assess the possibility of manipulation around the pooled threshold, I conduct a McCrary density test. Figure 7 presents the results. The estimated p-value of 0.258 indicates that I fail to reject the null hypothesis of continuity in the density of the running variable at the cutoff, providing no evidence of manipulation.

I next estimate a standard RD model using a two-stage least squares (2SLS) approach, as specified below:

$$B_{ict} = \pi_0 + \pi_1 Z_{ict} + \pi_2 TMcost_{ct} + \pi_3 Z_{ict} TMcost_{ct} + \epsilon_{ict},$$
$$b_{ict} = \beta_0 + \beta_1 \hat{B}_{ict} + \beta_2 TMcost_{ct} + \beta_3 Z_{ict} TMcost_{ct} + \mu_{ict},$$

where B_{ict} denotes the benchmark for county c that plan i serves in year t; $TMcost_{ct}$ is the per-enrollee, per-month cost of covering TM beneficiaries; and Z_{ict} is an indicator equal to 1 if county c, served by plan i, lies above the normalized pooled cutoff.

Using the estimated benchmark from the first stage, the 2SLS estimator of β_1 is obtained with the outcome variable b_{icst}, which represents the bid submitted by plan i in county c and state s at time t. The coefficient $\hat{\beta}_1$ captures the causal effect of a \$1 increase in the benchmark on bids: for example, the share of the benchmark retained by private insurers. Accordingly, $1 - \hat{\beta}_1$ measures the pass-through rate to beneficiaries. ϵ_{ict} and μ_{ict} are error terms that capture unobserved determinants of the benchmark and the bid, respectively. In all specifications, the running variable is restricted to the range $[-22.22]$ with a bin size of \$1. All models allow for a first-order polynomial.

Table 2. Balance Check

	(1)	(2)	(3)	(4)	(5)	(6)
	Hospitals	Nursing Facility (Dual)	Skilled Cursing Facility	Home Health Agency	X-Ray Supplier	Outpatient therapy
I(<Cutoff)	0.234	0.503	0.098	0.342	0.065	0.234
	(0.310)	(0.444)	(0.091)	(0.927)	(0.075)	(0.258)
Outcome mean	3.750	4.896	0.691	6.790	0.461	2.253
N	9,906	9,906	9,906	9,906	9,906	9,906
	(7)	(8)	(9)	(10)	(11)	(12)
	End Stage Renal Disease Facility	Nursing Facility	Intermediate Care Facility	Rural Health Clinic	Hospice	Federally Qualified Health Center
I(<Cutoff)	0.377	0.404	0.244	0.090	0.138	0.256
	(0.301)	(0.478)	(0.493)	(0.175)	(0.223)	(0.296)
Outcome mean	2.920	5.719	3.432	3.394	1.986	2.907
N	9,906	9,906	9,906	9,906	9,906	9,906
	(13)	(14)	(15)	(16)	(17)	(18)
	Male 65	Female 65	White (%)	Black(%)	Asian(%)	Hispanic(%)
I(<Cutoff)	535.822	712.727	0.914	1.048*	-0.021	-0.153
	(556.578)	(729.832)	(0.637)	(0.555)	(0.065)	(0.475)
Outcome mean	6567.732	8316.016	83.562	9.735	1.246	7.637
N	9,905	9,905	9,905	9,905	9,905	9,905

All medical facilities from column 1 through column 13 are scaled per 10,000 Medicare Advantage enrollment for each county. All regressions include year fixed effects, state fixed effects, and the weighted average of plan quality. Standard errors in parentheses and clustered at state-level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Since most private plans form networks by negotiating with health care providers and because these networks can influence the costs beneficiaries face, I conduct balance checks to test for differences in the number of health care providers at the threshold. I first create a county-level dataset and merge it with data on the number of health care providers by type. I then use these provider counts as outcome variables in the first stage of the RD design.

Column (1) of Table 2 shows that counties just above the cutoff have 0.2 fewer hospitals on average, though this difference is not statistically significant. The number of skilled nursing facilities dually certified for Medicare and Medicaid, as well as those certified only

for Medicare, also decreases above the threshold, but these effects are likewise not statistically significant. Counts for all other types of health care providers also decline above the cutoff, but none of these differences are statistically significant.

I also examine the effects on demographic variables. The number of men and women in the county both decline slightly above the cutoff, though the differences are not statistically significant. The share of Black residents decreases by about 1 percentage point, which is marginally statistically significant. All other demographic variables appear smooth across the threshold, suggesting that counties just above and below the cutoff are similar, supporting the validity of causal inference using this RD design.[13,14]

5 Results

5.1 Main Results

Figure 8 plots benchmarks against TM costs. Counties exhibit similar levels of TM costs, yet their benchmarks differ because benchmark levels also depend on plan characteristics. The average benchmark is approximately $840, with differences in multipliers of 7.5% at the first and second cutoffs and 5% at the third. The resulting variation is modest, about $20, or less than 2.5% of the average benchmark. This smaller variation arises because benchmark increases are capped at the pre-ACA rate, which is the amount that would have applied absent the policy change.

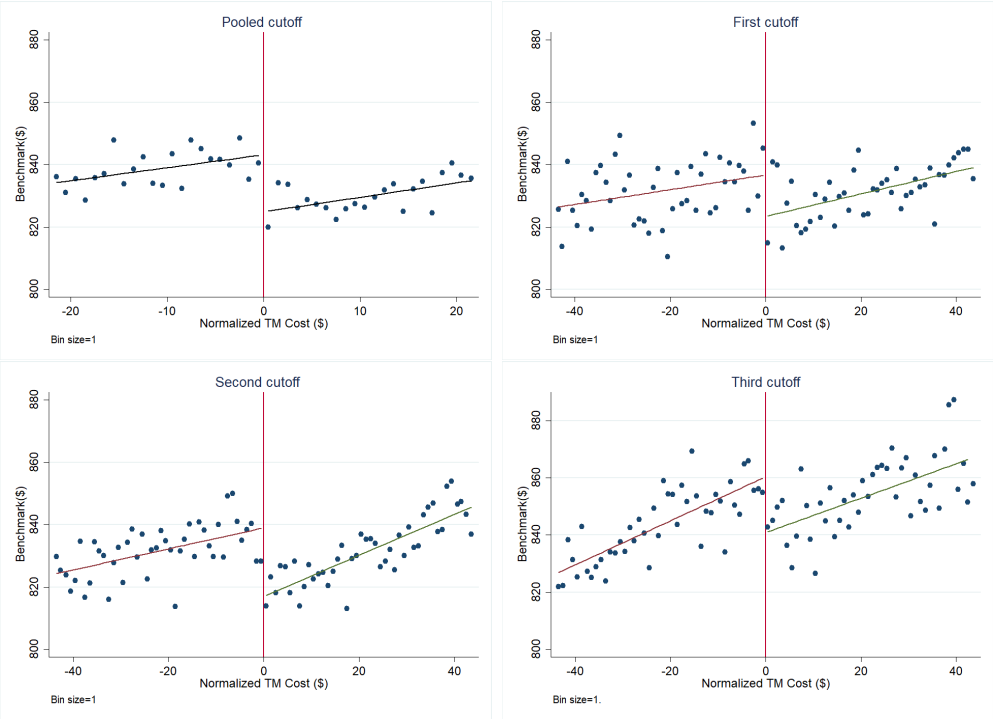
Panel (a), which stacks all cutoffs across years, displays a sharp and distinct discontinuity in benchmarks around the pooled cutoff, providing meaningful identifying variation. Benchmarks are higher below the cutoff, reflecting the government's assignment of greater multipliers to quartiles with lower TM costs. The subsequent panels stack the 25th, 50th, and 75th

[13] Appendix Figure B1 provides graphical evidence supporting the smoothness of covariates around the cutoff.

[14] Appendix Figure B2 presents additional graphical evidence using predicted values and residuals of the benchmark, obtained from a regression of the benchmark on covariates. As expected, the predicted values are smooth around the threshold, while the residuals exhibit a clear discontinuity.

percentile cutoffs across all years in the dataset. By applying higher multipliers to quartiles with lower TM costs, the policy narrows benchmark differences across counties.

Figure 8. First Stage: Variation in Benchmarks at the cutoff.



Note: The bandwidth in the first panel, based on the pooled threshold, is $22 to avoid presenting the same observations on both sides of the cutoff. For the remaining panels at each individual cutoff, the bandwidth is $44. The bin width is $1, and the predicted line is based on a first-order polynomial.

Table 3 presents the first-stage estimates of the effect of being assigned just below the quartile cutoff on the benchmark at each cutoff. I include several controls, beginning with year fixed effects, state fixed effects, insurer fixed effects, plan type, and plan quality as measured by star ratings. Year fixed effects account for inflationary trends, while state fixed effects capture variation in regulatory environments. Insurer fixed effects control for differences across insurers, such as brand, market power, and management practices. Plan type accounts for heterogeneity in benefit design and provider networks. Plan quality is measured by CMS star ratings, which are assigned on a 0.5-point scale by the federal government. Ratings are based on measures such as the percentage of enrollees receiving preventive care, management of chronic conditions, member experience, and medication

management, and are assigned at the contract level, which typically includes multiple plans.

Table 3. First stage: The variation in benchmarks around the cutoffs.

	(1)	(2)	(3)	(4)
Dependent Var: Benchmark ($)				
Panel 1: 25th percentile window				
I(<Cutoff)	13.545***	13.282***	12.760***	12.282***
	(2.581)	(2.357)	(2.279)	(2.245)
Mean	830.320	830.320	830.320	829.470
N	53,034	53,034	53,034	49,865
Panel 2: 50th percentile window				
I(<Cutoff)	21.618***	21.154***	20.682***	20.065***
	(2.178)	(2.004)	(1.866)	(1.901)
Mean	830.249	830.249	830.249	829.472
N	54,847	54,847	54,847	51,476
Panel 3: 75th percentile window				
I(<Cutoff)	18.432***	17.931***	17.384***	17.113***
	(2.520)	(2.515)	(2.397)	(2.436)
Mean	849.836	849.836	849.836	849.836
N	27,603	27,603	27,603	26,401
Panel 4: Pooled data				
I(<Cutoff)	17.550***	17.504***	16.918***	16.475***
	(1.486)	(1.383)	(1.351)	(1.365)
Mean	834.359	834.359	834.359	833.775
N	135,164	135,164	135,164	127,428
Year FE	X	X	X	X
State FE		X	X	X
Insurer FE			X	X
Plan Quality				X

Note: Cutoffs in each year are normalized to 0 to produce a single representative estimate. The range of the running variable, Traditional Medicare cost, is restricted to −$22 to $22. Plan type includes HMO, HMOPOS, Local PPO, Regional PPO, and PFFS. The MA plan's Star Rating serves as a proxy for plan quality. Standard errors are clustered at both the plan and county levels, and reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Across all panels, although the magnitudes of the estimates vary somewhat across quartiles, they consistently show that plans serving counties just below the cutoff face higher benchmarks. Column (1) of Panel 1 shows that plans in counties just below the cutoff at the 25th percentile face a $18.71 higher benchmark. If a plan serves 1,000 enrollees in a county just below the cutoff at the 25th percentile, this translates into $18,712 more in monthly government payments than it would if assigned above the cutoff. By stacking all cutoffs, I find that plans serving counties just below the pooled cutoff face an $18.25 increase in the benchmark, consistent with the discontinuity shown in panel (a) of Figure 8. Subsequent specifications include controls to account for potential differences across groups. However, because identification relies on an exogenous, policy-induced change, including controls has

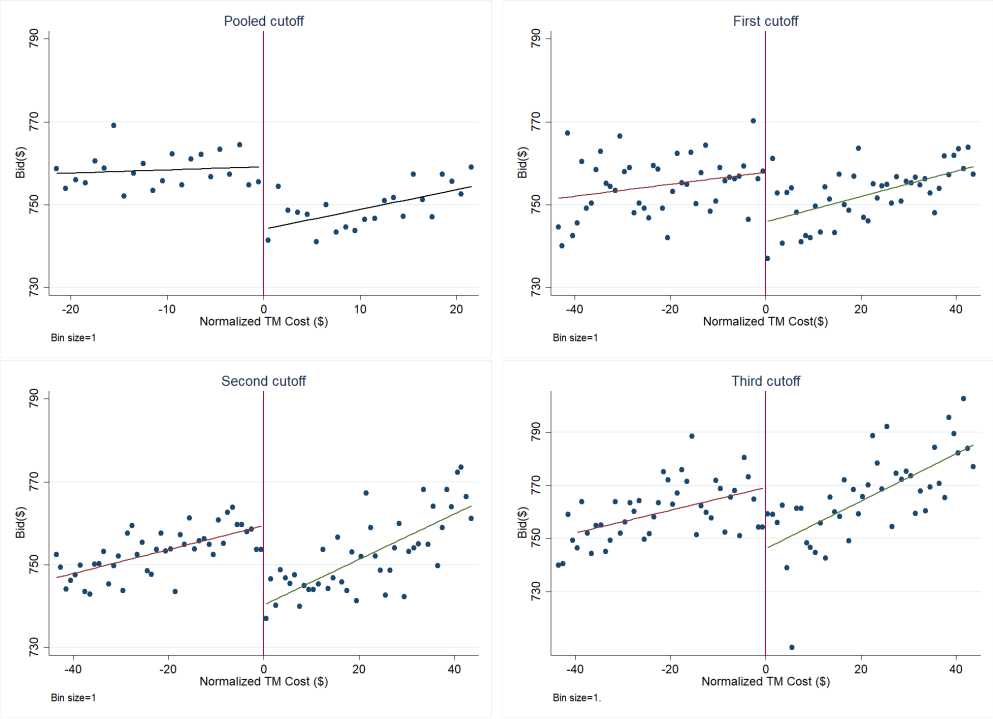
little effect on the estimated coefficients. After adding all controls, the estimated increase in the benchmark remains $16.48.

If a plan just below the cutoff does not adjust its bid, the corresponding increase in the rebate translates into $11.53 (= 0.7×16.48), assuming the plan's star rating is 4.5 or higher. The plan receives this rebate but must apply it toward reducing Part B or Part D premiums or toward offering an equivalent value of supplemental benefits. Assuming that private insurers do not raise their bids, a higher benchmark may increase enrollment, as the resulting rebate must be used to enhance the plan's attractiveness.

However, private insurers may have an incentive to raise their bids to maximize profits. Higher bids reduce the rebate amount available to reduce premiums and enhance benefits, potentially lowering enrollment. At the same time, increasing the bid directly raises the payments that private insurers receive from the government, thereby increasing revenue per enrollee. As there is a clear trade-off between higher revenue per enrollee and larger enrollment, insurers choose their bid levels to maximize profits.

Figure 9 illustrates how private insurers adjust bids at the pooled cutoff as well as at the 25th, 50th, and 75th percentiles of TM costs. In counties just below the cutoff where plans face an exogenously higher benchmark, bids increase by approximately $15. The pattern of variation in bids closely mirrors that of the benchmark, and the magnitude of the effect varies modestly across the 25th percentile, 50th percentile, and 75th percentile cutoffs.

Figure 9. Outcome variable: Variation in Bids at the cutoff.



Note: The bandwidth in the first panel, based on the pooled threshold, is $22 to avoid presenting the same observations on both sides of the cutoff. For the remaining panels at each individual cutoff, the bandwidth is $44. The bin width is $1, and the predicted line is based on a first-order polynomial.

Table 4 presents this pattern. Column (1) of Panel 1 shows that private insurers increase their bids by $13.33 in counties just below the cutoff. The estimated effects become somewhat larger at the second and third cutoffs, but the magnitudes remain stable across model specifications. Column (4) presents the estimate with all controls at the pooled cutoff, indicating that plans in counties just below the cutoff raise their bids by approximately $14.56 in response to the higher benchmark. The estimated effect remains stable across specifications with additional controls. This increase in bids reflects insurers' incentives to raise per-enrollee revenue at the risk of losing enrollment.

Table 5 reports the 2SLS estimates, which quantify how much private insurers increase their bids in response to a $1 increase in the benchmark. The estimates in the table can also be interpreted as the ratio of the coefficients reported in Table 4 to those in Table 3.

Table 4. The variation in bids around the cutoffs

	(1)	(2)	(3)	(4)
Dependent Var: Bid ($)				
Panel 1: 25th percentile window				
I(<Cutoff)	10.380***	10.910***	10.545***	10.263***
	(2.540)	(1.919)	(1.900)	(1.903)
Mean	752.316	752.316	752.316	753.640
N	53,034	53,034	53,034	49,865
Panel 2: 50th percentile window				
I(<Cutoff)	17.881***	16.502***	18.692***	18.205***
	(2.335)	(1.827)	(1.599)	(1.666)
Mean	751.835	751.835	751.835	753.645
N	54,847	54,847	54,847	51,476
Panel 3: 75th percentile window				
I(<Cutoff)	16.578***	14.066***	15.317***	14.846***
	(4.510)	(3.165)	(2.180)	(2.192)
Mean	759.903	759.903	759.903	761.582
N	27,603	27,603	27,603	26,401
Panel 4: Pooled data				
I(<Cutoff)	14.497***	14.167***	14.844***	14.550***
	(1.672)	(1.353)	(1.173)	(1.189)
Mean	753.708	753.708	753.708	755.330
N	135,164	135,164	135,164	127,428
Year FE	X	X	X	X
State FE		X	X	X
Insurer FE			X	X
Plan Quality				X

Note: Cutoffs in each year are normalized to 0 to produce a single representative estimate. The range of the running variable, Traditional Medicare cost, is restricted to −$22 to $22 to avoid having observations appear on both sides of the cutoff. Plan type includes HMO, HMOPOS, Local PPO, Regional PPO, and PFFS. The MA plan's Star Rating serves as a proxy for plan quality. Standard errors are clustered at both the plan and county levels, and reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Column (1) presents the 2SLS estimate without controls, indicating that a $1 increase in the benchmark raises bids by $0.83, statistically significant at the 1% level. After including all controls, the 2SLS estimates suggest that private insurers capture about 88% of additional government payments, using only the remaining portion to reduce premiums or expand supplemental benefits. The magnitude of the estimate is consistent with Duggan et al. (2016), who find that bids increase by roughly $0.90 for each $1 increase in the benchmark during 2007–2011. Although the benchmark formula changed after 2012, creating exogenous variation through policy-induced cutoffs rather than population thresholds, private insurers appear to exhibit similar behavioral responses.

The literature on pass-through rates typically uses premiums as the outcome variable to measure the extent of pass-through. However, in the MA market, when the benchmark

Table 5. Two-stage least square estimates: The effect of Benchmark on Bid

	(1)	(2)	(3)	(4)
Dependent Var: Bid ($)				
Benchmark ($)	0.826***	0.809***	0.877***	0.883***
	(0.078)	(0.052)	(0.026)	(0.026)
Outcome mean	753.708	753.708	753.708	755.330
N	135,164	135,164	135,164	127,428
Year FE	X	X	X	X
State FE		X	X	X
Insurer FE			X	X
Plan quality				X

Note: This table reports the 2SLS estimates, obtained by dividing the estimates in Table 6 by those in Table 4. The coefficients indicate how much private insurers raise bids in response to a $1 increase in the benchmark. Standard errors are clustered at both the plan and county levels, and reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

exceeds the bid, private insurers may either reduce premiums or expand supplemental benefits by an equivalent amount. Moreover, although about 93% of plans bid below the benchmark, more than 40% still charge an additional premium, suggesting that insurers often adjust benefits rather than lower premiums. From the enrollee's perspective, an increase in the benchmark, holding the bid fixed, raises the rebate, which can be used to either reduce premiums or enhance benefits of equivalent value.

For this reason, I use premium minus rebate as the outcome variable to measure the pass-through rate. It is useful to highlight two institutional aspects of the MA program. First, when the benchmark exceeds the bid, the rebate is not the full difference but depends on the plan's star rating: plans with a rating of 4.5 or higher retain 70% of the difference as a rebate; those rated 4 or 3.5 retain 65%; and those rated 3 or below retain 50%. Second, if the bid exceeds the benchmark, no rebate is provided, and plans must charge an additional premium to enrollees.

Table 6 reports the estimated pass-through rates. Without controls, enrollees capture about 16% of additional government payments. After including insurer fixed effects, the estimate declines to 8%. Based on Column (4), the standard errors rule out pass-through rates greater than 10.6% at the 95% confidence level. The estimates from Tables 5 and 6 do not sum to one because only a portion of rebates is used to benefit enrollees, while

Table 6. Two-stage least square estimates: Pass-through rate

	(1)	(2)	(3)	(4)
Dep: Premium($)				
Benchmark($)	-0.041	-0.039**	0.001	-0.017
	(0.041)	(0.017)	(0.008)	(0.016)
Outcome mean	24.442	24.442	24.442	25.242
Dep: Premium–Rebate ($)				
Benchmark ($)	-0.159***	-0.169***	-0.079***	-0.062***
	(0.082)	(0.045)	(0.023)	(0.022)
Outcome mean	-25.749	-25.749	-25.749	-23.023
N	135,052	135,052	135,052	127,316
Year FE	X	X	X	X
State FE		X	X	X
Insurer FE			X	X
Plan quality				X

Note: If the bid is below the benchmark, the rebate equals the difference between the benchmark and the bid, multiplied by a factor based on the plan's Star Rating. For plans with a Star Rating of 4.5 or higher, the multiplier is 0.70; for plans with a rating of 4.0 or 3.5, it is 0.65; and for plans rated below 3.5, it is 0.50. If the bid exceeds the benchmark, insurers must charge an additional premium equal to the difference between the bid and the benchmark. Cutoffs in each year are normalized to 0 to produce a single representative estimate. The range of the running variable, Traditional Medicare cost, is restricted to −$22 to $22 to avoid having observations appear on both sides of the cutoff. The MA plan's Star Rating serves as a proxy for plan quality. Standard errors are clustered at both the plan and county levels, and reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

the remainder is retained in the Medicare Trust Fund. The estimated pass-through rate of 6% is smaller than the 13% reported by Duggan et al. (2016), though their estimate is not statistically significant. They define enrollee cost as the sum of premiums and out-of-pocket spending minus rebates and find that pass-through primarily occurs through reductions in out-of-pocket costs. Because variation in premiums and rebates may correlate with out-of-pocket costs, the estimates in this paper are smaller in magnitude but more precisely estimated.

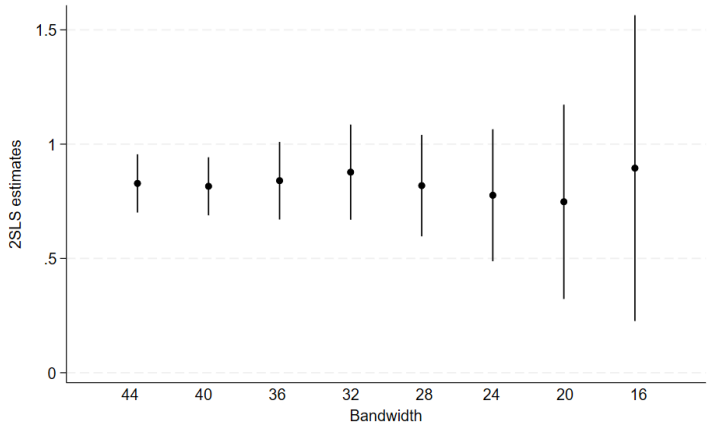
5.2 Robustness Check

As robustness checks, I study four alternatives. First, I estimate models using higher-order polynomials. Appendix Figure C1 reports results based on a second-order polynomial. Allowing greater flexibility in the functional form does not meaningfully alter the magnitude of the estimates. Under the baseline first-order specification, bids increase by $0.828 in

response to a $1 increase in government payments. The corresponding second-order estimate is $0.836.

Second, I conduct a sensitivity analysis by varying the bandwidth to assess whether the magnitude of the estimates is sensitive to the choice of window size. Figure 10 reports the 2SLS estimates obtained using a sequence of bandwidth values. The baseline estimate, shown on the far left, corresponds to a bandwidth of $44, which includes observations within $22 on either side of the threshold. I then decrease the bandwidth in $4 intervals and estimate the models accordingly. The estimates remain stable across different bandwidth choices, and all are statistically significant at the 1% level.

Figure 10. Sensitivity Check



Note: The bandwidth used in the main analysis is $44. Estimates are based on regressions without year fixed effects, state fixed effects, plan type fixed effects, and plan quality. The figure displays 95% confidence intervals.

Third, I implement an optimal bandwidth selection procedure. The optimal bandwidth minimizes the asymptotic mean squared error of the local linear regression estimator, balancing the trade-off between bias and variance. I use a triangular kernel, which assigns greater weight to observations closer to the cutoff and less weight to those farther away. This approach improves precision by prioritizing observations near the threshold, which are most informative for local identification. The optimal bandwidth is 6.596 under the baseline first-order specification and 8.97 under a second-order specification. These values are smaller

than the bandwidth used in the baseline specification, yet the resulting estimates remain similar in magnitude. The corresponding regression results are reported in Appendix Figure C2.

Last, I implement a donut-hole specification that excludes observations in a narrow range around the cutoff. This approach addresses the fact that counties near the threshold are more likely to switch quartile assignment based on annual ranking. Such switchers are assigned the average of multipliers from the previous and current years and thus face the same level of multiplier regardless of whether they fall just below or above the cutoff, meaning they do not contribute to variation in the benchmark. The regression results indicate that excluding these observations does not affect the magnitude of the estimates. Additionally, I exclude quartile switchers directly. The estimated magnitude increases slightly, but the difference is not statistically significant relative to the baseline 2SLS estimate. Appendix Figure C3 shows the results.

5.3 Role of Competition

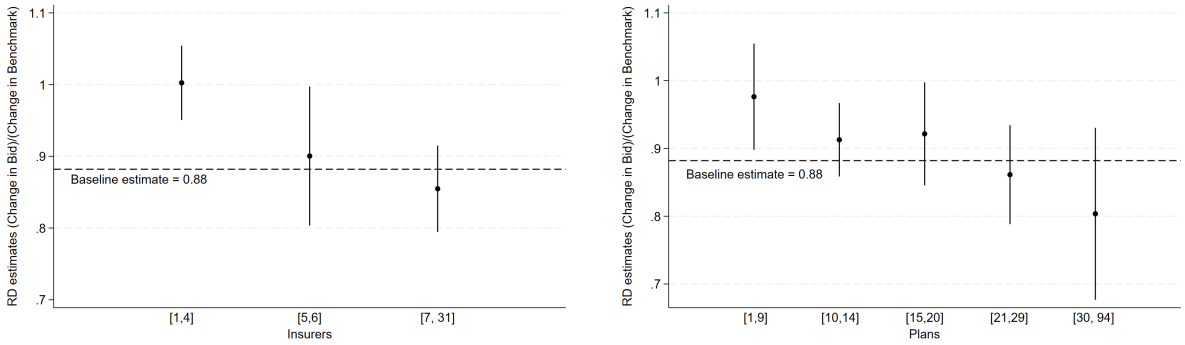
To study the role of market competition in determining the share of the benchmark captured by private insurers, I first group counties based on the number of local insurers operating within each county. Panel (a) of Figure 11 reports the 2SLS estimates for three groups: counties with one to four insurers, those with five or six, and those with seven or more.[15] The estimates decline as markets become more competitive, suggesting that consumers benefit more when private insurers raise their bids less to attract additional enrollees. The p-value for the difference between the estimates in concentrated and competitive markets is 0.0003, indicating they are statistically different at the 1% level.

Panel (b) presents the RD estimates by the number of plans offered in each county, grouped by equal quintiles. The estimates exhibit a similar declining pattern as the number

[15]These groups are defined assuming equal market shares. In industrial organization, an HHI above 2,500 denotes a highly concentrated market, 1,500–2,500 indicates moderate concentration, and less than 1,500 reflects a competitive market.

of plans increases. The p-value for the difference between the estimates in concentrated and competitive markets is 0.024, rejecting equality at the 5% level.

Figure 11. Market Competition by the number of insurers and plans



Note: According to the standard measure of market competitiveness, where an HHI of 2,500 or higher indicates a highly concentrated market and an HHI of 1,500 or lower indicates a competitive market, I divide the sample assuming that insurers hold equal market shares in each county. In the first figure, the first estimate indicates how much private insurers raise bids in response to $1 increase in the benchmark in a competitive market. In the second figure, each sample is grouped by 20% quintiles of the distribution.

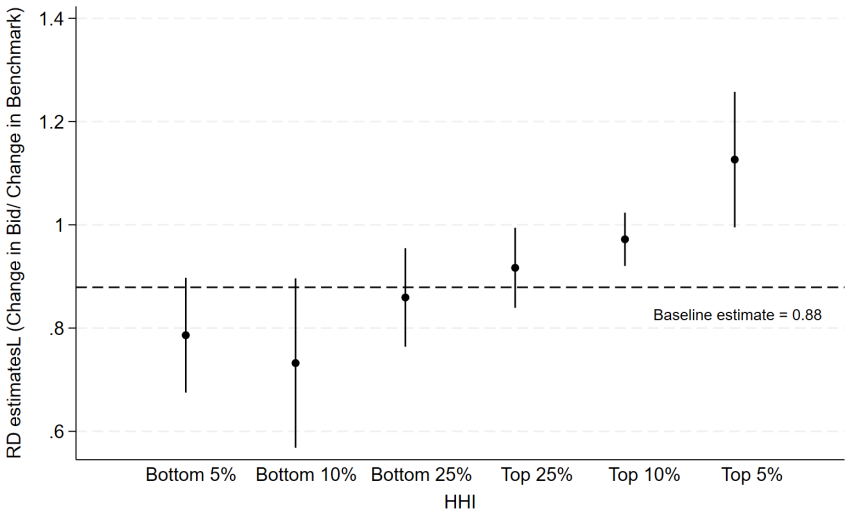
Table 7. Market competition by insurers and plans

	Insurers			Plans				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep var: Bid ($)	[1,4]	[5,6]	[7,31]	[1,9]	[10,14]	[15,20]	[21,29]	[30,94]
Benchmark($)	1.003***	0.900***	0.855***	0.976***	0.913***	0.922***	0.861***	0.804***
	(0.026)	(0.049)	(0.031)	(0.040)	(0.028)	(0.039)	(0.037)	(0.065)
Outcome mean	756.481	751.434	756.056	757.714	750.563	751.459	755.831	762.107
N	25668	22385	79378	22244	25698	27923	27394	24172

Note: Columns (1) to (3) classify the sample based on the number of insurer companies operating in each county. Columns (4) to (8) are defined at the plan level, where the included plans are HMO, HMOPOS, Local PPO, Regional PPO, and PFFS. Plans with special eligibility requirements are excluded from the count. All regressions include the full set of controls used in the main analysis, so the estimates are directly comparable to the baseline estimate of 0.88. Standard errors, clustered at the county or plan level, are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

However, even among counties with the same number of insurers, market dominance may differ depending on each insurer's enrollment shares. Similarly, simply grouping counties by the number of plans does not capture these differences. Therefore, I calculate the HHI based on each plan's enrollment share within a county and classify markets according to the degree of competition. Figure 12 presents the results. The bottom 5% of the HHI distribution represents highly competitive markets where HHI is less than 1,340. In these counties,

Figure 12. Market Competition by Herfindahl–Hirschman Index



Note: I compute the HHI using each plan's enrollment share in each county it serves. The first estimate reflects the extent to which private insurers raise their bids in response to a $1 increase in the benchmark in the most competitive market in the data. The difference in estimates between the bottom 5% and the top 5% is statistically significant at the 5% level.

private insurers retain approximately 79% of additional government payments, whereas in the most concentrated counties (the top 5% of the HHI distribution, where HHI exceeds 6,945), plans increase their bids by more than the increase in the benchmark.

Table 8 presents the estimates, all of which are statistically significant at the 99% confidence level. The p-value for the difference in estimates between the bottom 5% and top 5% is 0.0001, indicating that market competition plays a significant role in ensuring that government payments translate into consumer benefits.

Table 8. Market competition by HHI

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Var: Bid($)						
Benchmark($)	0.786***	0.732***	0.859***	0.917***	0.972***	1.126***
	(0.057)	(0.084)	(0.049)	(0.040)	(0.026)	(0.067)
Outcome mean	752.461	754.435	755.785	755.923	761.324	764.112
N	3,657	9,996	31,166	28,052	11,180	5,228

Note: HHI is measured at the year–county level, as the county is the unit of market in Medicare Advantage. It is calculated based on each plan's share of projected enrollment relative to the total projected enrollment within the county. At the cutoff for the bottom 5% of markets, the HHI is 1,340, indicating the most competitive markets, while at the cutoff for the top 5% the HHI is 6,945, reflecting the most concentrated markets. From columns (1) to (6), year and state fixed effects are included. Standard errors are two-way clustered at the plan and county levels and reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

There are two noteworthy points from this analysis. First, the MA market itself is highly concentrated. In the data, the HHI at the bottom 25th percentile is around 2,100, while it reaches about 4,200 at the 75th percentile. This concentration helps explain why the aggregate estimate in the main analysis is relatively high at 0.88. Second, although the results are consistent with the general prediction that greater supplier competition leads to a higher pass-through to consumers, the share of government payments captured by insurers still exceeds 73% even in the competitive market.

5.4 Bid Decomposition

In this section I decompose plan bids to identify which components private insurers adjust in response to increases in the benchmark. Before turning to the decomposition, I first examine whether the observed bid increases can be explained by changes in risk scores or enrollment.

From the balance test, counties near the cutoff exhibit similar observable characteristics. However, when a plan in a county below the cutoff faces a higher benchmark, insurers may respond by attracting riskier enrollees, thereby increasing the average risk of their pool. Because payments are risk-adjusted, insurers can receive greater reimbursements from the government by doing so. Such a response could explain the observed bid increases, as a higher share of riskier enrollees or an increase in overall enrollment can raise the frequency and intensity of health care service utilization, thereby increasing the underlying cost components that determine the bid.

However, Table 9 shows no evidence that plans, which serve counties facing higher benchmarks below the cutoff, experienced increases in projected risk scores or projected enrollment. This result suggests that not only are counties near the cutoff are similar in their observable characteristics, but insurers also continue to target comparable populations across these counties. It further indicates that insurers are not inflating risk scores or adjusting projected enrollment to capture higher payments. I also examine whether actual enrollment increased or decreased differently from projected enrollment, but the magnitudes of their

Table 9. The effect on projected risk score and enrollment

	(1) Risk Score	(2) Projected Enrollment	(3) Realized Enrollment
I(<Cutoff)	0.002	23.316	19.011
	(0.002)	(27.316)	(27.412)
Outcome mean	1.091	316.114	423.185
N	127,428	98,668	98,668

Note: This table reports the effect of being assigned below the pooled cutoff on risk score, projected enrollment, and realized enrollment. In the enrollment data, observations with fewer than ten enrollees are suppressed for privacy reasons and appear as missing. For comparison, when using projected enrollment as the outcome variable, I restrict the sample to observations where realized enrollment is not missing. All regressions include year, state, and insurer fixed effects, as well as controls for plan type and plan quality. Standard errors are clustered at both the plan and county levels, and reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

responses are not statistically different.

Although MA operates as a private market in which insurers offer a variety of plans, it functions within a highly regulated framework established by the federal government. That is, private insurers cannot freely raise their bids, as the bids must reflect the projected cost of providing Medicare Part A and Part B benefits to enrollees, including administrative expenses and a modest profit margin. To decompose these bids, I separate each plan's bid into three main components: health care costs for private insurers, administrative costs, and the gain/loss margin. Health care costs include projected costs for inpatient care, skilled nursing facilities, home health services, ambulance transport, outpatient care, and professional services. These projections incorporate the prices of health care services as well as expected utilization rates, intensity of care, and underlying assumptions. Administrative costs include sales, marketing, direct and indirect administration costs, and the net cost of private reinsurance. The gain/loss margin represents the profit of private insurers, defined as the additional revenue required beyond health care and administrative costs.

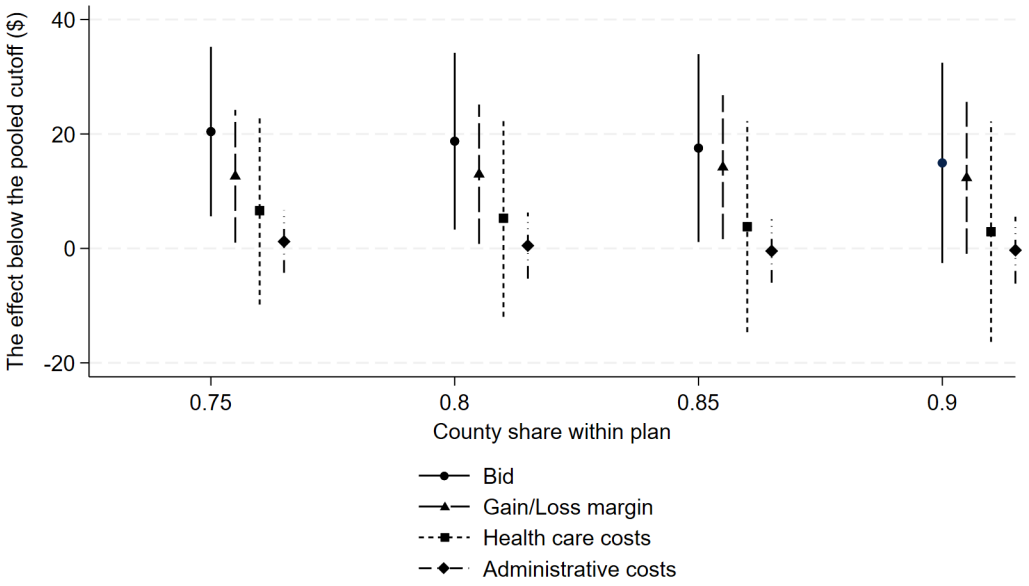
Understanding the source of the bid increase is crucial, as each component implies a different policy implication. If increases in bids are primarily driven by rising health care costs, it is important to examine whether the portion captured by private insurers

is ultimately shared with health care providers and translated into benefits for enrollees. Conversely, when bid increases are attributed to administrative costs, they are less likely to translate into improved enrollee benefits, though they may still lead to better access to care through enhanced customer service, outreach, or network management. Last, if the increase in bids is driven by a higher gain/loss margin, it can be interpreted simply as an increase in profit.

A key challenge in this analysis is that variables of health care costs, administrative costs, and gain/loss margin are reported at the plan level, whereas the main RD design relies on plan–county level variation at the county-level cutoffs. Because plans serve 14 counties on average, plan-level measures exhibit limited variation near the cutoffs. To address this issue, I select one main driving county for each plan—the county that has the greatest influence within the plan—and match each plan with that single county. This approach ensures that each plan is linked to only one county even when it operates in multiple counties. Although the counties belonging to a plan differ from one another, the selected county represents the largest share of the plan's enrollment, and because plan characteristics are largely determined by enrollment weights, this county exerts the most influence on the plan-level outcomes.

Figure 13 presents the regression results using samples based on the enrollment share of the main county within each plan. Plans are included in the sample only if they serve counties whose enrollment shares exceed a given threshold. The first plot focuses on plans in which a single county accounts for at least 90% of total enrollment. The solid line depicts how bids change below the pooled cutoff, showing an increase of approximately \$17.5 when the sample includes counties with an enrollment share greater than 0.85. It is comparable to the estimate obtained from the full sample. This increase is largely driven by the gain/loss margin, which rises by about \$14.2 below the cutoff, suggesting that roughly 81% of the bid increase can be attributed to higher profit margins. In contrast, health care costs increase by about \$3 but are not statistically significant, while administrative costs remain close to zero. As the threshold for the main county's enrollment share decreases, allowing other counties

Figure 13. Bid Decomposition



Note: In this analysis, the bid is measured at the plan level. To express each component, health care costs, administrative costs, and the gain/loss margin on the same scale, each component is divided by a conversion factor calculated as Conversion Factor = (1 - Medicare Secondary Payer percentage) × Weighted Average Risk Factor. This factor serves as an adjustment that aligns government payments with individual risk and applicable reductions. The sum of the adjusted health care costs, administrative costs, and gain/loss margin equals the plan-level bid. Because each plan is already matched to a single county, restricting the sample to counties with an enrollment share above 0.75 yields 1,384 observations, and further restricting it to those above 0.9 leaves 1,090 observations. Given the limited within-insurer variation, I include plan-type fixed effects as a proxy for insurer fixed effects. For the same reason, standard errors are clustered at the county level.

within the plan to have greater influence, the overall pattern remains similar.

Table 10 presents the regression results based on the sample including counties where the enrollment share within a plan exceeds 0.85. This sample includes only plans in which a single county represents the vast majority of total enrollment. The outcome variable is the plan-level bid, calculated as the weighted average of each county's bid using enrollment shares as weights. By definition, the plan-level bid equals the sum of the gain/loss margin, health care costs, and administrative costs. As shown in the outcome decomposition, health care costs account for the largest portion of the bid, while the component corresponding to the gain/loss margin is relatively small. Although it may not be fully representative of all plans, the finding that the increase in bids is driven by a rise in the gain/loss margin

Table 10. Bid Decomposition

Sample: County share>0.85	(1) Plan level bid	(2) Gain/Loss margin	(3) Health care costs	(4) Administrative costs
I(<Cutoff)	17.537*	14.200*	3.794	-0.457
	(9.500)	(7.447)	(10.441)	(3.971)
Outcome mean	725.718	16.210	620.776	88.732
N	1,174	1,174	1,174	1,174

Note: The sample includes counties where the enrollment share is greater than 0.85. The outcome variable, the plan-level bid, is measured at the plan level and calculated as the weighted average of bids across the counties a plan serves. The sum of the gain/loss margin, health care costs, and administrative costs equals the plan-level bid. All regressions include the full set of controls; however, insurer fixed effects are replaced with plan type fixed effects because the number of observations is limited. Standard errors are clustered at the county level and reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

rather than in health care costs suggests that the additional payments are not passed on to enrollees as benefits but instead flow into insurers' profits.

Because MA is a regulated health insurance market, there are strict rules governing insurers' bids. In principle, each bid is calculated using the Bid Pricing Tool, which incorporates a wide range of variables affecting estimated costs. Under the medical loss ratio (MLR) requirement, each insurer must maintain an MLR of at least 85%, meaning that at least 85% of total revenue must be spent on health care services and quality improvement activities.[16] According to the regression results, the share of the gain/loss margin in the plan-level bid is approximately 2.2% on average. When incorporating the change in bids below the benchmark cutoff into the calculation, the ratio becomes 4.1%, which still falls within the range permitted under the MLR requirement.

5.5 Plan Characteristics

To investigate whether increases in bids are linked to expansions in plan generosity, I analyze plan characteristics including maximum of out-of-pocket costs and copayments for primary care, urgent care, specialist visits, outpatient visits, and ambulance services.[17] The

[16]This ratio does not correspond exactly to the share of health care costs in the bid, as certain administrative costs may also qualify the two are closely related. If CMS finds that an MA organization's MLR for a contract is below 0.85 for three or more consecutive years, the organization will not be allowed to enroll new members under that contract in the second following year. If the MLR remains below 0.85 for five consecutive years, CMS will terminate the contract in the second following year (based on § 422.510).

[17]They represent the maximum copayments, which are likely those for out-of-network services.

identification strategy mirrors that of the bid decomposition analysis in Section 5.4, as these outcomes are also observed at the plan level. To address limited variation, I restrict the sample to plan–county pairs in which a single county accounts for more than 85% of the plan's enrollment. Table 11 presents how plan characteristics change when a county is subject to a higher benchmark, that is, when it falls below the cutoff. The benchmark increases by approximately $12, and all the estimates are statistically insignificant. In particular, copayments for various health care services are effectively zero given the small magnitudes of the estimates relative to their standard errors.

Table 11. Plan characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Max OOPC	Primary care	Urgent care	Specialist care	Outpatient	Ambulance	Star Rating
I(<Cutoff)	-44.150	-0.547	0.036	-0.278	-1.390	-2.832	0.029
	(189.079)	(1.019)	(2.052)	(1.639)	(16.300)	(9.053)	(0.066)
Outcome mean	5,060.1	7.537	35.158	30.568	195.686	223.279	3.929
N	1,173	898	1,034	1,011	760	828	1,174

Note: The sample includes counties where the enrollment share is greater than 0.85. For column (2) to (6), the outcome variables are maximum copayments indicating the amount from out-of-network. Star Rating is measured at the contract level, which may include multiple plans. All regressions include the full set of controls; however, insurer fixed effects are replaced with plan type fixed effects because the number of observations is limited. Standard errors are clustered at the county level and reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

However, because not all variables are observable, it is difficult to conclude that the increase in benchmarks and bids translates into greater enrollee benefits based solely on a few observable variables. Therefore, I examine the plan's quality, measured by the star rating, as an alternative outcome. Unlike in the commercial market, where consumers evaluate plans, in the MA market the government assesses star ratings. It reflects a composite of several factors, including the rate of health care service utilization and member experience surveys, and is reported on a five-point scale in 0.5-point increments. The average star rating among the plans in this sample is 3.9. Although higher benchmarks lead to higher bids, they do not result in higher star ratings, consistent with the earlier finding that bid increases are not driven by higher health care or administrative costs. Taken together, the evidence provides little indication that increases in bids are clearly associated with enhancements in plan benefits or plan quality.

5.6 Market Entry

The policy increases benchmarks for counties in the lower TM-cost quartile to encourage private insurers' market entry, as plans in these areas face relatively lower benchmarks and not enough revenue to manage plans. At the same time, the policy reduces excessive payments in the highest TM-cost quartile by applying a multiplier of 0.95. Understanding insurer response is therefore essential to assessing the policy's effectiveness. However, as shown in the main results, the benchmark increase for counties just below the cutoff is approximately $18 per member per month, with an average projected enrollment of 350 beneficiaries. This yields an annual extra payment of about $75,600 per plan, which may be small compared with the fixed costs required to expand provider networks and administer operations in new markets. Furthermore, because quartile assignment is based on annually updated county rankings, plans face substantial uncertainty when making entry decisions, which require consideration of longer dynamics.

Table 12. Market Entry

	(1) Insurers	(2) Insurers	(3) Plans	(4) Plans
I($< Cutoff$)	0.063	0.073	0.245	0.291
	(0.169)	(0.123)	(0.399)	(0.272)
Outcome mean	6.159	6.170	13.513	13.539
N	10,003	9,981	10,003	9,981
Year FE		X		X
State FE		X		X
Insurer FE		X		X
Plan quality		X		X

To analyze whether private insurers enter the market around the cutoff, I aggregate county-level data based on enrollee shares and examine the number of insurers and plans in each county. The number of plans is defined using representative plan types employed in the main analysis, including HMO, HMOPOS, Local PPO, Regional PPO, and PFFS. Standard errors are clustered at the county level and reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

To test whether higher government payments lead to increased market participation, I aggregate the data to the county level based on the enrollment share of each plan within

a county. Table 12 reports the regression results. Column (1) shows that the number of insurers increases by 0.063 if a county is below the cutoff (i.e., receives the higher benchmark). Even after including year fixed effects, state fixed effects, and plan quality as controls, the magnitude remains approximately 0.073 and is not statistically significant. Given that the outcome mean indicates there are about six insurers per county, the effect of benchmark changes on the number of insurers appears to be close to zero. Because expected revenue gains are limited and future returns are uncertain depending on a county's position relative to the cutoff, insurers may not expand if the costs of establishing provider networks and marketing outweigh the potential benefits. This suggests that the policy did not achieve its objective of increasing competition among insurers. The number of plans also increases by about 0.25 when counties receive higher benchmarks below the cutoff and remains roughly 0.29 after controlling for year fixed effects, state fixed effects, and plan quality. Compared with the outcome mean, this represents less than 5%, and it is likewise not statistically significant.

However, even if the total number of plans does not change, this measure does not capture potential shifts in the types of plans offered. Therefore, I examine how the number of plans changes by plan type when counties receive higher benchmarks below the cutoff. Higher benchmarks provide insurers with additional resources, enabling them to offer more generous benefits that can make plans more attractive to enrollees with greater health care needs. This increased flexibility may allow insurers to better accommodate higher-risk individuals and create incentives for plans to expand their offerings. However, Table 13 shows that the response largely manifests through increases in the availability of HMO plans only. Below the cutoff with the higher benchmark, the number of HMO plans increases by 0.43, which corresponds to roughly 8% of the average of 5.36 plans per county, and it is statistically significant at the 95% confidence level. When benchmarks increase, even small gains in rebates and any improvements in benefits resulting from higher bids make plans more attractive to TM beneficiaries. Because TM beneficiaries tend to be less healthy than

current MA enrollees, insurers have a strong motive to avoid high-risk individuals. As a result, they predominantly expand HMO products, which allow tighter network control and more effective risk selection, rather than broadening offerings across other plan types.

Table 13. Market Entry by plan types

	(1) HMO	(2) HMOPOS	(3) Local PPO	(4) Regional PPO	(5) PFFS	(6) 4.5 and 5 star	(7) 3.5 and 4 star	(8) 3.0 and under
I($< Cutoff$)	0.428**	0.039	-0.045	-0.003	-0.038	0.278**	0.168	-0.012
	(0.201)	(0.052)	(0.097)	(0.011)	(0.038)	(0.138)	(0.187)	(0.094)
Outcome mean	5.360	1.018	3.365	2.922	0.865	3.168	7.911	2.433
N	10,003	9,981	9,981	9,981	9,981	10,003	10,003	10,003

To analyze whether private insurers enter the market around the cutoff, I aggregate county-level data based on enrollee shares and examine the number of insurers and plans in each county. The number of plans is defined using representative plan types employed in the main analysis, including HMO, HMOPOS, Local PPO, Regional PPO, and PFFS. Star rating is measured at 0.5 interval. From column (1) to (5), year fixed effect, state fixed effect and plan quality are are included. From column (6) to (8), only year fixed effect and state fixed effect are included. Standard errors are clustered at the county level and reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Columns (6)–(8) report how the number of plans changes below the cutoff by star rating categories. I categorize plans into three groups because the rebate share depends on these thresholds: 4.5 stars or higher, 3.5 and 4 stars, and 3 stars or below. Plans rated 4.5 and above retain 70% of the difference between the benchmark and the bid; plans rated 3.5 and 4 stars retain 65%; and plans rated 3 stars or below retain 50%, with the remainder returned to the Medicare Trust Fund. Therefore, to capture larger rebates, insurers may either improve plan quality to attract enrollees or strategically raise bids to increase revenue from the government within a given rating tier.

Consistent with this prediction, Column (6) shows that the number of high-quality plans (4.5 stars or higher) increases by 0.28 below the cutoff, which represents about 9% of the average of 3.17 per county, and the estimate is statistically significant. By contrast, the number of mid- and low-quality plans does not exhibit meaningful changes. It is important to note that these results reflect extensive margin adjustments in the availability of different types of plans, while the findings in Table 11 capture intensive margin changes within existing plans. Although plan quality within a given plan does not appear to change, existing enrollees

may still experience an improvement in quality if they switch to newly available, higher-rated plans.

Table 14. Enrollment change by plans

	(1) Total Enrollment	(2) 4.5 and 5 star	(3) 3.5 and 4 star	(4) 3.0 and under
I($< Cutoff$)	785.080	-243.891	878.552	150.419
	(1063.543)	(408.463)	(711.698)	(152.300)
Outcome mean	8201.037	2399.469	4836.610	964.958
N	3,253	3,253	3,253	3,253

To analyze whether private insurers enter the market around the cutoff, I aggregate county-level data based on enrollee shares and examine the number of insurers and plans in each county. The number of plans is defined using representative plan types employed in the main analysis, including HMO, HMOPOS, Local PPO, Regional PPO, and PFFS. Star rating is measured at 0.5 interval. Enrollment is measured using actual enrollment counts. If a plan has fewer than 10 enrollees, the value is suppressed for privacy reasons and therefore recorded as missing. Standard errors are clustered at the county level and reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 14 examines whether the increase in high-quality plans is associated with changes in enrollment. Enrollment counts are suppressed and recorded as missing when a plan has fewer than 10 enrollees due to privacy restrictions. Therefore, these cases are excluded from the analysis. In addition, some counties do not offer plans in every rating category, so I restrict the sample to county-year observations in which enrollment is observable for plans with all star ratings. Although Table 13 shows an increase in the number of high-quality plans below the cutoff, this expansion does not lead to higher enrollment in those plans. This indicates that benchmark-induced improvements in measured plan quality do not translate into changes in beneficiaries' choices, suggesting that insurers' adjustments reflect financial incentives rather than enhancements that improve plan benefits that enrollees value.

6 Discussion

The finding of this paper that bids rise almost in parallel with benchmark increases helps explain why MA spending continues to rise in the United States. The intent of the benchmark-

bid system was to encourage private insurers to compete by submitting lower bids, thereby reducing program costs and improving efficiency. However, this paper shows that when benchmarks rise, insurers choose not to compete by lowering bids to reduce premiums or expand supplemental benefits, instead tending to raise their bids to the equivalent amount. Even if this entails the risk of losing some enrollees, they prefer to secure the stable government payment associated with higher bids. This suggests that the incentive structure of the system is not functioning as intended.

One reason for this is that private insurers directly determine the payments they receive through the bids they submit, whereas the gap between the bid and the benchmark must, by regulation, be passed on to consumers. Cabral et al. (2018) find a 45% pass-through rate to premiums and a 9% pass-through to plan benefits in the MA market prior to the benchmark-bid system. In that period, government payments to private insurers were fixed, based on historical TM expenditures in each county. As these payments increased, insurers could retain the higher payment and strategically determine how much to reduce premiums depending on the competitiveness of the local market. However, under the current benchmark-bid system, if the benchmark exceeds the bid, the insurer is required to use the difference as a rebate to either reduce the premium or provide supplemental benefits. Thus, the additional payment cannot be captured directly as profit and instead must be returned to beneficiaries. While plans may benefit from increased enrollment if lower premiums or expanded benefits attract more consumers, this comes at the cost of reduced revenue per enrollee. These results suggest that private insurers prefer stable revenue per enrollee over the uncertain gains from attracting additional beneficiaries. Thus, the policy designed to lower premiums paradoxically generates the opposite outcome.

Another reason may be inertia in health insurance markets. Beneficiaries rarely switch plans because of the complexity of comparing plan options, switching costs, or loyalty to established insurers (Ericson, 2014; Handel, 2013; Heiss et al., 2021). An additional factor is that lowering bids to expand supplemental benefits may attract higher-cost beneficiaries,

raising insurers' expenses if the risk adjustment fails to fully offset this selection. Together, these mechanisms help explain why higher benchmarks do not translate into lower premiums.

This paper is subject to a few limitations. First, because it relies on an RD design, the estimates are local in nature. The magnitude of the estimated effects are similar across cutoffs, but the impact of benchmark increases on bids may vary at the lower or upper extremes of the cost distribution. Moreover, in the bid decomposition analysis, the available data on medical expenses, non-medical expenses, margins, and plan characteristics are measured at the plan level, while plans typically operate in around 14 counties on average. Thus, selecting only the county with the largest enrollment share for each plan and conducting the analysis based solely on those pairs may not yield generalizable results. Nevertheless, the magnitudes documented in this paper provide meaningful evidence on the incentive structure of the benchmark-bid system and offer important implications for potential policy improvements.

In particular, these results raise the question of whether bids themselves represent correct estimates of projected costs. An important component of the bid is the assumption of costs for the following contract year such as anticipated utilization, demographic changes, provider payments, or unit costs of treatment. These projections may serve as instruments of profit maximization rather than neutral forecasts of medical spending.

Finally, incentive structures should be designed at the plan or insurer level rather than at the county level to foster greater competition. This paper shows that applying multipliers to benchmarks based on county-level rankings of TM costs does not effectively induce insurers' market entry. This is because multipliers are determined at the county level, whereas a typical plan operates across 14 counties on average. As a result, a higher multiplier in a single county does not provide a sufficiently strong incentive from the perspective of the plan as a whole.

7 Conclusion

Medicare Advantage (MA) relies on the participation of private insurers to deliver a variety of health insurance plans, with the dual goals of expanding consumer choice and reducing program costs. However, the fact that payments have increased far more rapidly than enrollment has raised concerns over whether private insurers are gaming the MA program. To examine how private insurers adjust their bids in response to increased government payments, I exploit policy-induced variation in benchmarks across counties, stemming from quartile cutoffs based on each county's ranking of TM costs.

By pooling these cutoffs, I find that a $1 increase in the benchmark leads to a $0.88 increase in bids. This estimate remains consistent across a range of control variables and alternative specifications. I also find that market competition plays a modest role, as private insurers capture less in more competitive markets. However, the persistence of substantial bid increases even in competitive markets suggests that frictions such as consumer inertia or the complexity of plan design may reduce beneficiaries' responsiveness, thereby dampening the effectiveness of competition.

The rise in bids is driven primarily by an increase in profit margin for private insurers rather than health care costs. Therefore, there is limited evidence that the increase in bids is associated with lower premiums and better cost-sharing structures. These results suggest that insurers seek to increase revenues from stable, recurring sources while minimizing cost exposure by offering more narrow-network plans targeting healthier individuals. The current benchmark-bid system also appears to have little effect on insurer market entry. In this study, the combination of annually changing cutoff values and high fixed costs does not provide sufficient incentives to attract new insurers. However, because it is less costly for insurers to introduce or discontinue plans in counties where they already operate than to enter new markets, I find increases in the number of HMO and high-quality plans but no evidence that enrollees sort into these plans.

An important area for further investigation is whether and how insurers strategically

terminate less profitable plans and steer enrollees toward more profitable ones. When facing higher benchmarks, insurers may seek to avoid lowering premiums or expanding benefits by consolidating enrollees into more lucrative plan offerings. Understanding this behavior offers insight into how insurers can maximize profits while maintaining market concentration. Another promising direction is to explore how increases in the benchmark affect plan networks. For instance, higher government payments may allow insurers to expand their provider networks, though the extent and nature of such adjustments remain unclear.

8 Reference

References

Amber Batata. The effect of hmos on fee-for-service health care expenditures: evidence from medicare revisited. *Journal of health economics*, 23(5):951–963, 2004.

Jason Brown, Mark Duggan, Ilyana Kuziemko, and William Woolston. How does risk selection respond to risk adjustment? new evidence from the medicare advantage program. *American Economic Review*, 104(10):3335–3364, 2014.

Marika Cabral, Michael Geruso, and Neale Mahoney. Do larger health insurance subsidies benefit patients or producers? evidence from medicare advantage. *American Economic Review*, 108(8):2048–2087, 2018.

Zhun Cao and Thomas G McGuire. Service-level selection by hmos in medicare. *Journal of health economics*, 22(6):915–931, 2003.

Colleen Carey. Technological change and risk adjustment: Benefit design incentives in medicare part d. *American Economic Journal: Economic Policy*, 9(1):38–73, 2017.

Leemore Dafny, Mark Duggan, and Subramaniam Ramanarayanan. Paying a premium on your premium? consolidation in the us health insurance industry. *American Economic Review*, 102(2):1161–1185, 2012.

Leemore S Dafny. Are health insurance markets competitive? *American Economic Review*, 100(4):1399–1431, 2010.

Francesco Decarolis. Medicare part d: Are insurers gaming the low income subsidy design? *American Economic Review*, 105(4):1547–1580, 2015.

Francesco Decarolis and Andrea Guglielmo. Insurers' response to selection risk: Evidence from medicare enrollment reforms. *Journal of Health Economics*, 56:383–396, 2017.

Mark Duggan, Amanda Starc, and Boris Vabson. Who benefits when the government pays more? pass-through in the medicare advantage program. *Journal of Public Economics*, 141:50–67, 2016.

Keith M Marzilli Ericson. Consumer inertia and firm pricing in the medicare part d prescription drug insurance exchange. *American Economic Journal: Economic Policy*, 6(1):38–64, 2014.

Richard G Frank, Jacob Glazer, and Thomas G McGuire. Measuring adverse selection in managed health care. *Journal of Health economics*, 19(6):829–854, 2000.

Michael Geruso and Timothy Layton. Upcoding: evidence from medicare on squishy risk adjustment. *Journal of Political Economy*, 128(3):984–1026, 2020.

Michael Geruso and Timothy J Layton. Selection in health insurance markets and its policy remedies. *Journal of Economic Perspectives*, 31(4):23–50, 2017.

Benjamin R Handel. Adverse selection and inertia in health insurance markets: When nudging hurts. *American Economic Review*, 103(7):2643–2682, 2013.

Florian Heiss, Daniel McFadden, Joachim Winter, Amelie Wuppermann, and Bo Zhou. Inattention and switching costs as sources of inertia in medicare part d. *American Economic Review*, 111(9):2737–2781, 2021.

Kate Ho and Robin S Lee. Insurer competition in health care markets. *Econometrica*, 85

(2):379–417, 2017.

Kate Ho, Joseph Hogan, and Fiona Scott Morton. The impact of consumer inattention on insurer pricing in the medicare part d program. *The RAND Journal of Economics*, 48(4): 877–905, 2017.

Sonia Jaffe and Mark Shepard. Price-linked subsidies and imperfect competition in health insurance. *American Economic Journal: Economic Policy*, 12(3):279–311, 2020.

Mark Shepard. Hospital network competition and adverse selection: evidence from the massachusetts health insurance exchange. *American Economic Review*, 112(2):578–615, 2022.

Zirui Song, Mary Beth Landrum, and Michael E Chernew. Competitive bidding in medicare advantage: effect of benchmark changes on plan bids. *Journal of health economics*, 32(6): 1301–1312, 2013.

Karen Stockley, Thomas McGuire, Christopher Afendulis, and Michael E Chernew. Premium transparency in the medicare advantage market: implications for premiums, benefits, and efficiency. Technical report, National Bureau of Economic Research, 2014.

Erin E Trish and Bradley J Herring. How do health insurer market concentration and bargaining power with hospitals affect health insurance premiums? *Journal of health economics*, 42:104–114, 2015.

Appendix

Appendix A. Details of the Benchmark Construction

The benchmark formula and the corresponding descriptions of its components are presented below

$$\begin{aligned}\text{Benchmark} = \min\Big[&(\text{FFS rate excluding phase-out IME}) \times (1 + \text{quartile\%} + \text{bonus\%}),\\ &\max\Big(\text{Pre-ACA Rate including all IME}\\ &\quad \times \text{2018 Adjusted Growth Factor} \times \text{Risk Score Model Adjustment Factor},\\ &\quad \text{FFS USPCC} \times \text{AGA} \times (1 - \text{GME factor})\\ &\quad \times (\text{DOD Adjustment Factor} + \text{VA Adjustment Factor} - 1)\\ &\quad \times \text{Credibility Factor}\Big)\\ &- (\text{2018 Phase-out IME Dollar Amount})\Big] \end{aligned} \tag{1}$$

FFS rate excluding phase-out IME: Traditional Medicare cost net of the phased-out Indirect Medical Education (IME) payment.

Quartile% + Bonus%: Adjustment based on county spending quartile and quality bonus payments based on plan star rating.

Pre-ACA Rate including all IME: Legacy benchmark rate prior to the ACA reform, including all Indirect Medical Education payments.

Adjusted Growth Factor: Factor capturing growth in per-capita costs.

Risk Score Model Adjustment Factor: Adjustment for differences in beneficiary risk scores.

FFS USPCC × AGA: U.S. Per-capita cost (USPCC) multiplied by the annual growth adjustment.

GME factor: Reduction for graduate medical education costs.

DOD/VA Adjustment Factors: Adjustments for Department of Defense and Veterans Affairs beneficiaries.

Credibility Factor: Weight applied to stabilize estimates in small counties.

Phase-out IME Dollar Amount – Deduction reflecting the phased-out indirect medical education payment.

Table A below shows cutoff values for each 25th percentile, 50th percentile and 75th percentile.

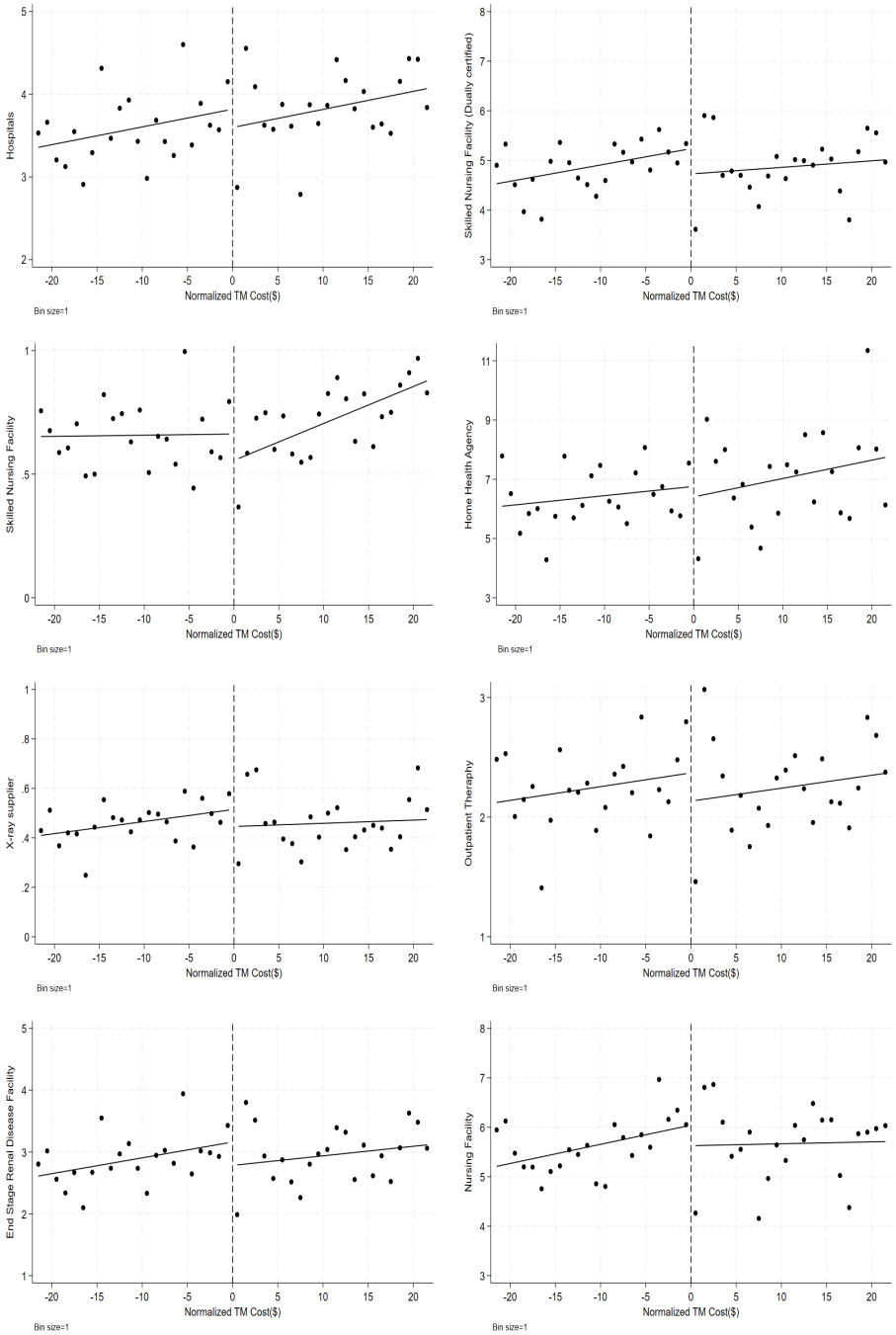
Table A1. Annual Cost Thresholds for Traditional Medicare

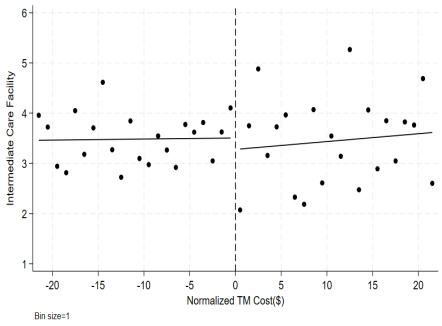
	Cutoff Value ($)					
Year	2014	2015	2016	2017	2018	2019
25th percentile	668.71	698.81	678.11	709.13	741.92	763.7
50th percentile	716.83	745.6	722.34	755.38	790.05	812.31
75th percentile	777.12	805.71	782.38	817.99	854.07	881.6

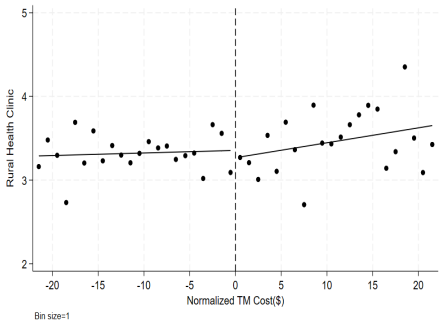
Note: The threshold values were backed out using the reported cost data for each county and their assigned quartile, by sorting county-level annual average cost for Traditional Medicare in ascending order.

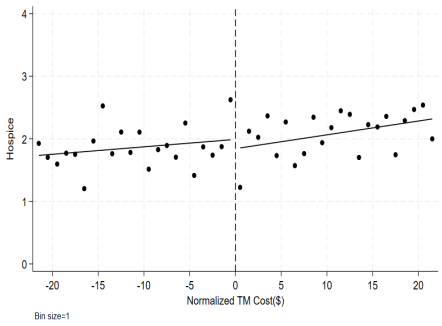
Appendix B. Balance Check Figures

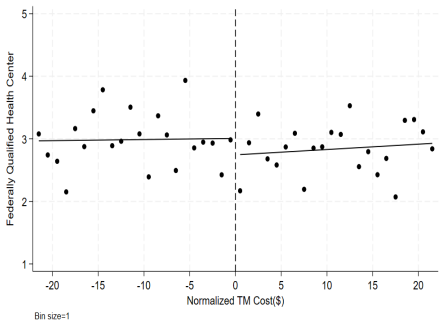
Figure B1. Balance Check Figures

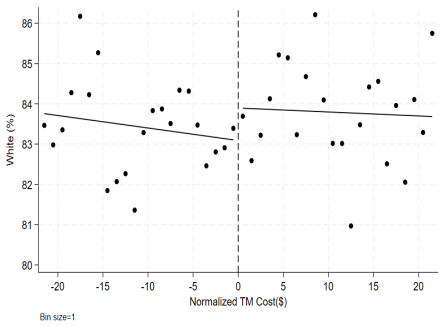


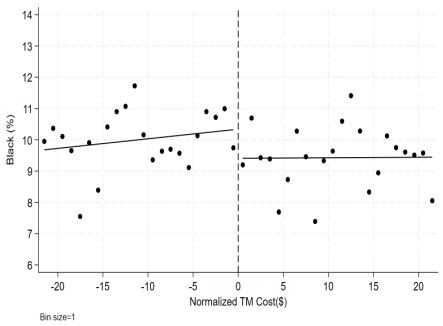

Intermediate Care Facility
Normalized TM Cost($)
Bin size=1

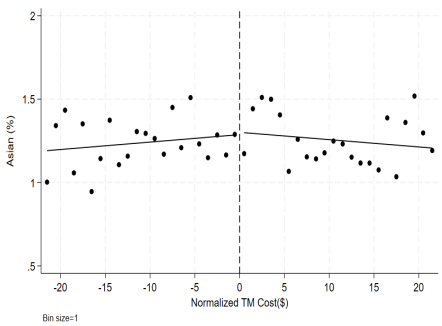

Rural Health Clinic
Normalized TM Cost($)
Bin size=1


Hospice
Normalized TM Cost($)
Bin size=1


Federally Qualified Health Center
Normalized TM Cost($)
Bin size=1


White (%)
Normalized TM Cost($)
Bin size=1


Black (%)
Normalized TM Cost($)
Bin size=1


Asian (%)
Normalized TM Cost($)
Bin size=1

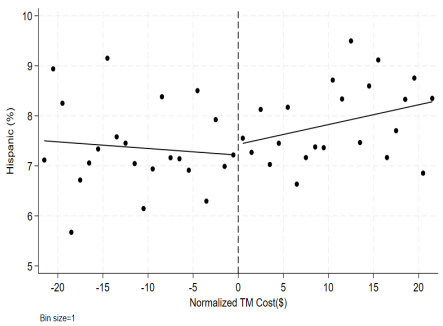
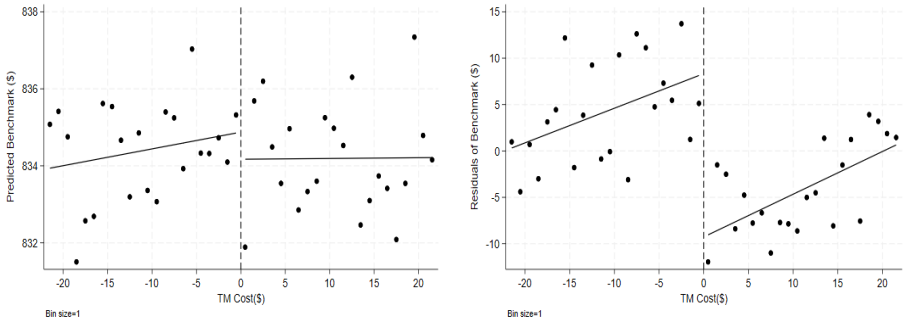

Hispanic (%)
Normalized TM Cost($)
Bin size=1

Figure B2. Predicted benchmark and Residuals of benchmark.



Appendix C. Robustness Check

C.1 Regression Discontinuity under second-order polynomials

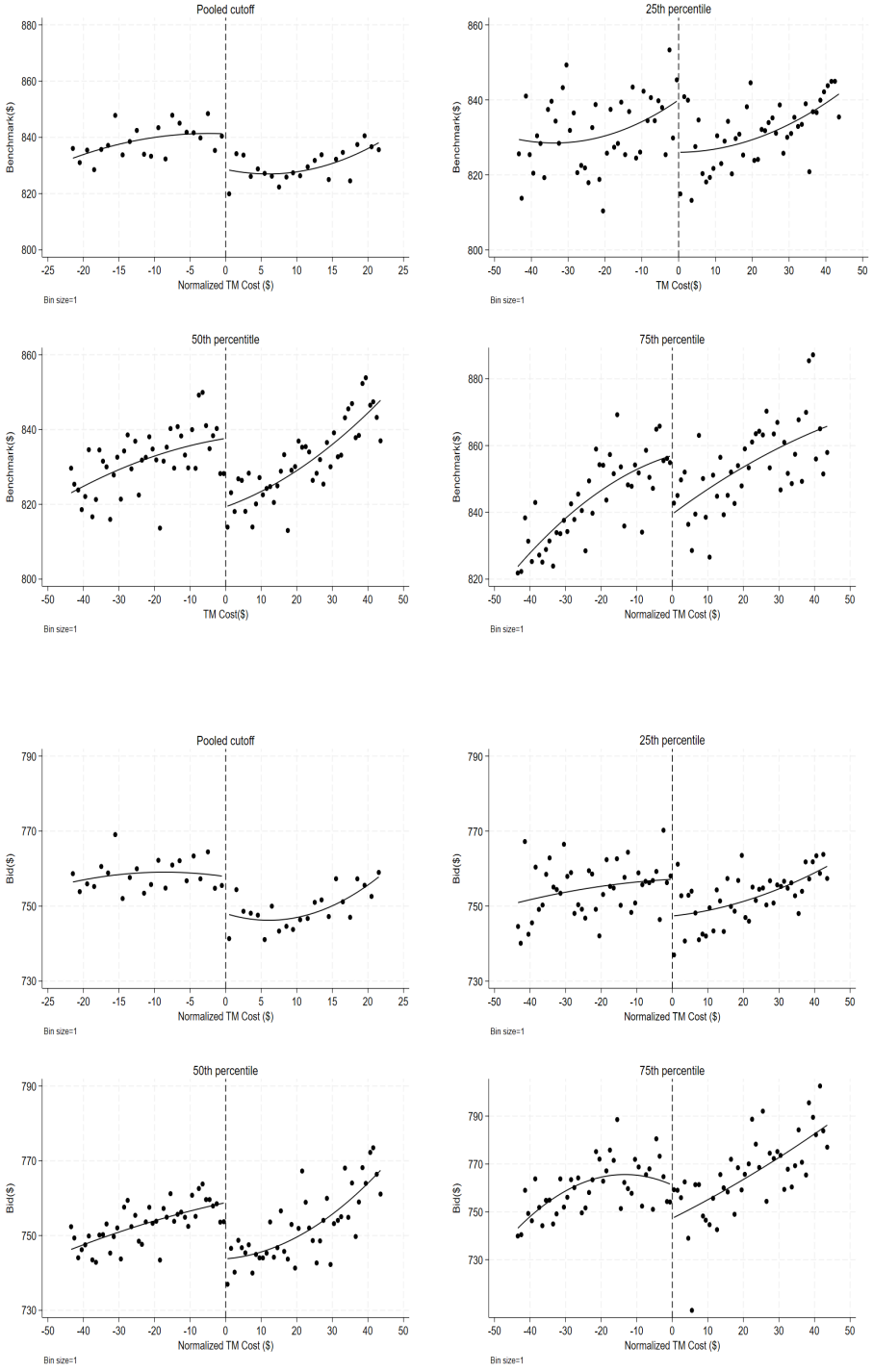


Table C1. Reduced form estimates at 25th, 50th and 75th percentile

Dep: Bid ($)	(1)	(2)	(3)	(4)
Cutoff:	Pooled cutoff	25th percentile	50th percentile	75th percentile
Benchmark ($)	0.836***	0.718***	0.834***	0.861***
	(0.140)	(0.126)	(0.179)	(0.143)
Outcome mean	753.708	753.713	752.153	760.137
R-squared	0.238	0.276	0.241	0.216
N	135,167	100,856	103,008	55,920

Note: This table reports 2SLS estimates using second-order polynomials. As in the baseline analysis, the bandwidth for the regression at the pooled cutoff is 44 dollars, and the bandwidth at the single cutoff is 88 dollars. All specifications exclude year fixed effects, state fixed effects, plan type, and plan quality controls. Standard errors, clustered at the state level, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

C2.Estimates Using the Optimal Bandwidth Method

Table C2. Estimates using the optimal bandwidth method

	1st order polynomial		2nd order polynomial	
	(1)	(2)	(3)	(4)
Specification	1st stage	2SLS	1st stage	2SLS
I(>Pooled cutoff)	-18.4***		-20.65***	
	(0.726)		(0.909)	
Benchmark		0.735***		0.802***
		(0.096)		(0.106)
N	135,167	135,167	135,167	135,165

Note: This table reports estimates obtained using the optimal bandwidth method with a triangular kernel. Columns 1 and 2 present results based on a first-order polynomial, using an optimal bandwidth of 6.596. Columns 3 and 4 report results from a second-order polynomial specification, where the optimal bandwidth is 8.97. Columns 1 and 3 report the change in Benchmark at the policy-induced threshold, while Columns 2 and 4 report the change in Bids in response to a $1 increase in Benchmark.All specifications exclude year fixed effects, state fixed effects, plan type, and plan quality controls. Standard errors, clustered at the state level, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

C3.Estimates Using Donut-hole analysis

Table C3. Estimates using Donut-hole analysis

	(1)	(2)	(3)	(4)	(5)
Dep: Bid ($)	Drop ±0.5	Drop ±1	Drop ±1.5	Drop ±2	Drop Quartile Switchers
Benchmark	0.816***	0.861***	0.890***	0.867***	0.946***
	(0.069)	(0.084)	(0.103)	(0.106)	(0.067)
Outcome means	753.805	753.947	753.922	753.914	754.189 .
N	132,160	128,935	126,025	122,784	102,314

Note: Column 1 to 4 present 2SLS estimates excluding observations near the pooled threshold. Column 1 excludes observations in the interval $[-\$0.5, \$0.5]$. In Column 5, I drop "Quartile Switcher" observations—counties that switch from one quartile to another across the threshold—because these counties face the same level of multipliers on both sides, which eliminates variation in the Benchmark. All specifications exclude year fixed effects, state fixed effects, plan type, and plan quality controls. Standard errors, clustered at the state level, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$