NBER WORKING PAPER SERIES

THE EFFECTS OF PRIVATE PLAN SUBSIDIES ON ENROLLMENT: EVIDENCE FROM MEDICARE ADVANTAGE

Paul D. Jacobs
Timothy J. Layton

Working Paper 36042
http://www.nber.org/papers/w36042

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2026

We thank Melinda Buntin (Johns Hopkins Bloomberg School of Public Health), Jeffrey Marr (Brown University), Daria Pelech (Congressional Budget Office), and Karen Stockley (Medicare Payment Advisory Commission) for helpful comments and discussions, as well as participants at the American Society of Health Economists (ASHE-con) Conference in Minneapolis, MN (June 2026). The views expressed in this paper are those of the authors and should not be interpreted as those of Johns Hopkins University, the University of Virginia, the Congressional Budget Office, or the Medicare Payment Advisory Commission. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research. Paul D. Jacobs acknowledges funding from the Peterson Center on Healthcare and the Commonwealth Fund [grant number 26085]. Timothy J. Layton acknowledges funding from Arnold Ventures, the National Institute on Aging, and the Russell Sage Foundation. The funders had no role in study design; data collection, analysis, or interpretation; writing; or the decision to submit.

At least one co-author has disclosed additional relationships of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w36042

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2026 by Paul D. Jacobs and Timothy J. Layton. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

The Effects of Private Plan Subsidies on Enrollment: Evidence from Medicare Advantage
Paul D. Jacobs and Timothy J. Layton
NBER Working Paper No. 36042
November 2026
JEL No. H44, H51, I11, I13, I18

ABSTRACT

Some social insurance programs, including the Medicare program in the United States, allow beneficiaries to choose between public and private insurance firms. In these types of programs, the government will often provide a voucher to those who opt for private firms. A key policy decision is the level of this voucher, and a key input to that decision is the elasticity of enrollment in the public versus the private option with respect to the size of the voucher. We estimate this elasticity in the Medicare program, where recent work has provided remarkably little evidence on how Medicare Advantage (MA) enrollment responds to payment despite the program's rapid growth. We leverage the fact that the formula determining county-level "benchmark" payments (vouchers) involves coarse thresholds based on county-level spending and counties occasionally move across these thresholds, providing a shock to the vouchers unrelated to beneficiary demand. Using several complementary estimation strategies, we find that a 1 percent increase in a county's benchmark raises MA enrollment by roughly half a percentage point within three years, implying an elasticity of about 2. Applying this estimate to a policy counterfactual that sets benchmarks at 95 percent of county fee-for-service spending implies a decline in MA enrollment of about 6.1 percentage points, or roughly one-quarter of our sample. We also provide suggestive evidence, consistent with prior work, that the increase in benchmarks gets passed through in the form of additional benefits, providing a mechanism for the enrollment increases

Paul D. Jacobs
Johns Hopkins University
Bloomberg School of Public Health
Department of Health Policy and Management
pauljacobs@jhu.edu

Timothy J. Layton
University of Virginia
and NBER
timothyjlayton@gmail.com

1 Introduction

Many social health insurance programs allow beneficiaries to choose between a public option and provision of benefits by private firms. When program beneficiaries are given the option to opt out of public programs and instead enroll in private health insurance plans, the government often provides some kind of "voucher" to those who opt out of the public option. In the United States, Medicare and Medicaid often consist of both public and private options, but other examples of public-private choice exist in Chile, Germany, and elsewhere (Vargas Bustamante and Méndez, 2016; Panthöfer, 2016).

A key question facing program administrators is the amount of that voucher. A larger voucher results in more beneficiaries opting out of the public plan and thus lowering aggregate public plan costs, but it also results in higher program spending in the form of the vouchers.[1] Program administrators thus want to set the voucher where the marginal benefit of an extra dollar of voucher is equal to its marginal cost. The extent to which enrollment responds to the size of the voucher – the elasticity of demand with respect to the voucher – is thus a critical parameter for policymakers trying to determine the level of the voucher for two key reasons. First, it determines how many beneficiaries switch in response to a given voucher change, and therefore its net fiscal cost. Second, a strong enrollment response contradicts the idea that higher MA payments are purely a transfer to MA insurers, as an increase in enrollment implies that at a minimum some of the surplus was shared with the marginal MA enrollees to attract them to the program.

In this paper, we estimate this parameter in one of the largest public-private health insurance programs in the world: the US Medicare program. This program provides health insurance to older adults and people with disabilities in the United States, and it has long been a pioneer of this type of joint public-private provision. The private arm of the program is known as Medicare Advantage (MA) in contrast to the public program, often referred to as Traditional Medicare (TM).

MA payment works quite similarly to the type of voucher-based payment design described above. Under the MA payment architecture, the Centers for Medicare and Medicaid Services (CMS) sets a county-level benchmark — the maximum government payment per enrollee — against which private plans submit bids. If MA plans bid higher than the benchmark,

[1]Another consideration involves any fiscal externality of the choice of a private plan, though this type of externality is more relevant when the private plan is more of a "supplement" to a public option rather than a replacement (Palme et al. 2026).

beneficiaries must pay a premium to enroll. If plans bid below the benchmark, they are required to pass through most of the difference to the beneficiary in the form of additional benefits known as rebates, which may take the form of lower cost-sharing, supplemental benefits like vision and dental, and reductions in mandatory Medicare premiums. A smaller portion of any difference is passed back to the government. Thus, the benchmark behaves like a voucher.

Despite extraordinary growth of MA enrollment over the last two decades, remarkably little is known about whether and how government payment generosity has driven this expansion. Over the last two decades, MA enrollment has increased from a low of 19% in 2007 to a high of 55% in 2026, accounting for $462 billion, or 54 percent, of total federal Medicare spending in 2024 (Freed et al., 2026; Congressional Budget Office (CBO), 2024). Estimates suggest that payment relative to public program costs has fluctuated substantially, but in recent years, when much of the MA growth has occurred, MA payments grew at an annual rate of 3.5% over 2013–23, compared with 2.6% for Traditional Medicare (Garrett et al., 2025). But are these two trends linked? Despite a large literature studying this program, there is remarkably little evidence pointing to a non-zero elasticity of demand for MA with respect to the level of the benchmark. This makes it difficult for policymakers and the Congressional Budget Office (CBO) to make accurate predictions regarding proposed reforms to MA payment, either in the positive or negative direction (CBO, 2022). This has greatly limited the ability of policymakers to make legislation around MA, because the costs or savings of this legislation are difficult to model.

Understanding how beneficiaries respond to voucher payments is also critical for evaluating the effectiveness of proposals that rely on insurer competition. In the MA context, it is not obvious that enrollees respond significantly to higher benchmark payments, as there is considerable evidence of inertia in health plan choice (Handel, 2013; Heiss et al., 2021) potentially driven by information frictions or search costs. Further, MA plans are constrained in their ability to pass through higher benchmark payments in the form of lower premiums due to the zero-premium lower bound. Related evidence suggests that several non-pecuniary factors — such as perceived plan quality, provider network breadth, and brand recognition — may weigh as heavily as price in beneficiaries' coverage decisions (Rice et al., 2014), and that it may be increasingly costly to attract enrollees via these types of benefits due to their weaker salience and to satiation of consumer preferences on these dimensions (Stockley et al., 2014). Further, the MA market is highly concentrated, resulting in little pressure to pass higher payments through to enrollees leading to the possibility that insurers, instead of consumers, capture much of the surplus from higher benchmarks (Song et al., 2013; Cabral et al., 2018; Curto

et al., 2021; Pelech and Song, 2025). Further, payments may already be so high that there are few additional benefits that insurers can provide to newly entice beneficiaries to join MA. In other words, we may already be in the "flat of the curve" of the enrollment-payment relationship. At the same time, it seems unlikely that large cuts to MA payment would not have any impact on enrollment. Yet, the current evidence base provides little guidance on this question.

We seek to fill this gap by leveraging exogenous, intertemporal variation in county-level benchmark payment rates to identify the causal effect of government payments on MA enrollment. Under the Patient Protection and Affordable Care Act of 2010 (ACA), all U.S. counties are ranked every year into quartiles based on projected TM per capita spending, with the benchmark percentage declining from 115% to 95% as counties move from the lowest- to the highest-spending quartile.

This formula generates plausibly exogenous changes over time through the annual re-ranking of counties based on updated Medicare fee-for-service (FFS) spending projections. As a result of the yearly re-ranking, counties regularly cross these quartile thresholds and receive discrete payment changes unrelated to local demand for coverage or changes in underlying health needs. These changes are not transitory: 86.7% of counties that cross a threshold after 2015 remain on the other side of it for at least three years. We exploit this persistence by identifying cohorts of counties that cross these payment thresholds over time and comparing their enrollment trajectories to those of counties that remain in their quartiles, identifying the causal effect of payment generosity on enrollment.

Ultimately, we find that a 1 percent increase in the benchmark raises MA enrollment by about half a percentage point three years after the change, an elasticity of about 2 (2.1 in our preferred event-study specification; 95% confidence interval roughly 1.6 to 2.6). Event studies show that the enrollment effects of crossing the threshold occur precisely when payments increase, with little evidence of differential pre-trends between treatment and control counties. Enrollment effects were also approximately symmetric for positive versus negative payment shocks, further validating the design. These results are supported by several alternative specifications including a simulated quartile crossing instrument constructed solely based on variation in national trends in FFS spending, a test for reverse causality, and an augmented synthetic control approach. Our key results are also supported by two separate placebo tests. The first examines counties whose benchmarks are capped, so that moving to a more generous quartile produces no increase in payment; the second examines counties that cross terciles of the FFS spending distribution instead of the quartiles used to set payment. Neither test shows an enrollment response. Further confirming our estimates, the results exhibit a dose-response relationship: counties experiencing above-median changes in their benchmark exhibit larger

effects than those experiencing below-median changes. In all cases, our results indicate a significant enrollment response to benchmarks.

Another contribution of our paper is examining *how* payment changes may translate into enrollment. We explore a range of potential explanations linking payment to enrollment, including: changes in the value of additional benefits (through plan rebates), plan entry and exit, and premiums. Our event-study estimates show that benchmark crossings change average plan rebates by about $5.50 per enrollee per month and change the number of plans by roughly two plans in the direction of the crossing,[2] consistent with insurers responding to payment changes through changes in benefits and market participation. When we split the results by direction of the crossing, the rebate response is driven almost entirely by benchmark increasers, and the entry/exit response is driven mostly by plan exit in benchmark-decreasing counties. This suggests friction in adjusting benefits downward that insurers respond to with the extreme event of closing the plan altogether, producing similar enrollment effects to changes in benefits but via a different mechanism.

In the last part of the paper, we explore the extent to which our results can help project how MA enrollment would change under an alternative policy scenario. Specifically, we use our elasticity estimate to simulate how MA enrollment would have changed during our period if payments were reduced to 95% of public program (Traditional Medicare) spending instead of the current estimated 107% (MedPAC, 2019). The simulation implies that payment reductions of this magnitude would reduce MA enrollment by about one quarter or 6.1 percentage points in our estimation sample. Thus, MA enrollment is responsive to payment, but this type of significant funding cut is unlikely to lead to the total demise of MA.

These results fill a major gap in the literature on the Medicare Advantage program. Early work in the MA context exploited changes in payment rates in the Balanced Budget Act of 1997 (Cawley et al., 2005) and discontinuities in the minimum payment floors for urban and rural counties (Duggan et al., 2016; Cabral et al., 2018) and found that enrollment generally responds to payment. More recent studies of the current benchmark system, however, have found little enrollment response to changes in payment (Schwartz et al., 2023; Murray et al., 2024). Prior work leveraging county benchmark thresholds relied on cross-sectional benchmark variation generated by these thresholds, finding no association between payment variation and MA enrollment (Murray et al., 2024). However, that variation is potentially confounded by persistent differences across counties in provider market structure, beneficiary

[2]The MA program supports several types of plans including Special Needs Plans (SNPs) that serve specific segments of the Medicare population that we have excluded from our sample because they do not face the same plan choices as most beneficiaries. As a result, throughout, we calculate changes in the number of plans among non-SNP plans.

composition, and other factors. Our paper contributes to this literature by providing robust evidence that demand-responsiveness to benchmark increases remains strong in the current era. We also provide novel evidence on how the mechanisms behind enrollment responses may differ for benchmark increases versus decreases, with increases resulting in better plan benefits and decreases resulting in exit, a result that is consistent with recent large-scale exits by MA insurers in response to slower-than-expected payment growth (Meiselbach et al., 2026; Wall Street Journal, 2025).

The remainder of the paper proceeds as follows. Section 2 provides background on the Medicare Advantage program and the relevant components of the benchmark setting process; Section 3 discusses our empirical strategies; Section 4 describes the data and sample definition; Section 5 presents our results; Section 6 discusses the implications of our findings and concludes.

2 Background

2.1 The Medicare Advantage Program

Medicare Advantage is a publicly subsidized market in which private insurers offer managed care plans to Medicare beneficiaries — primarily adults aged 65 and older and those with long-term disabilities — as an alternative to Traditional Medicare (TM). Insurers define the service area for each plan as a collection of counties, and beneficiaries may choose plans whose service area includes their county of residence. During the annual open enrollment period, beneficiaries choose between TM, which imposes few restrictions on utilization or provider choice, and one of several MA plans, which typically offer supplemental benefits not covered by TM in exchange for prior authorization requirements and limited provider networks.

2.2 MA Payment Policy

Under the MA payment architecture, CMS sets a county-level benchmark representing the maximum per-enrollee payment the government will make to a plan operating in that county. The benchmark is determined by ranking all U.S. counties into quartiles based on projected per capita FFS spending and assigning each quartile a fixed percentage of underlying FFS costs — 115, 107.5, 100, and 95 percent, declining from the lowest- to highest-spending quartile. Plans in low-cost counties thus receive benchmarks that exceed local FFS costs, while plans in high-cost counties receive benchmarks below them.

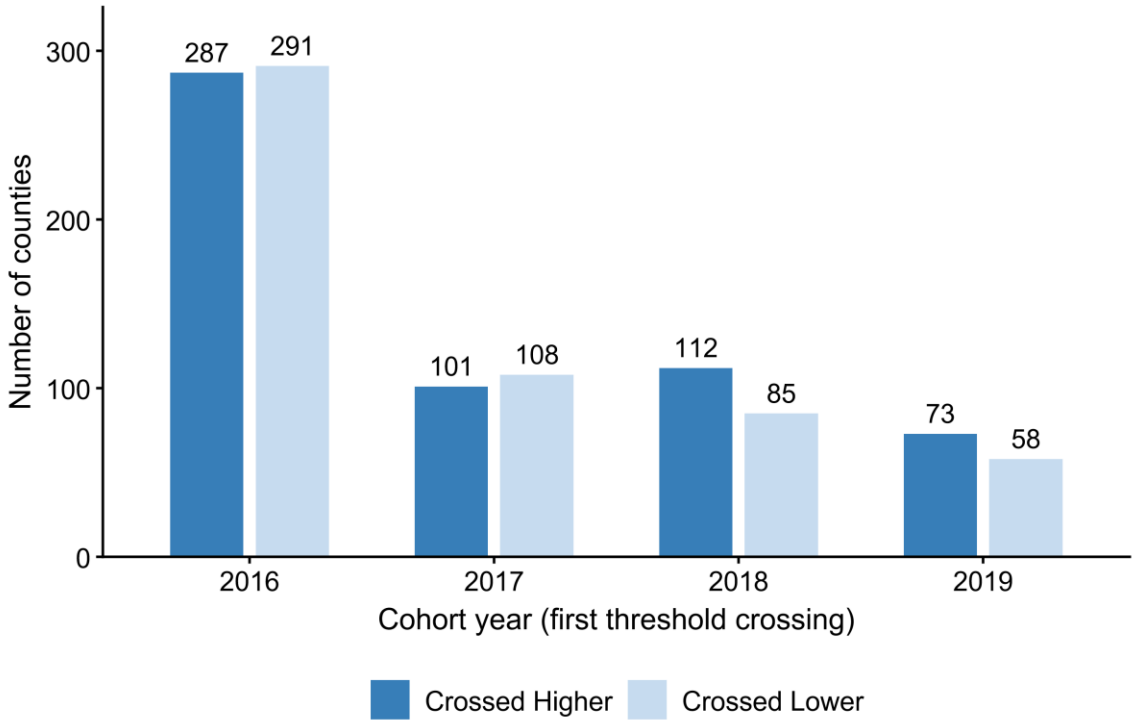
Private plans submit bids reflecting their expected cost of providing the standard TM benefit package. When a plan's bid is below the benchmark, it retains a share of the difference as a rebate, which must be returned to enrollees in the form of lower premiums for prescription drug coverage, reduced cost-sharing, or supplemental benefits — such as dental, vision, or hearing coverage. Because rebate dollars must by regulation be directed toward additional benefits, the rebate serves as a direct measure of the magnitude of benefits provided to enrollees above and beyond the standard TM benefit package. When a plan bids above the benchmark, beneficiaries pay a monthly premium equal to the difference. In practice, the vast majority of MA plans bid below the benchmark, making the rebate channel the primary margin through which payment generosity reaches enrollees.

Crucially for our identification strategy, quartile assignments are recalculated each year based on updated FFS spending projections. Counties whose projected spending crosses a quartile threshold receive a discrete change in their benchmark percentage, independent of any change in local demand for MA coverage or shifts in the underlying health of the population. These threshold crossings generate variation in payment generosity that is both plausibly exogenous and symmetric: a county whose FFS spending falls relative to a threshold may cross into a lower-spending quartile and receive a *higher* benchmark percentage, while a county whose spending rises may cross into a higher-spending quartile and experience a *reduction* in the benchmark. In our primary specifications, we estimate a single pooled effect of benchmark increases and decreases; however, we also test our hypothesis separately in each direction.

2.3 Descriptive Evidence on Quartile Crossings

Figure 1 reports the annual number of counties first crossing quartile thresholds between 2016 and 2019, measured relative to each county's assigned quartile in 2015 and disaggregated by direction of crossing. Excluding counties that cross thresholds after 2019 yields a sample of 2,370 counties over the 2015–2019 period, of which 1,115 (47.0 percent) experience at least one quartile crossing, with movements distributed roughly evenly between upward and downward directions. This frequency ensures sufficient variation to support our empirical design while confirming that threshold crossings are a regular feature of the annual recalibration process rather than isolated events.

Figure 1: Counties crossing benchmark quartiles over time, 2016-2019



Note: Number of counties by year of first quartile threshold crossing relative to each county's 2015 quartile assignment, by direction of the resulting benchmark change.

3 Empirical strategy

3.1 Two-way Fixed Effects (TWFE) Difference in Differences

As a first approximation of the effect of benchmark changes on MA enrollment and related outcomes, we estimate the canonical two-way fixed effects (TWFE) difference-in-differences model that compares outcomes in counties that cross a quartile threshold to those that do not, before and after the crossing occurs. Because a threshold crossing can raise or lower a county's benchmark, our primary specification pools upward and downward crossings into a single treatment scaled by the direction of the county's benchmark percentage change (positive for increases, negative for decreases). The pooled estimator includes two stacks: one for counties whose crossing raised their benchmark together with control counties, one for counties whose crossing lowered their benchmark together with the same control counties. This specification recovers an average post-treatment effect under the assumption of parallel trends and homogeneous treatment effects across cohorts and time periods. In Section 2.2, we relax this assumption by allowing for staggered, heterogeneous treatment effects.

We estimate:

$$Y_{cts} = \alpha_{cs} + \alpha_{ts} + \delta(\sigma_s D_{cts}) + X'_{cts}\beta + \varepsilon_{cts} \quad (1)$$

where Y_{cts} denotes the outcome in county c in year t in stack $s \in \{increase, decrease\}$; $\alpha_{cs}$and α_{ts} are county-by-stack and year-by-stack fixed effects, respectively; $D_{cts}$is an indicator equal to one if county c crossed a quartile threshold in the stack's direction by year t and zero otherwise; σ_s equals +1 in the increase stack and −1 in the decrease stack; and ε_{cts} is the error term. Scaling the treatment indicator by σ_s signs each effect in the direction of the benchmark change, so that δ measures the average response to a benchmark change, which we report throughout.

X'_{cts} is a vector of time-varying county-level covariates. In all specifications, we control for the running variable that determines quartile placement — county-year FFS per capita spending — and its square. We also control for potential time-varying determinants of healthcare supply and demand at the county level drawn from the Area Health Resources Files (AHRF) for 2015–19 (Health Resources and Services Administration, 2015-2019). From these files, we include: 1) number of active physicians per 100,000 residents; 2) total hospital beds per 100,000 residents; 3) total county population; and 4) median household income.

While the TWFE estimator provides a transparent and easily interpretable baseline, work has shown that it can yield biased estimates of the average treatment effect in staggered adoption settings when treatment effects vary across cohorts or over time (de Chaisemartin and D'Haultfœuille, 2020; Goodman-Bacon, 2021). To address this concern and recover dynamic treatment effects that are robust to heterogeneity, we turn to the staggered event study design of Sun and Abraham (2021).

3.2 Staggered Event Study Design

Because counties cross thresholds at different points over the study period, we estimate a staggered-adoption event study within each stack. We define cohorts of counties g based on the first year each county crosses into a different quartile compared to the one it was assigned in 2015.

$$Y_{cts} = \alpha_{cs} + \alpha_{ts} + \Sigma_g \Sigma_{l \neq -1} \delta_{g,l} \sigma_s \, (1[G_c = \mathrm{g}] \cdot 1[\mathrm{t} - \mathrm{g} = \mathrm{l}]) + X'_{cts}\beta + \varepsilon_{cts} \quad (2)$$

where all terms are defined as in (1) except that treatment is now parameterized as a set of cohort-by-relative-time interactions: l indexes time relative to treatment, and $\delta_{g,l}$ captures cohort-specific effects at each relative time period. As in equation (1), σ_s signs effects in the direction of the benchmark change, and we also estimate the model separately by stack. We

aggregate the cohort-specific estimates $\delta_{g,l}$ as a weighted average across cohorts at each relative time period, yielding average treatment effects that are robust to the heterogeneous treatment effects that bias conventional two-way fixed effects estimators (Goodman-Bacon, 2021).

3.3 Simulated Quartile Crossing Instrument

To further address the concern that county-level FFS spending trends may confound our estimates, we construct an alternative model using a simulated quartile crossing instrument in the spirit of Currie and Gruber (1996). The intuition is straightforward: counties near a quartile boundary may cross a threshold either because their own FFS spending changed or because the national distribution of FFS spending shifted the threshold values themselves. Only the latter source of variation is unambiguously exogenous to local enrollment dynamics, since movements in the national quartile cutoffs reflect the aggregate spending behavior of all counties and are not meaningfully influenced by any single county. We isolate this variation by freezing each county's FFS spending at its year $t-1$ value and testing whether it would nonetheless be assigned to a different quartile under the updated national thresholds in year t. Formally, we define the simulated crossing indicator:

$$\widetilde{D}_{ct} = 1\big[\, q\big(FFS_{c,t-1}, \theta_t\big) \neq q\big(FFS_{c,t-1}, \theta_{t-1}\big) \,\big] \tag{3}$$

where $FFS_{c,t-1}$is county c's per capita FFS spending in year $t-1$, $\theta_t$denotes the vector of national quartile cutoff values in year t, and $q(\cdot,\cdot)$ maps a spending level and a set of cutoffs to a quartile assignment. The indicator $\widetilde{D}_{ct}$ equals one when a county would have crossed a quartile threshold even if its own spending had not changed, isolating variation driven entirely by movements in the national thresholds. The exclusion restriction requires that shifts in the national cutoffs affect county-level MA enrollment only through their effect on the benchmark — a plausible assumption given that no individual county's spending materially influences the national quartile boundaries. We redefine benchmark-increasing cohorts using $\widetilde{D}_{ct}$ in equation (2) and estimate the reduced-form effect directly.

3.4 Addressing Endogeneity

Our identifying variation comes from counties crossing quartile thresholds because of annual changes in county level estimates of FFS spending. A potential concern is that the FFS spending trajectory itself may be correlated with outcomes of interest, leading to bias. We address this concern in four complementary ways.

First, all models control for the running variable, the county FFS spending measure CMS uses to set benchmarks, in levels and as a squared term. Because benchmarks are a deterministic function of this measure and quartile assignment, controlling flexibly for it isolates the discontinuous variation in benchmarks that arises purely from crossing a quartile threshold. Conditional on the level and trajectory of FFS spending, the only remaining source of benchmark variation is the discrete change in the applicable quartile percentage rate — precisely the policy-induced variation we seek to exploit.

A concern with controlling for FFS spending is that it may be a "bad control" in the language of Angrist and Pischke (2009). If beneficiaries select into MA and higher benchmarks raise MA enrollment, a benchmark crossing may itself change average FFS spending among the remaining beneficiaries, and controlling for it could absorb part of the effect we seek to estimate. We therefore re-estimate all models replacing the benchmark-setting measure with predicted FFS spending, constructed only from each county's spending growth in a pre-sample base period (2007–2015), which cannot respond to post-crossing changes in enrollment. The results using this approach are similar to, though somewhat smaller than, our main results (Appendix Figure A.10).

Second, we estimate alternative models to confirm that our estimates reflect the change in benchmark payment at quartile thresholds rather than a spurious correlation with changes in FFS spending. We first conduct a placebo test estimating the effect of moving to a higher-percentage quartile among counties whose benchmarks were capped by the ACA's benchmark ceiling, so that the crossing did not change the payment MA insurers actually received. In a second placebo test, we replace the statutory quartile cutoffs of the FFS spending distribution with tercile cutoffs, which carry no change in benchmark payment, but importantly test for whether the estimates from our primary specification may be driven by counties with differential rates of FFS spending growth. Additionally, we estimate models using the simulated crossings described in Section 2.3, which are driven entirely by the nationwide shift in the distribution of FFS spending rather than any individual county's spending trajectory.

Third, we implement the augmented synthetic control estimator for staggered adoption settings developed by Ben-Michael et al. (2021). This approach constructs county-specific counterfactuals by reweighting untreated counties to match each treated county's pre-treatment outcome trajectory, while an augmentation step corrects for imperfect fit using a ridge-regularized model. We include FFS spending and its square as covariates in the matching procedure. The augmented synthetic control estimator relaxes the unconditional parallel trends assumption required by standard difference-in-differences, instead requiring only that a weighted combination of untreated counties can approximate the treated county's pre-

treatment path — a weaker and arguably more credible assumption in our setting, where heterogeneous FFS spending growth across counties generates both the treatment variation and a potential source of confounding.

Fourth, as an explicit test of reverse causality, we estimate a model of threshold crossings as a function of lagged MA enrollment. We estimate a discrete-time hazard model of first crossing on lagged MA enrollment, with county and year fixed effects and controls for lagged FFS spending and its square.

Finally, while we do not observe several plausible determinants of MA enrollment, including provider network composition and the local marketing and brokerage environment, our county fixed effects absorb the persistent components of these factors. Because our identifying variation comes from intertemporal variation in counties crossing quartile thresholds, omitted determinants would confound our estimates only if they shifted discretely in the same year a county crossed.

3.5 Tests of Heterogeneity

We conduct two subgroup analyses, re-estimating our staggered event study specifications on split samples. First, to test for a dose-response relationship, we divide treated counties by whether the change in their benchmark after crossing a threshold was above or below the median, separately for counties whose benchmark increased and decreased. Larger effects among counties with above-median benchmark changes — absent differential trends in underlying FFS spending — would support a causal interpretation of our estimates. Second, to examine whether responses vary with insurer market concentration, we split counties at the median county-level Herfindahl-Hirschman Index (HHI), computed across MA parent companies serving each county in 2015. Differential responses by MA parent insurer concentration would be consistent with evidence that plans in less competitive markets capture a larger share of benchmark increases (Pelech and Song, 2025).

4 Data and Sample

4.1 Data

Our primary outcome is the MA enrollment share in each county-year, which we construct from the Medicare Beneficiary Summary Files (MBSF) for 2015-19. We restrict the denominator to beneficiaries enrolled in both Medicare Parts A and B, not enrolled in an employer-sponsored MA plan, without end-stage renal disease, and not dually eligible for Medicare and

Medicaid. These restrictions limit the sample to beneficiaries facing the standard choice between MA and TM and who did not face special incentives to enroll in particular plan types. We define MA enrollment as enrollment in an MA contract as of July of each calendar year.

We supplement the MBSF with five additional CMS data files, which we use to construct the crossing indicators and running variable and to examine mechanisms of MA enrollment. First, we construct the county-year level crossing instruments using CMS county-level ratebook files for each year (CMS, 2015-2019b). Second, we merge data from the Plan Payment Data files for 2015–19, which report plan-level amounts allocated to supplemental benefits and reduced cost-sharing ("rebates") (CMS, 2015-2019e). Third, we match county-year observations to plan-level data from the Bid Pricing Tool (BPT), available for 2015–19, which contains plan bids (CMS, 2015-2019a). Fourth, we merge average plan premiums and the number of plans from the Landscape Files for 2015-19 (CMS, 2015-2019c). Fifth, for our alternative measure of the running variable based on projected per capita FFS spending, we project forward each county's 2007–15 spending history in the CMS Geographic Variation Public Use File (PUF) to the years 2015–19 (CMS, 2015-2019d).

4.2 Sample Definition

We examine cohorts crossing county quartiles between 2016 and 2019. We limit our analysis to the 2016–19 cohorts, which enables sufficient pre- and post-period analyses without relying on data potentially impacted by the COVID-19 pandemic. All estimation is conducted at the county-year level. Our final analytic sample consists of 11,485 county-year observations spanning 2,297 counties over the 2015–2019 period, of which 573 counties (24.9%) experienced an increase in the benchmark and 519 counties (22.6%) experienced a decrease in the benchmark after crossing a quartile threshold.[3] The remaining 1,205 counties serve as never-treated comparisons.

4.3 Summary Statistics

Table 1 reports summary statistics over the 2015-19 period for never-treated counties and for benchmark increasers and decreasers in our sample. Benchmark increasers — counties whose falling FFS spending pushes them into a higher-percentage reimbursement quartile — have notably lower benchmark levels ($807.36) and FFS spending ($775.51) than both never-treated

[3]Figure 1 is based on all 2,370 counties in the CMS ratebooks that did not cross a quartile threshold after 2019, including 573 benchmark increasers, 542 benchmark decreasers, and 1,255 counties that never crossed. Our analytic sample excludes 73 of these counties (23 decreasers and 50 never-treated counties) because their county-level MA enrollment share could not be measured in every year from 2015 to 2019.

counties ($843.67 and $821.24) and benchmark decreasers ($832.16 and $815.62), whose values are more comparable to one another. Benchmark decreasers — counties whose rising FFS spending pushes them into a higher-spending quartile — have notably lower baseline MA enrollment (16.4 percent) and smaller average monthly rebates ($42.36) than both never-treated counties (23.6 percent and $52.69, respectively) and benchmark increasers (28.9 percent and $60.13), consistent with weaker MA market penetration in areas where benchmark percentages are less generous relative to underlying costs.

Table 1: Summary Statistics by Group, Full Sample, 2015-2019

Variable	Never Treated	Benchmark Increasers	Benchmark Decreasers
Benchmark ($)	843.67	807.36	832.16
FFS Spending ($)	821.24	775.51	815.62
MA enrollment Rate (share)	0.236	0.289	0.164
Rebate ($)	52.69	60.13	42.36
Number of Plans	24.82	29.54	14.18
Premium ($)	24.62	20.36	35.60
Plan Bid ($)	759.34	746.17	761.82
N (county-year obs)	6,025	2,865	2,595

Source: MBSF 2015-19, CMS Plan Payment Data 2015-19, CMS Plan Bid Data 2015-19, CMS Landscape Data 2015-19. All dollar amounts are monthly.

5 Results

5.1 MA Enrollment

Table 2 reports estimates from the simple TWFE difference-in-differences specification in equation (1). The results are directionally consistent with our hypotheses: quartile crossing resulted in a $22.94 change in the benchmark in the direction of the crossing relative to controls ($p < 0.01$) and a 0.49 percentage point increase in MA enrollment ($p < 0.01$). These estimates imply that a $100 increase in the monthly benchmark raises MA enrollment by 9.8 percent, an elasticity of 0.76.[4]

[4]Calculated from Table 2's benchmark crossing estimates and pre-treatment means among treated counties. The elasticity was calculated as (0.0049/0.2167) ÷ (22.94/774.40) = 0.76. The percent change in enrollment per $100 of benchmark semi-elasticity is: (0.0049/0.2167) ÷ (22.94/100) = 9.8 percent (rounded to the nearest tenth of a percent).

Table 2: TWFE Benchmark Crossing Difference-in-Differences Estimates, 2015-2019

	Bench-mark ($)	MA Enroll-ment Rate	Rebate ($)	Number of Plans	Premium ($)	Plan Bid ($)
Benchmark crossing	22.94***	0.0049***	1.36**	0.74***	-0.37	3.05***
	(0.67)	(0.0012)	(0.55)	(0.21)	(0.30)	(0.84)
Pre-period mean (treated)	774.40	0.2167	41.59	21.32	27.05	731.45
N	15,720	15,720	15,306	15,506	15,506	15,259
FFS controls	FFS Running Variable	FFS Running Variable	FFS Running Variable	FFS Running Variable	FFS Running Variable	FFS Running Variable

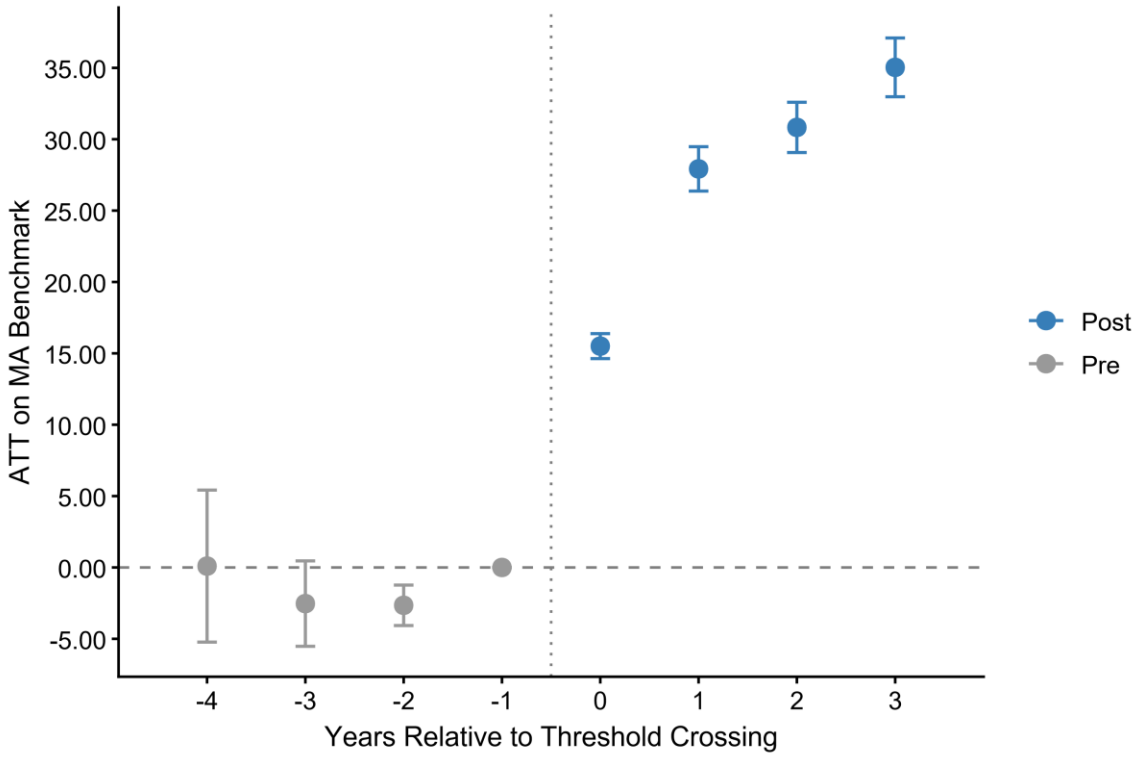
** p < 0.05; *** p < 0.01. Standard errors, clustered at the county level, are in parentheses. Stacked regression with counties whose benchmark increased stacked with sign-flipped counties whose benchmark decreased. Pre-period mean is the mean of the outcome (not sign-flipped) among treated counties in the years before their benchmark crossing. N counts county-by-stack observations over the five years; sample restrictions equivalent to Table 1. FFS Running Variable controls for the county's FFS spending rate used to rank counties into quartiles (level and square). All dollar amounts are monthly.

Figure 2 presents our "first-stage" estimates of the effect of quartile threshold crossings on county-level MA benchmarks. Counties crossing quartiles experience a sharp, persistent increase in their benchmark, rising from roughly $0 relative to untreated counties prior to the crossing to $15.51 at t = 0 and $35.03 at t = 3 higher in the post period. The effect at $t = 0$ is approximately half the magnitude of the $t = 1$ estimate, consistent with the CMS payment rule that in the first year a county crosses a threshold, its benchmark percentage is set equal to the average of the old and new quartile rates. After that, the benchmark effect levels out at the higher level.

Figure 3 presents the corresponding event study estimates for the MA enrollment rate estimated from equation (2). Pre-treatment estimates are statistically indistinguishable from zero at all pre-period lags, supporting the parallel trends assumption. Post-treatment, MA enrollment rises steadily in the direction of the benchmark change, with effects near zero at t=0 and building to approximately 2.1 percentage points by t=3. Scaling this t = 3 effect by the Figure 2 benchmark effect implies that a $100 increase in the monthly benchmark raises MA enrollment by 27.4 percent, an elasticity of 2.12.[5] This pattern of increasing effects over time is consistent with (1) the initial phase-in of the new benchmark in year t=0 and (2) a large stock of inertial enrollees who only occasionally look at how their options have changed and react via switching.

[5]Calculated from Figure 3 and Figure 2's t = 3 estimates and pre-treatment means among treated counties. The elasticity was calculated as (0.0208/0.2167) ÷ (35.03/774.40) = 2.12. The semi-elasticity, the percent change in enrollment per $100 of benchmark, is (0.0208/0.2167) ÷ (35.03/100) = 27.4 percent.

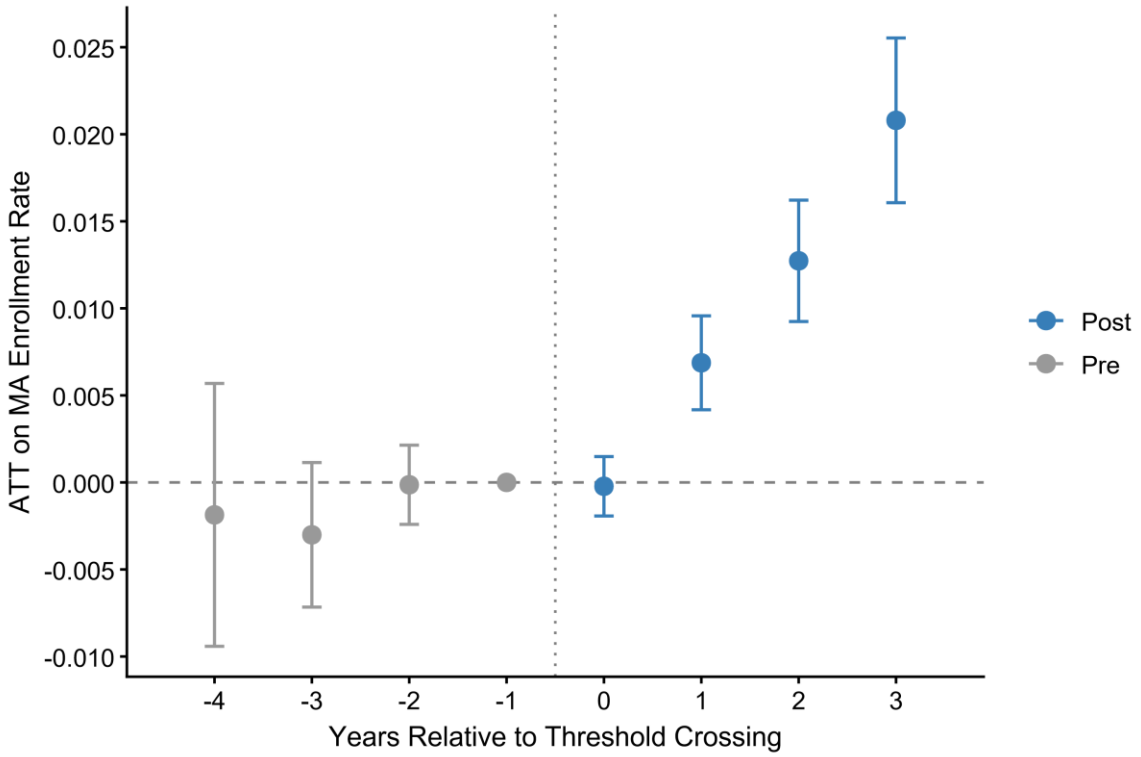
Figure 2: Event Study Estimates of Effect of Benchmark Threshold Crossing on MA Benchmark



Note: Sun and Abraham (2021) staggered event-study estimates. Estimates pool benchmark increases and decreases and are signed in the direction of the benchmark change. ATT = average treatment effect on the treated. Reference year is t = −1. Bars show 95% confidence intervals based on standard errors clustered by county.

Figure 4 presents event study estimates of the effect of threshold crossing on MA enrollment rates by direction of the benchmark change that induced the crossing. Pre-period estimates are indistinguishable from zero in both panels (though the decreasers show some signs of a slight pre-trend, consistent with the first-stage trends in Appendix Figure A.5). Post-crossing, participation rises for increasers and falls for decreasers, reaching about +1.7 and -2.5 percentage points, respectively, at t=3. Participation moves in the direction our design predicts in both panels, rising where benchmarks rise and falling where they fall, which is the sign restriction the research design implies and which supports a causal reading of the pooled estimate. The two effects are also of broadly comparable magnitude in percentage points, and we cannot reject that they are equal and opposite at the 5% level ($p = 0.096$).

Figure 3: Event Study Estimates of Effect of Benchmark Threshold Crossing on MA Enrollment Rate



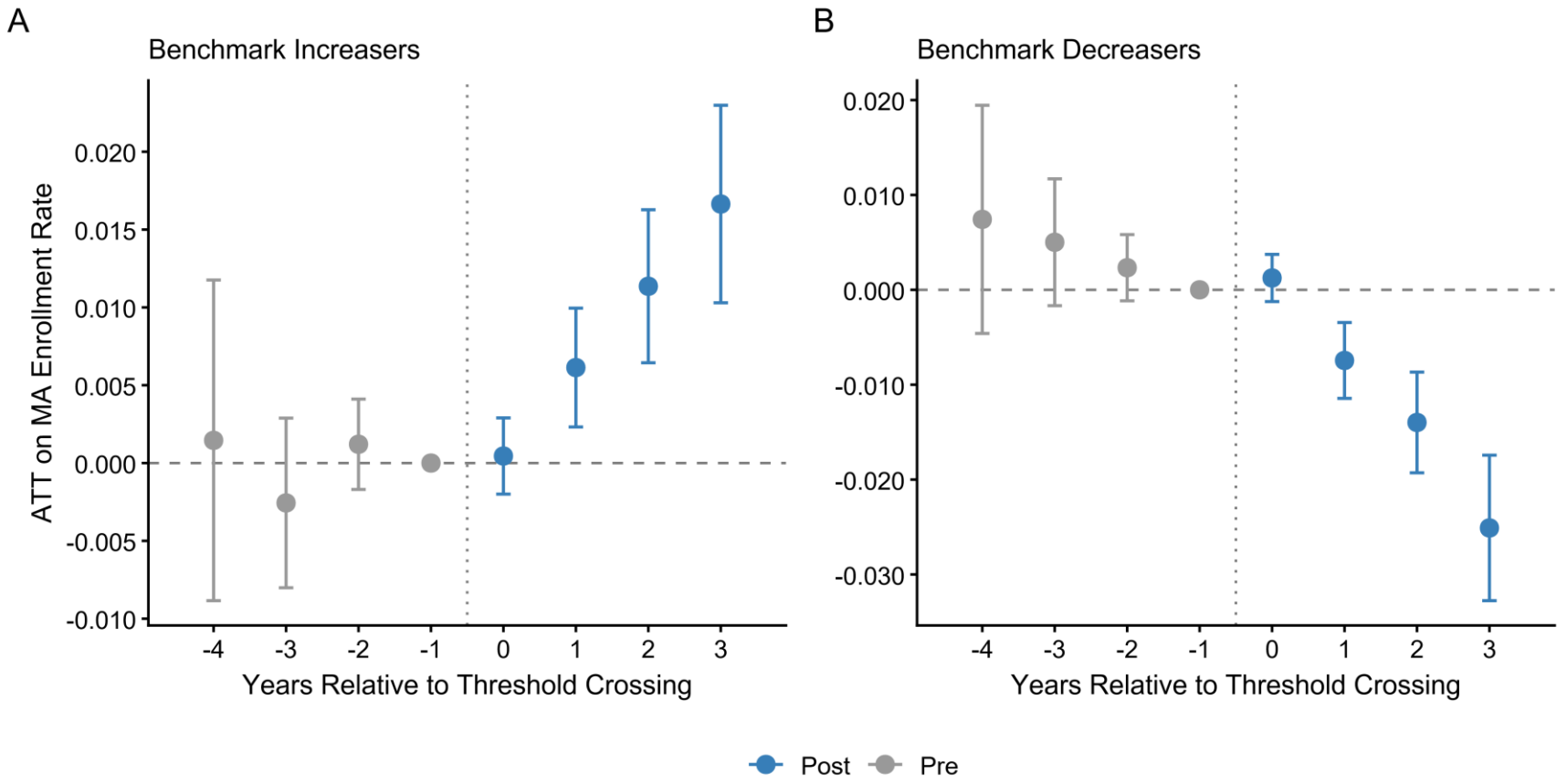
Note: Sun and Abraham (2021) staggered event-study estimates. Estimates pool benchmark increases and decreases and are signed in the direction of the benchmark change. ATT = average treatment effect on the treated. Reference year is t = −1. Bars show 95% confidence intervals based on standard errors clustered by county.

Table 3 summarizes the MA enrollment estimates from the baseline, preferred, robustness, and placebo specifications discussed below. A key concern is that the changes in FFS spending that trigger the quartile crossing and thus the benchmark increase themselves cause the increase in MA enrollment. To test whether this is the case, we identify a set of counties that cross the threshold (due to the same changes in FFS spending) but *do not* receive an increase in their benchmark. Appendix Figure A.6 reports this placebo test for counties whose benchmarks were capped by the ACA ceiling, where crossing into a higher-percentage quartile does not raise the payment MA insurers receive. In these counties, post-crossing estimates of MA enrollment are small, change sign across years, and are statistically indistinguishable from zero at every horizon with no upward trend. Pre-period estimates are also small and largely not statistically different from zero. Participation does not respond when a quartile crossing leaves payments unchanged, suggesting that our main results are driven by payment changes rather than secular trends in FFS spending. This test provides convincing evidence of the validity of our design.

Appendix Figure A.7 reports a second placebo test, which replaces the statutory quartile cutoffs with tercile cutoffs of the FFS spending distribution that carry no change in benchmark payment. Participation shows no response to tercile crossings: estimates are small and are statistically indistinguishable from zero in every pre- and post-period. Beneficiaries do not respond when their county moves across other arbitrary cutoffs in the FFS spending distribution, which supports attributing the response in the main estimates to the payment change at the statutory cutoffs.

Appendix Figure A.8 presents event study estimates of the effect on MA enrollment using the simulated quartile crossing instrument, which isolates shifts in the national quartile cutoffs while holding county-specific FFS spending fixed across years. Pre-treatment estimates are indistinguishable from zero, consistent with the identifying assumption that movements in national quartile cutoffs are uncorrelated with county-level enrollment trends prior to the crossing. Post-treatment, participation rises over time, reaching about 1.0 percentage points by t=3, with estimates statistically significant from t=1. The timing mirrors the main results, while the magnitude is roughly half that of the corresponding estimate in Figure 3 (about 2.1 percentage points).

Figure 4: Event Study Estimates of Effect of Benchmark Threshold Crossing on MA Enrollment Rate



Note: Sun and Abraham (2021) staggered event-study estimates, estimated separately for counties whose benchmark increased (Panel A) and decreased (Panel B). ATT = average treatment effect on the treated. Reference year is t = −1. Bars show 95% confidence intervals based on standard errors clustered by county.

Table 3: Stability Analyses

	Baseline	Preferred	Robustness			Placebo tests	
	TWFE	Event Study	Synthetic Control	Projected FFS Control	Simulated Crossing	Capped Benchmark Counties	Terciles of FFS Spending
Post-period average	0.0049***	0.0080***	0.0055***	0.0044***	0.0048**	-0.0016	-0.0003
	(0.0012)	(0.0014)	(0.0015)	(0.0014)	(0.0018)	(0.0039)	(0.0036)
t = 3 coefficient	-	0.0208***	0.0116***	0.0148***	0.0098***	0.0018	0.0020
		(0.0024)	(0.0035)	(0.0025)	(0.0030)	(0.0062)	(0.0069)
Elasticity†	0.76	2.12	1.18	1.50	0.93	0.19	0.19
Semi-elasticity per $100 (%)†	9.8	27.4	15.3	19.4	11.9	2.5	2.5
Pre-period mean	0.217	0.217	0.217	0.217	0.235	0.210	0.238
N (2015-2019)	15,720	15,663	15,700	15,663	8,934	4,317	11,088

** $p < 0.05$; *** $p < 0.01$. Standard errors, clustered at the county level, are in parentheses. Outcome is MA enrollment. Columns pool increases with sign-flipped decreases except the simulated-crossing and capped-benchmark columns, which cover increases only. † Elasticities and semi-elasticities are calculated from the t = 3 coefficients in all specifications except TWFE, which uses its static post-period coefficient. For these calculations, the enrollment effect is scaled by the Figure 2 benchmark effect, so columns differ only in the enrollment response. Enrollment effects are expressed relative to each column's pre-period mean among treated counties; the semi-elasticity is the percent change in enrollment per $100 of benchmark. Placebo tests: Capped Benchmark Counties examines counties whose benchmark is capped, where moving to a more generous quartile does not raise payment; Terciles of FFS Spending examines crossings of FFS spending terciles, which carry no payment change.

To test for reverse causality – whether increases in MA enrollment induce a county to cross the threshold, possibly due to corresponding changes in the TM risk pool – we estimate whether the probability of a county crossing a quartile threshold is driven by prior MA growth. As shown in Appendix Table A.1, a within-county increase in lagged MA enrollment of one standard deviation is associated with an insignificant change of about +0.3 percentage points in the probability of an upward crossing and about −0.4 percentage points for a downward crossing, relative to baseline crossing rates of roughly 5 percent, providing evidence that crossings are not driven by prior changes in MA enrollment.

Appendix Figure A.9 presents augmented synthetic control event studies of the effect of benchmark threshold crossing on MA enrollment. Pre-period estimates are indistinguishable from zero. Post-crossing, MA enrollment rises from near zero at t=0 to a peak of about 1.2 percentage points at t=3, corroborating the direction and significance of our main event study estimates. Appendix Figure A.10 re-estimates the main event study control-ling for projected FFS spending, based on each county's 2007-2015 spending history, rather than the FFS spending measure used for benchmark setting in our main specification. The results closely track the main estimates: participation rises to about 1.5 percentage points by t=3, compared with about 2.1 percentage points in the main specification.

Our results exhibit a dose-response pattern consistent with a causal interpretation, as evidenced in Appendix Figures A.11-A.12. Counties with above-median benchmark increases show a larger (average post-period effect = 1.7 percentage points) and statistically significant enrollment response relative to below-median counties (average post-period effect = 0.6 percentage points; $p < 0.01$ for difference; Appendix Figure A.11). Enrollment declines were likewise larger for counties with steeper decreases in the benchmark (average post-period effect = -2.1 percentage points) compared with below-median counties (-1.3 percentage points; $p < 0.05$ for difference; Appendix Figure A.12).[6]

Prior work has found that benefit pass-through increases with competition, suggesting enrollment effects could be stronger in more competitive markets (Cabral et al. 2018). To assess whether responses to benchmark changes vary with market concentration, we compared results for counties above and below the median HHI, measured at the MA parent-company level (Appendix Figures A.15-A.16). All comparisons across outcomes and event-

[6]Among benchmark increasers, FFS spending actually falls more in below-median counties than in above-median counties (about −$95 versus −$81 between t=−4 and t=3; Appendix Figure A.13), yet above-median counties show a larger enrollment response, suggesting that the response reflects the size of the benchmark change rather than the underlying FFS spending trend. Among decreasers, FFS spending rises roughly 1.7 times as much in above-median counties (about +$74 versus +$43; Appendix Figure A.14), so the dose-response pattern for decreasers cannot be fully separated from differences in FFS spending growth.

time periods were statistically insignificant, indicating broadly similar responses regardless of market concentration. For benchmark increases, the 95 percent confidence interval for the difference between more and less competitive markets is −1.5 to 0.6 percentage points at t = 3. The upper bound implies that the differential response in more competitive markets is at most 0.6 percentage points larger. That represents approximately 30 percent of our main effect of 2.1 percentage points, indicating that responses in more competitive markets are, at most, moderately larger than in less competitive markets. One possible explanation for the similar responses is that insurers in both types of markets already bid above costs at baseline, leaving less room to respond at the margin.

Our results provide new quasi-experimental evidence that government payment generosity causally increases MA enrollment, a relationship that most prior work has been unable to establish convincingly. Studies of earlier payment floors found enrollment responses; our elasticity of about 2 is somewhat larger than, but in the same general range as, the 1.6 implied by Cabral et al. (2018), suggesting that beneficiary sensitivity has persisted as the program has grown. More recent studies, however, have found little response. Murray et al. (2024), the study most closely related to ours, used cross-sectional variation in quartile thresholds but found no association between payment and enrollment. By exploiting repeated crossings within counties over time, our design absorbs time-invariant county differences and draws on many more threshold events, improving precision. Schwartz et al. (2023) also found a null result using a difference-in-differences approach leveraging the ACA's benchmark reductions. But the ACA changes coincided with secular increases in coding intensity and the expansion of quality bonus payments, both of which independently increased effective plan payments and may confound estimates from that policy episode. Our design avoids this concern by exploiting ongoing, annual threshold crossings within a stable policy regime, where the only source of benchmark variation conditional on FFS spending is the discrete change in the applicable quartile percentage rate.

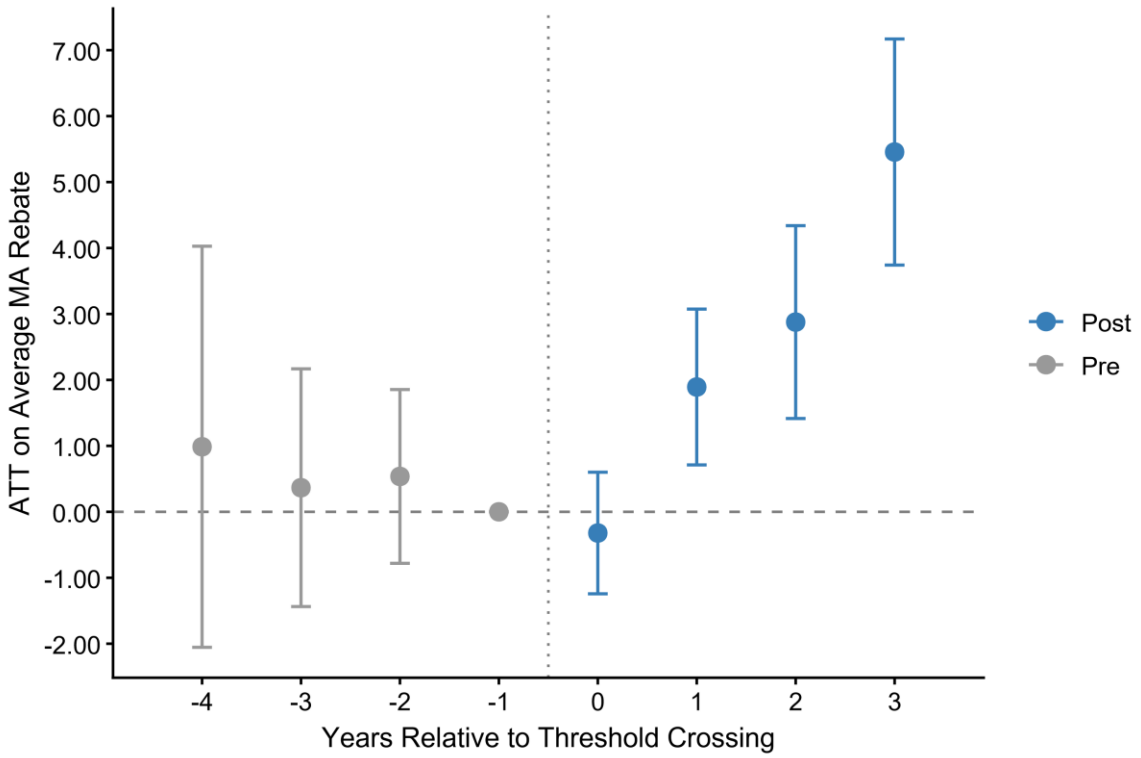
5.2 Explanations of MA Enrollment

While our results make it clear that benchmark changes have significant impacts on MA enrollment, it remains unclear how these changes actually affect beneficiaries. We thus explore the mechanisms behind the enrollment changes.

Figure 5 examines rebates, our proxy for the generosity of additional benefits, as a potential mechanism linking benchmarks to enrollment growth. Pre-treatment estimates are statistically indistinguishable from zero, supporting a causal interpretation. The post-treatment effect is not statistically distinguishable from zero at t=0, then rises steadily, reaching

approximately $5.46 per enrollee per month by t=3. Appendix Figure A.1 shows that the results in Figure 5 are largely driven by counties where benchmarks increased (Panel A), where rebates rise steadily from about $0.90 at t=0 to about $9.10 at t=3. Three of the four post-period estimates for counties where benchmarks decreased (Panel B) show no statistically significant rebate change.

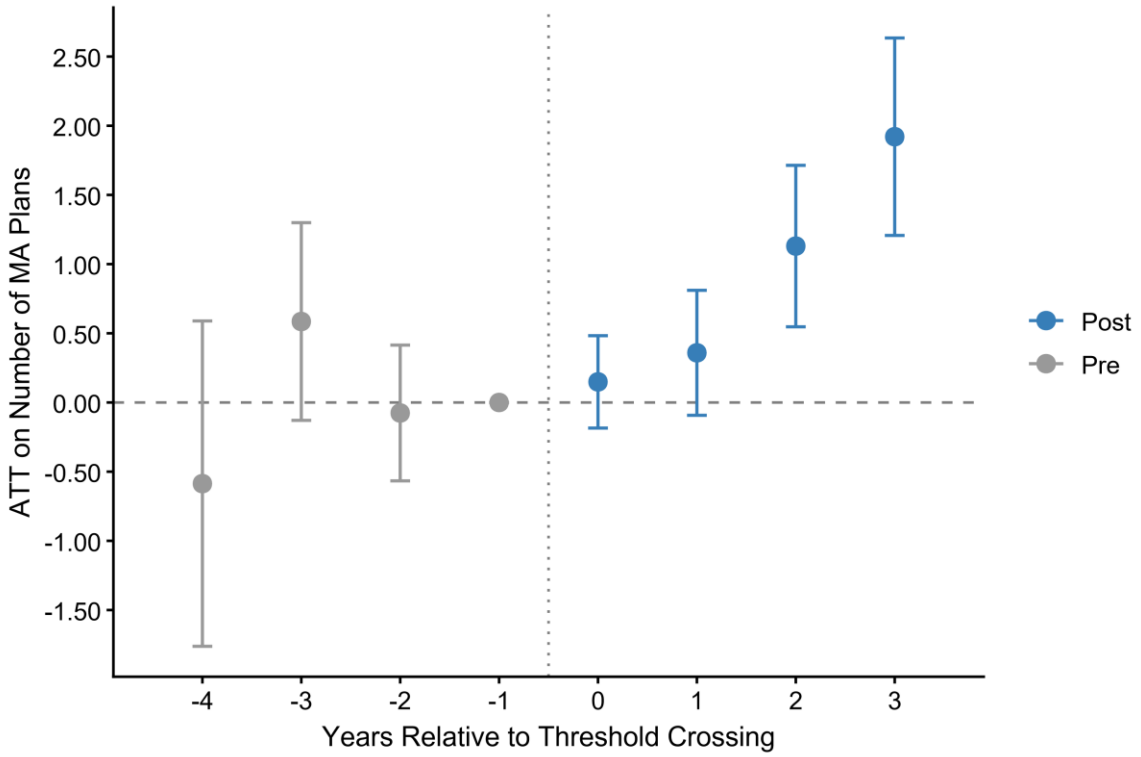
Figure 5: Event Study Estimates of Effect of Benchmark Threshold Crossing on Average MA Rebate



Note: Sun and Abraham (2021) staggered event-study estimates. Estimates pool benchmark increases and decreases and are signed in the direction of the benchmark change. ATT = average treatment effect on the treated. Reference year is t = −1. Bars show 95% confidence intervals based on standard errors clustered by county.

Figure 6 presents the pooled event study estimate of the effect of benchmark threshold crossing on the number of MA plans operating in a county. Pre-treatment estimates are statistically indistinguishable from zero. Post-treatment, the effect on the number of plans grows in the direction of the benchmark, from about 0.1 at t=0 to approximately 1.9 plans at t=3. Appendix Figure A.2 shows that the response is larger and statistically significant in benchmark-decreasing counties, where the number of plans falls by about 2.7 by t=3, and smaller and not statistically significant in benchmark-increasing counties, where it rises by about 1.1 plans by t=3.

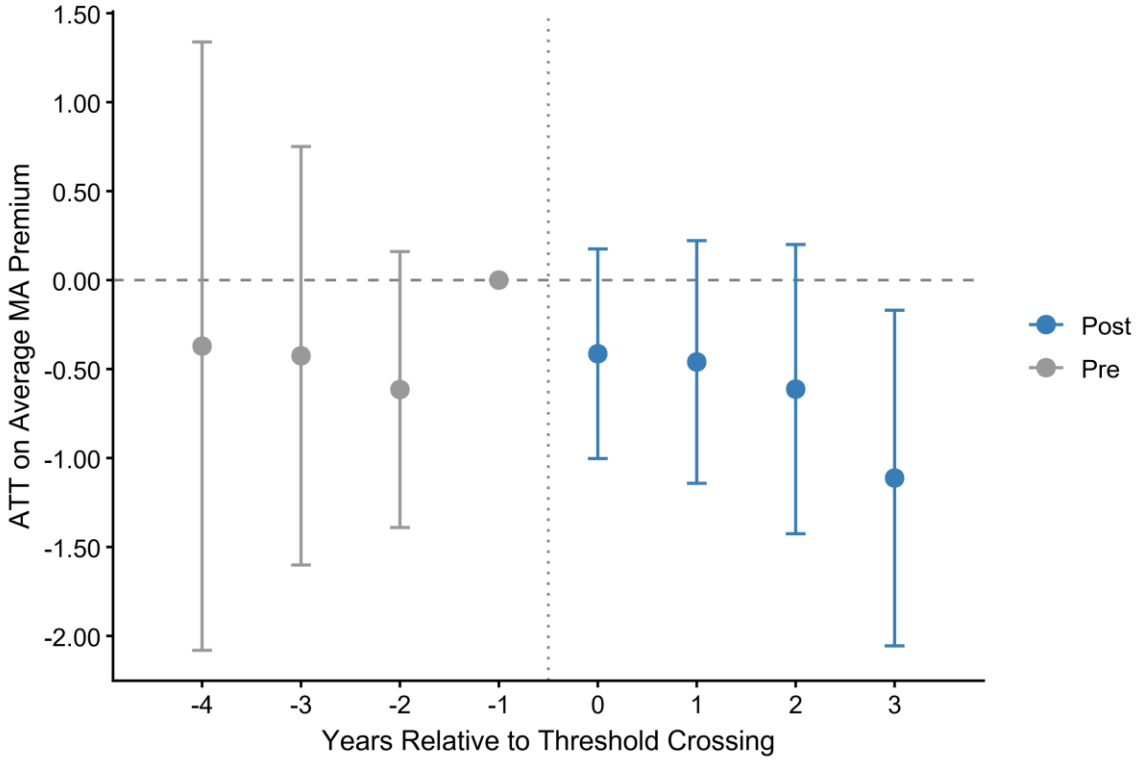
Figure 6: Event Study Estimates of Effect of Benchmark Threshold Crossing on Number of MA Plans



Note: Sun and Abraham (2021) staggered event-study estimates. Estimates pool benchmark increases and decreases and are signed in the direction of the benchmark change. Plan counts exclude special needs plans. ATT = average treatment effect on the treated. Reference year is t = −1. Bars show 95% confidence intervals based on standard errors clustered by county.

Figure 7 presents the pooled event study estimate of the effect of benchmark threshold crossing on average monthly MA premiums. Pre-treatment estimates are statistically indistinguishable from zero. Post-treatment, the effect is muted and imprecisely estimated before t=3 with a -$1.11 negative point estimate at t=3. Appendix Figure A.3 shows no consistent effect on premiums in benchmark-increasing counties and an increase of about $2.20 per month in benchmark-decreasing counties emerging at t=3.

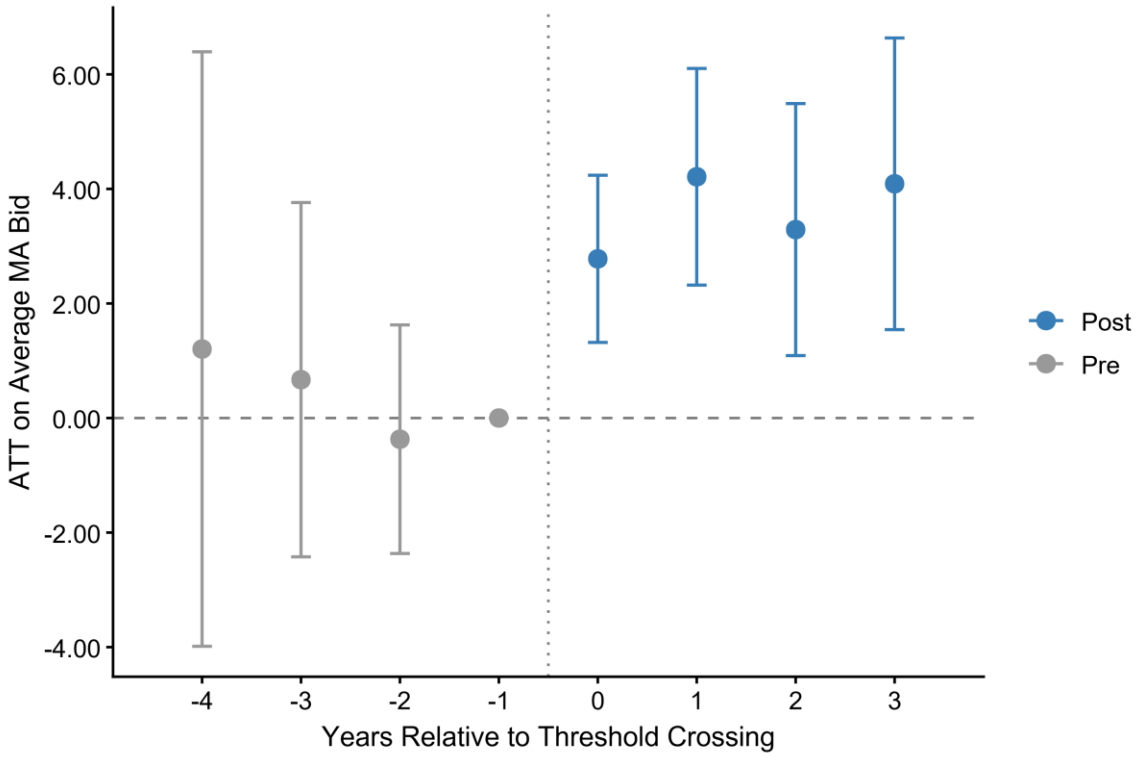
Figure 7: Event Study Estimates of Effect of Benchmark Threshold Crossing on Average MA Premium



Note: Sun and Abraham (2021) staggered event-study estimates. Estimates pool benchmark increases and decreases and are signed in the direction of the benchmark change. ATT = average treatment effect on the treated. Reference year is t = −1. Bars show 95% confidence intervals based on standard errors clustered by county.

Figure 8 presents the pooled event study estimate of the effect of benchmark threshold crossing on the average MA bid. Pre-treatment estimates are not statistically different from zero. Post-treatment, bids move in the direction of the benchmark change, with an increase of about $2.80 at t=0 and estimates of roughly $3.30 to $4.20 at t=1 through t=3. Appendix Figure A.4 shows a similar pattern in both groups: bids rise by about $2.50 to $4.50 in benchmark-increasing counties and fall by about $3 to $6 in benchmark-decreasing counties. Some of these estimates are imprecise and some confidence intervals include zero (t=3 for increasers and t=2 for decreasers).

Figure 8: Event Study Estimates of Effect of Benchmark Threshold Crossing on Average MA Bid



Note: Sun and Abraham (2021) staggered event-study estimates. Estimates pool benchmark increases and decreases and are signed in the direction of the benchmark change. ATT = average treatment effect on the treated. Reference year is t = −1. Bars show 95% confidence intervals based on standard errors clustered by county.

These results indicate that while enrollment responds to both payment increases and decreases, the mechanisms differ by direction. Benchmark decreases do not produce a corresponding reduction in rebates, and instead plans consider the more extreme action of market exit as a primary vehicle for adjusting to lower margins. Thus, the competitive dynamics of MA markets may operate differently in the two directions: plans compete to attract enrollees through expanded benefits and plan entry when resources expand, but prefer to exit markets and increase premiums when resources contract. These results may indicate more friction in adjusting benefits downward than upward. From a policy standpoint, this pattern suggests that reducing benchmark generosity may operate in part through plan exits that disrupt some beneficiary choices rather than solely through incremental adjustments to benefits received. The magnitude of this channel is likely modest, however: the average beneficiary had access to about 32 plans in 2026, so the loss of one or two plans in a county leaves substantial choice in most markets. Recent experience is nonetheless consistent with market exit as one way insurers respond when margins compress (Meiselbach et al., 2026).

We find some evidence that plan bids shadow changes in benchmarks, consistent with a substantial body of evidence documenting incomplete pass-through of government subsidies in MA. In our main specification, bids rise by roughly $3 to $4 per enrollee per month following a benchmark increase, alongside rebate increases of about $5.50. However, this bid response is less robust than our rebate and enrollment results and is attenuated or not statistically significant in some alternative specifications. Song et al. (2013) found that plans increase bids by roughly half of an increase in the benchmark, implying that only about half of the subsidy reaches enrollees through rebates. Pelech and Song (2025) confirm a bid response of approximately $0.57–$0.62 per dollar of benchmark change and show that plans in less competitive markets capture a larger share of benchmark increases, consistent with the exercise of market power. To the extent our bid estimates are smaller or less precise than those in prior work, one explanation is that MA has become more competitive over time (Layton et al., 2026), leading to more pass-through than in prior periods. Another explanation is that plans submit one bid for a service area that typically spans multiple counties, so a crossing in one county shifts the payment a plan faces by only a fraction of the county-level change. This attenuation likely accounts for much of the gap between the $35 benchmark increase and the $4 bid response.

6 Discussion

In social insurance programs with choice between public and private options, the size of the voucher paid for those who opt out is *the* primary instrument with which policymakers can influence enrollment and spending in the program. Set the voucher too high and program spending can spiral out of control. Set the voucher too low, and the program will fail to achieve the savings and efficiency gains that are the primary goal of allowing the private option in the first place. A critical parameter for policy design is thus the extent to which enrollment responds to the size of the voucher. This responsiveness is also a signal of who benefits from MA payment increases. If enrollment did not respond to payment, then it is possible that higher payments are purely transfers to MA insurers, or at most, transfers to MA insurers and current MA enrollees. But the significant enrollment responses we estimate suggest that at least some of the surplus is being shared with marginal enrollees, since they switch into MA only when higher payments make it the better choice.

Our results provide new evidence on that relationship in the Medicare Advantage program. Prior work has been remarkably silent on this question, especially in recent years. Our estimates indicate a significant demand response: for every $100 per-member-per-month

increase in the benchmark, MA enrollment grows by 5.9 percentage points (27.4 percent relative to the baseline penetration rate of 21.7%), implying a semi-elasticity of enrollment with respect to the benchmark of 0.274 per $100. Alternatively, a 1 percent increase in the benchmark increases enrollment by 2.1 percent (0.5 percentage points), an elasticity of enrollment with respect to the benchmark of 2.1. Further, we find enrollment effects for benchmark increases as well as decreases, suggesting that cuts to MA payments can have real effects on the size of the MA program.

To gauge the magnitude of this response, we use our estimates to simulate a policy counterfactual in the spirit of recent reform proposals: setting benchmarks at 95 percent of local FFS spending and eliminating the quality bonus program. In 2019, benchmarks including quality bonuses averaged 107 percent of FFS spending (MedPAC, 2019), so this policy would lower benchmarks by 12 percentage points of FFS spending, or approximately 11.2 percent. Applying our estimated elasticity of 2.1, this implies a 23.5 percent reduction in MA enrollment relative to baseline. At the 2019 mean MA enrollment of 26.1% among counties in our sample, that corresponds to a decline of 6.1 percentage points or roughly one quarter.

These magnitudes place our estimates between two views in the existing literature. Benchmark reductions could meaningfully contract the MA program, contrary to evidence of no enrollment response to the ACA cuts (Schwartz et al., 2023), but roughly three-quarters of baseline enrollment would remain under a relatively large cut to benchmarks. This is far from the large contraction in MA enrollment projected when the ACA payment reductions were enacted (Foster, 2010). Our estimate places a quantitative bound on the enrollment consequences that accompany the federal savings such proposals would generate.

These results have important implications for policymakers considering reforms to the MA program. The paucity of evidence for a link between payment and MA growth in recent years suggested that MA growth stemmed from exogenous factors rather than policy design. If this were the case, reforms that cut payments to MA plans could be predicted to have little effect on MA enrollment. Our results suggest otherwise. Instead, we find that enrollment is quite responsive to payment. This link is important not just for predicting the effects of MA payment reforms but also interpreting those effects. If enrollment responses to payment changes are as substantial as our results indicate, beneficiaries are likely experiencing real increases in surplus when MA payments go up. Similarly, beneficiaries are likely to experience real decreases in surplus when payments go down. This suggests that MA payment cuts are no "free lunch" with beneficiaries likely suffering as a result. Instead of assuming that cuts to MA plans can be made without any trade-off, policymakers will need to grapple with the question

of whether the savings generated by payment cuts exceed the corresponding loss in consumer surplus.

Reform of MA payment is likely in the future. Our results provide policymakers with solid estimates of what those payment reforms are likely to do to MA enrollment and Medicare costs.

References

Angrist, J.D., Pischke, J.-S., 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, Princeton, NJ.

Ben-Michael, E., Feller, A., Rothstein, J., 2021. The augmented synthetic control method. Journal of the American Statistical Association 116 (536), 1789–1803.

Cabral, M., Geruso, M., Mahoney, N., 2018. Do larger health insurance subsidies benefit patients or producers? Evidence from Medicare Advantage. American Economic Review 108 (8), 2048–2087.

Cawley, J., Chernew, M.E., McLaughlin, C., 2005. HMO participation in Medicare+Choice. Journal of Economics & Management Strategy 14 (3), 543–574.

[dataset] Centers for Medicare and Medicaid Services, 2015–2019a. Medicare Advantage bid pricing tool (BPT) data. Available at: https://www.cms.gov/medicare/payment/medicare-advantage-rates-statistics/data.

[dataset] Centers for Medicare and Medicaid Services, 2015–2019b. Medicare Advantage ratebooks and supporting data. Available at: https://www.cms.gov/medicare/payment/medicare-advantage-rates-statistics/ratebooks-supporting-data.

[dataset] Centers for Medicare and Medicaid Services, 2015–2019c. Medicare Advantage and Part D landscape source files. Available at: https://www.cms.gov/medicare/coverage/prescription-drug-coverage.

[dataset] Centers for Medicare and Medicaid Services, 2015–2019d. Medicare geographic variation public use file: national, state, and county. https://data.cms.gov/summary-statistics-on-use-and-payments/medicare-geographic-comparisons/medicare-geographic-variation-by-national-state-county.

[dataset] Centers for Medicare and Medicaid Services, 2015–2019e. Plan payment data. Available at: https://www.cms.gov/medicare/health-drug-plans/plan-payment-data.

Congressional Budget Office, 2022. Reduce Medicare Advantage benchmarks. In: Options for Reducing the Deficit, 2023 to 2032—Volume I: Larger Reductions. https://www.cbo.gov/budget-options/58626 (accessed September 16, 2026).

Congressional Budget Office, 2024. Medicare: CBO's June 2024 baseline projections. Congressional Budget Office, Washington, DC. https://www.cbo.gov/system/files/2024-06/51302-2024-06-medicare.pdf (accessed October 7, 2026).

Currie, J., Gruber, J., 1996. Health insurance eligibility, utilization of medical care, and child health. Quarterly Journal of Economics 111 (2), 431–466.

Curto, V., Einav, L., Levin, J., Bhattacharya, J., 2021. Can health insurance competition work? Evidence from Medicare Advantage. *Journal of Political Economy* 129 (2), 570–606. https://doi.org/10.1086/711951

de Chaisemartin, C., D'Haultfœuille, X., 2020. Two-way fixed effects estimators with heterogeneous treatment effects. American Economic Review 110 (9), 2964–2996.

Duggan, M., Starc, A., Vabson, B., 2016. Who benefits when the government pays more? Pass-through in the Medicare Advantage program. Journal of Public Economics 141, 50–67.

Foster, R.S., 2010. Estimated financial effects of the "Patient Protection and Affordable Care Act," as amended. Memorandum, Centers for Medicare and Medicaid Services, Office of the Actuary, Baltimore, MD, April 22, 2010.

Freed, M., Biniek, J.F., Damico, A., Ochieng, N., Neuman, T., 2026. Medicare Advantage in 2026: Enrollment Update and Key Trends. KFF, San Francisco, CA. https://www.kff.org/medicare/medicare-advantage-in-2026-enrollment-update-and-key-trends/ (accessed 16 September 2026).

Garrett, B., Holahan, J., Zuckerman, S., 2025. Medicare at 60: Many successes but more work to do. Journal of Health Politics, Policy and Law 50 (4), 573–602.

Goodman-Bacon, A., 2021. Difference-in-differences with variation in treatment timing. Journal of Econometrics 225 (2), 254–277.

Handel, B.R., 2013. Adverse selection and inertia in health insurance markets: When nudging hurts. American Economic Review 103 (7), 2643–2682. https://doi.org/10.1257/aer.103.7.2643

[dataset] Health Resources and Services Administration, 2015–2019. Area Health Resources Files (AHRF). U.S. Department of Health and Human Services, Rockville, MD. https://data.hrsa.gov/topics/health-workforce/nchwa/ahrf (accessed April 15, 2026).

Heiss, F., McFadden, D., Winter, J., Wuppermann, A., Zhou, B., 2021. Inattention and switching costs as sources of inertia in Medicare Part D. American Economic Review 111 (9), 2737–2781. https://doi.org/10.1257/aer.20170471

Layton, T.J., Maini, L., McWilliams, J.M., 2026. Substitutes for success? Public versus private competition in Medicare Advantage. Journal of Economic Perspectives 40 (2), 143–170. https://doi.org/10.1257/jep.20251469

Medicare Payment Advisory Commission, 2019. Report to the Congress: Medicare Payment Policy, March 2019. MedPAC, Washington, DC.

Meiselbach, M.K., Lavallee, M., Zahn, M., Xu, J., Polsky, D., 2026. Forced disenrollments among Medicare Advantage beneficiaries following 2026 plan exits. *JAMA* 335 (10), 861–862. https://doi.org/10.1001/jama.2026.0028

Murray, R.C., Meyers, D.J., Fuse Brown, E.C., Williams, T.C., Ryan, A.M., 2024. Association between the Medicare Advantage quartile adjustment system and plan behavior and enrollment. JAMA Health Forum 5 (1), e234822.

Palme, M., Seim, D., Spinnewijn, J., Wikström, J., 2026. The impact and incidence of supplemental health insurance: Evidence from Sweden. Working Paper, London School of Economics. https://personal.lse.ac.uk/spinnewi/SHI_draft.pdf

Panthöfer, S., 2016. Risk selection under public health insurance with opt-out. Health Economics 25 (9), 1163–1181.

Pelech, D., Song, Z., 2025. Private Medicare plans' responses to benchmark changes and competition before and after the Affordable Care Act's payment cuts. Health Services Research 60 (2), e14392.

Rice, T., Jacobson, G., Cubanski, J., Neuman, T., 2014. The private health insurance choices of Medicare beneficiaries: How much does price matter? Medical Care Research and Review 71 (6), 661–689. https://doi.org/10.1177/1077558714556896

Schwartz, A.L., Kim, S., Navathe, A.S., Gupta, A., 2023. Growth of Medicare Advantage after plan payment reductions. JAMA Health Forum 4 (6), e231744.

Song, Z., Landrum, M.B., Chernew, M.E., 2013. Competitive bidding in Medicare Advantage: Effect of benchmark changes on plan bids. Journal of Health Economics 32 (6), 1301–1312.

Stockley, K., McGuire, T.G., Afendulis, C.C., Chernew, M.E., 2014. Premium transparency in the Medicare Advantage market: Implications for premiums, benefits, and efficiency. NBER Working Paper No. 20208. National Bureau of Economic Research, Cambridge, MA. https://doi.org/10.3386/w20208.

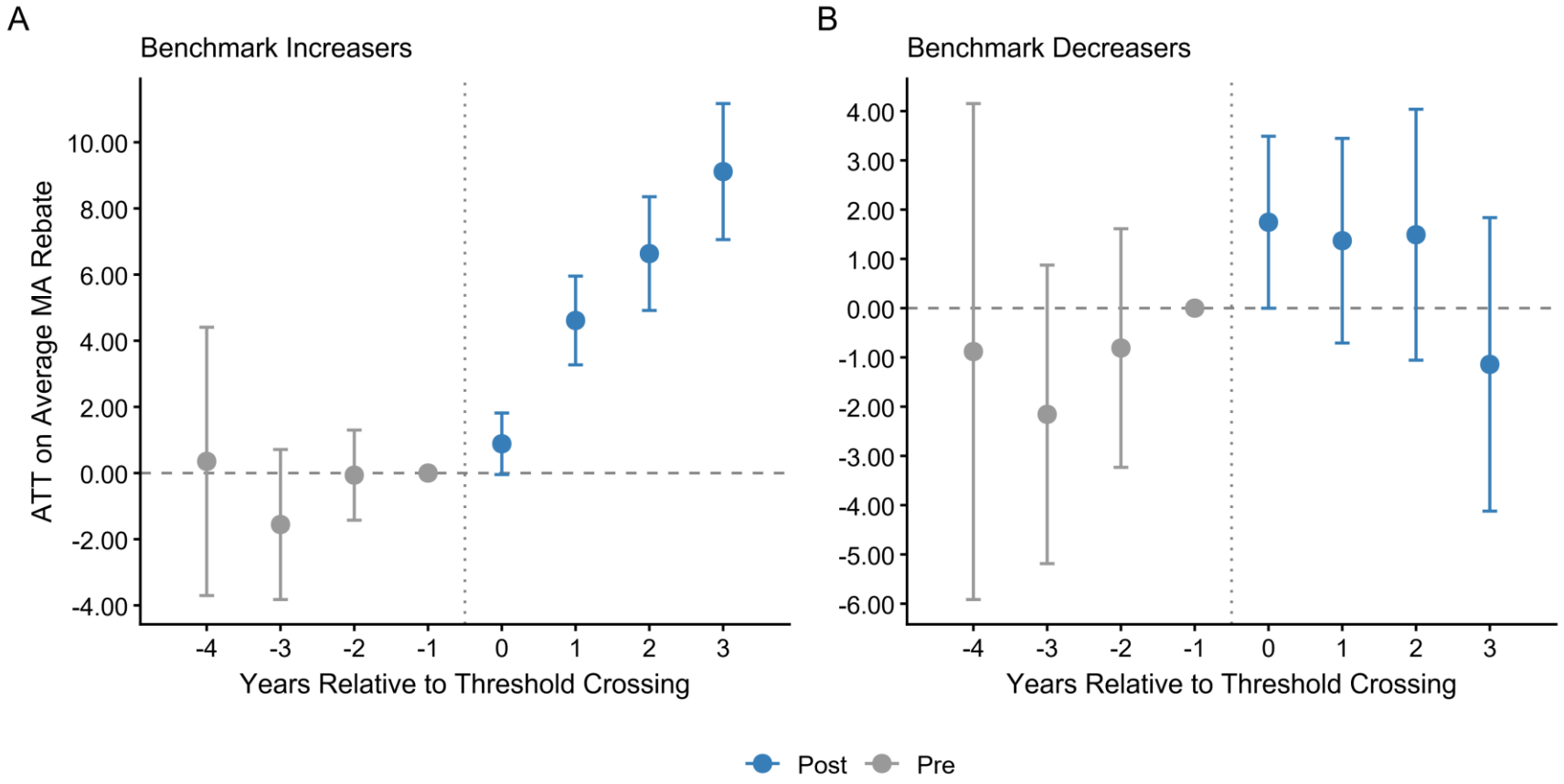
Sun, L., Abraham, S., 2021. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. Journal of Econometrics 225 (2), 175–199.

Vargas Bustamante, A., Méndez, C.A., 2016. Regulating self-selection into private health insurance in Chile and the United States. International Journal of Health Planning and Management 31 (3), e219–e234.

Wall Street Journal, 2025. The Medicare pullback is here. *Heard on the Street*, August 5, 2025. https://www.wsj.com/health/healthcare/the-medicare-pullback-is-here-e6ab37da (accessed November 4, 2026).

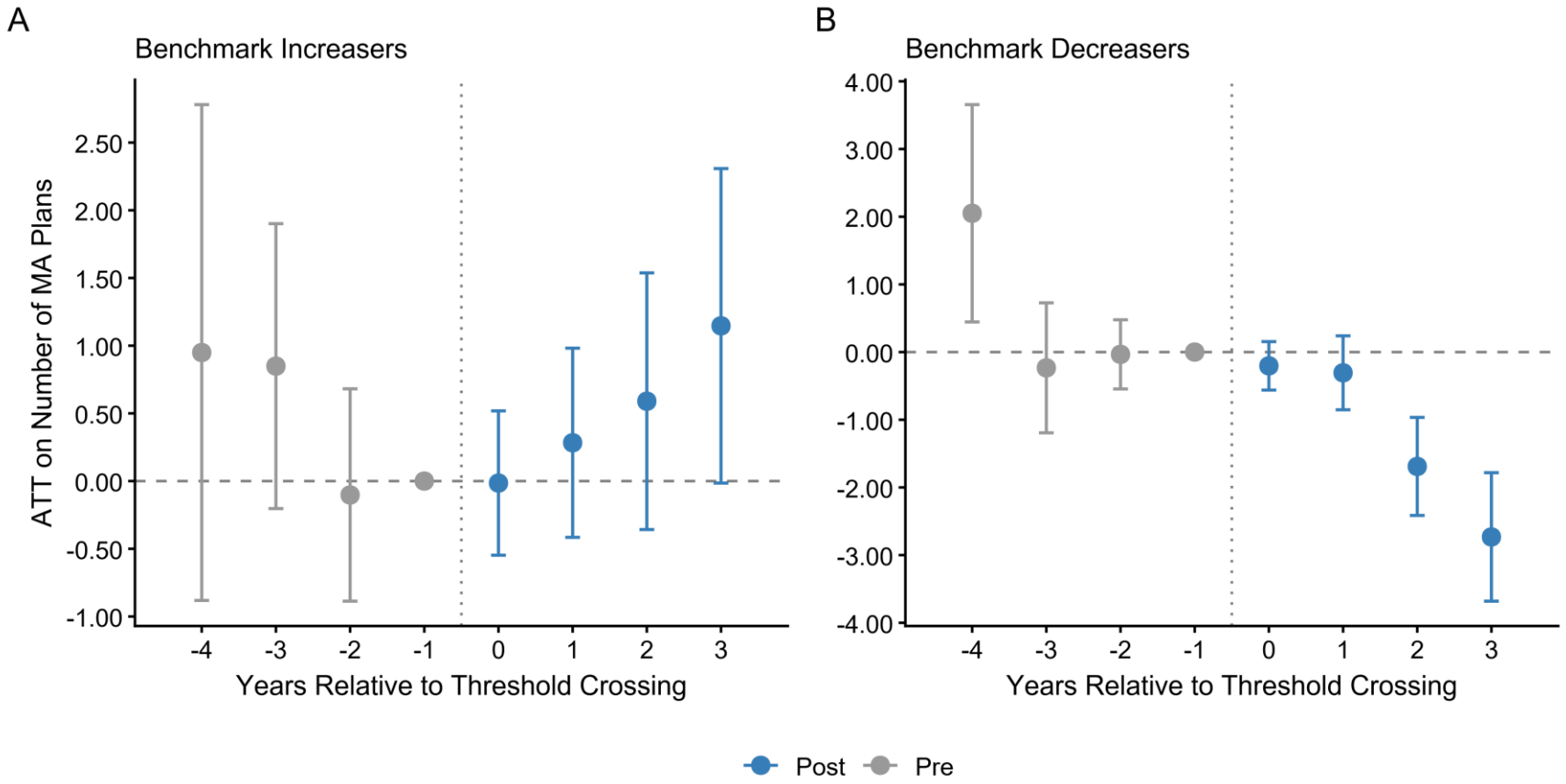
Appendix

Figure A.1: Event Study Estimates of Effect of Benchmark Threshold Crossing on Average MA Rebate



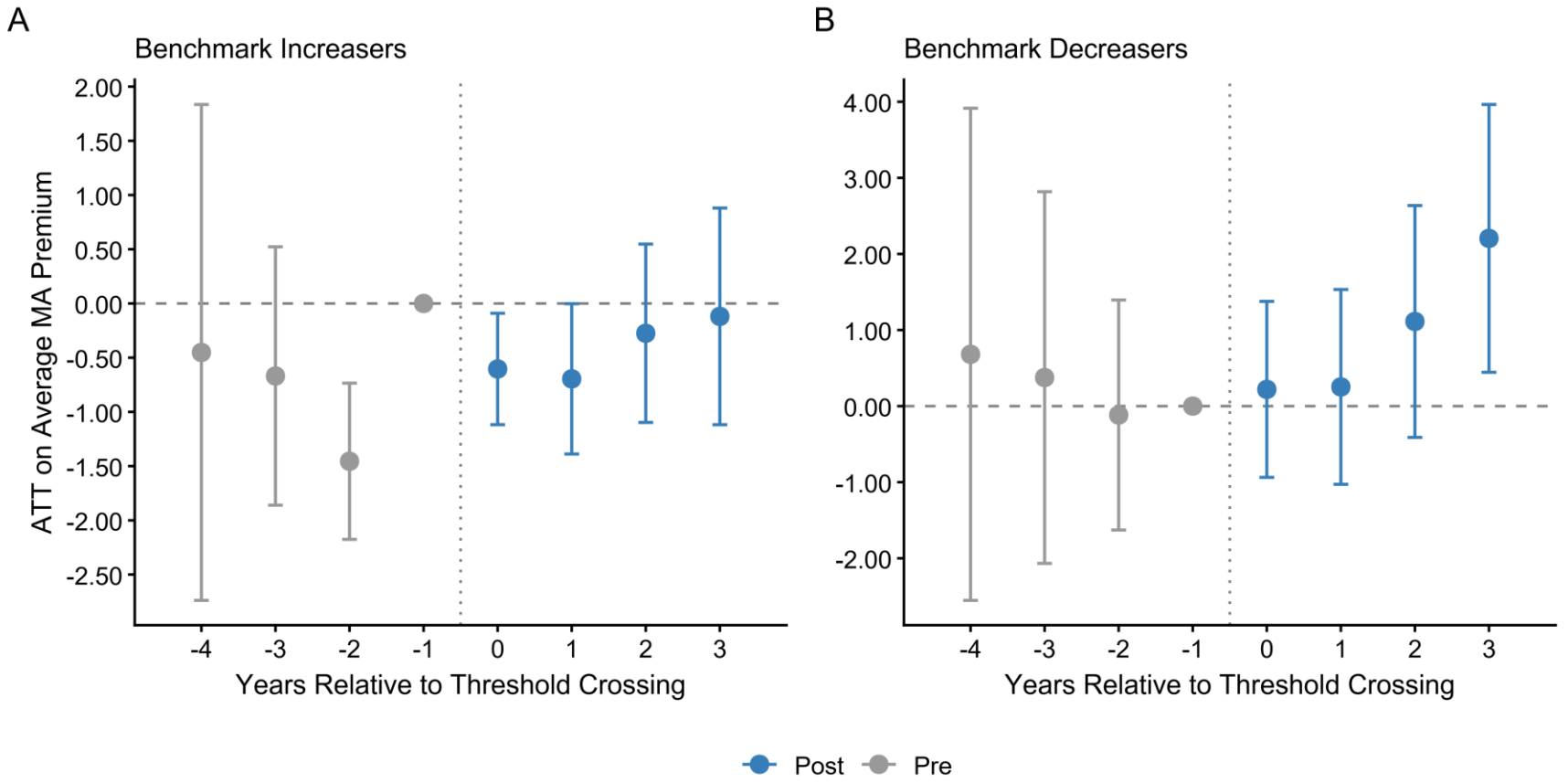
Note: Sun and Abraham (2021) staggered event-study estimates, estimated separately for counties whose benchmark increased (Panel A) and decreased (Panel B). ATT = average treatment effect on the treated. Reference year is t = −1. Bars show 95% confidence intervals based on standard errors clustered by county.

Figure A.2: Event Study Estimates of Effect of Benchmark Threshold Crossing on Number of MA Plans



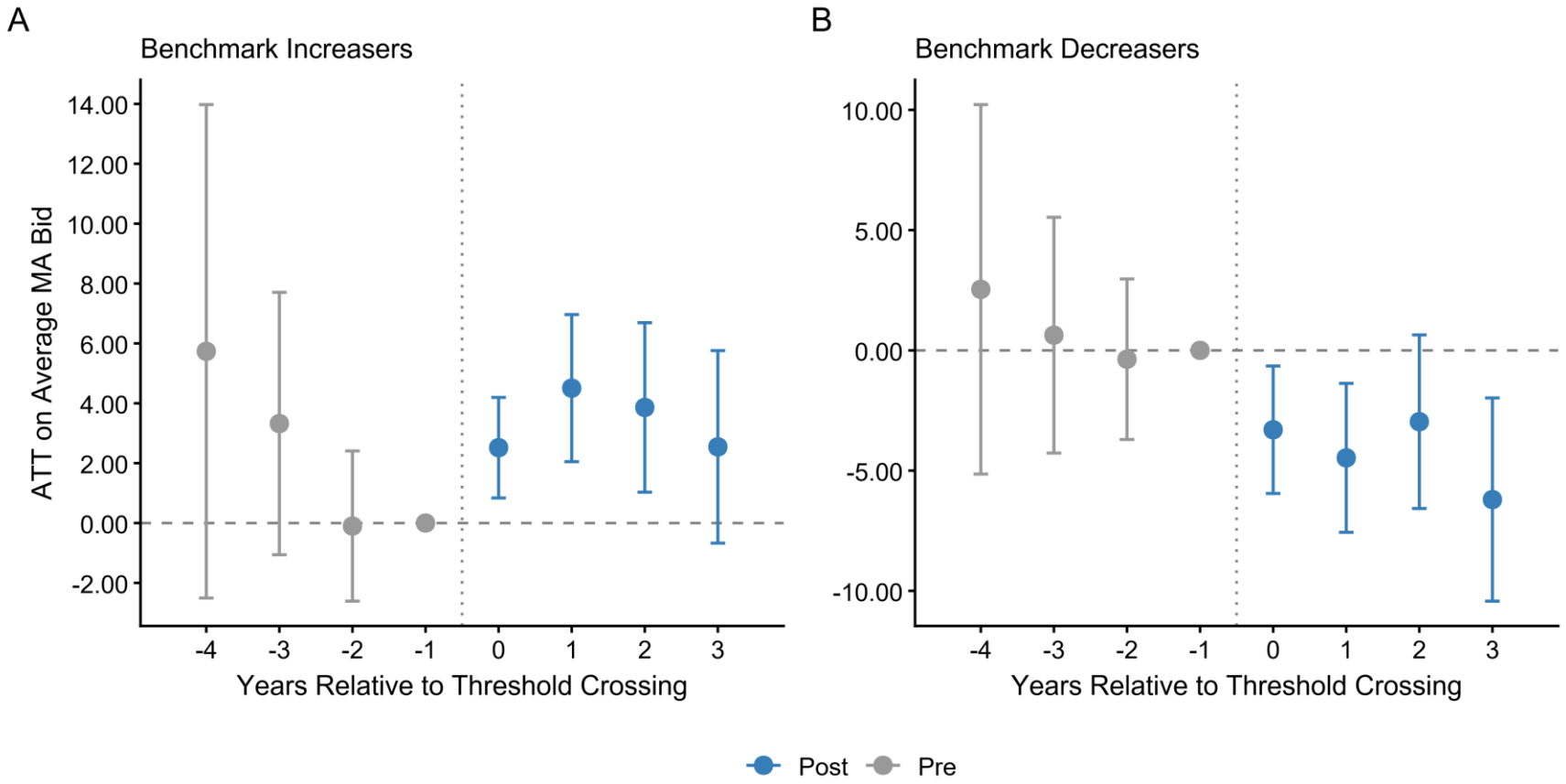
Note: Sun and Abraham (2021) staggered event-study estimates, estimated separately for counties whose benchmark increased (Panel A) and decreased (Panel B). Plan counts exclude special needs plans. ATT = average treatment effect on the treated. Reference year is t = −1. Bars show 95% confidence intervals based on standard errors clustered by county.

Figure A.3: Event Study Estimates of Effect of Benchmark Threshold Crossing on Average MA Premium



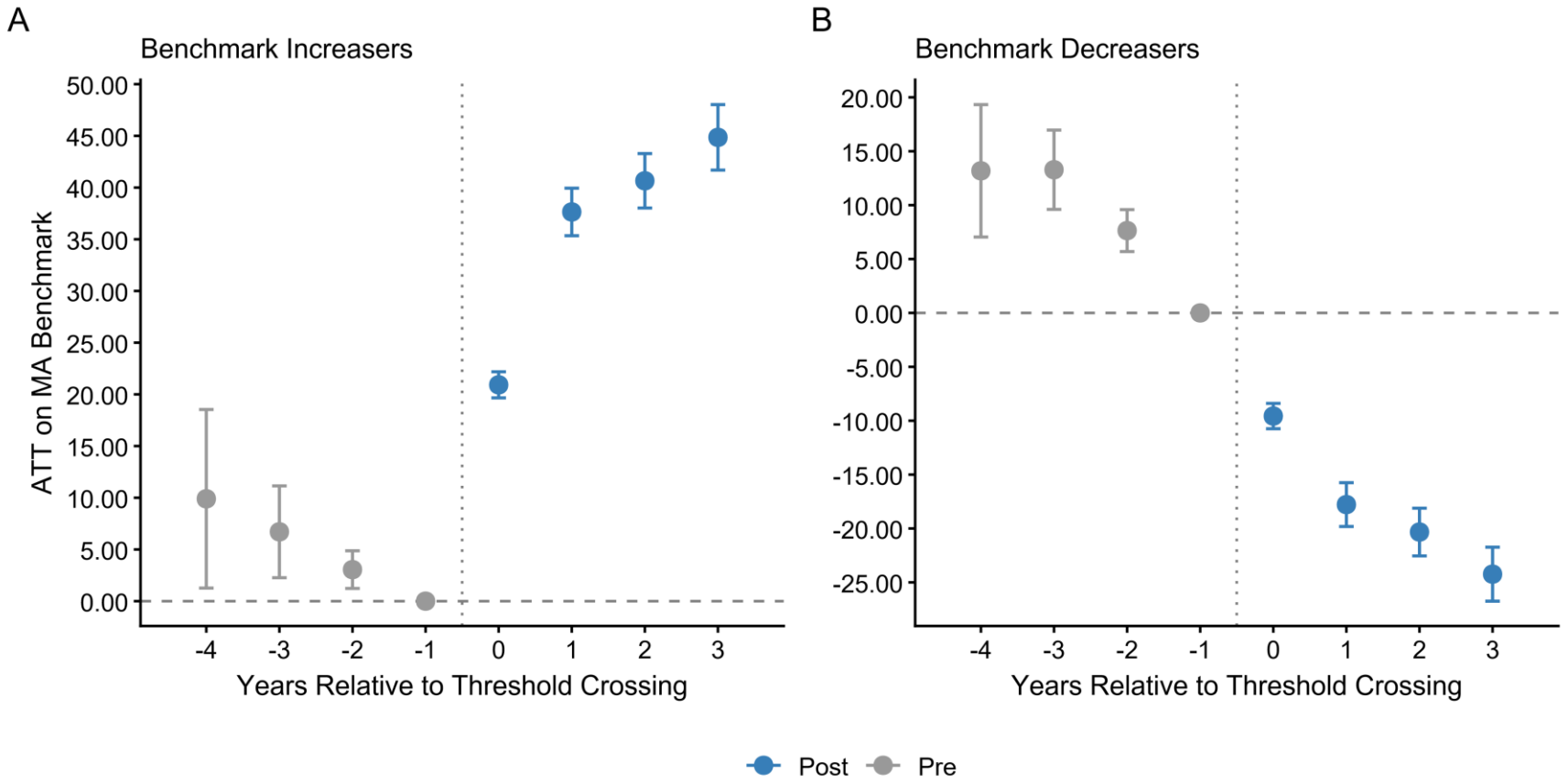
Note: Sun and Abraham (2021) staggered event-study estimates, estimated separately for counties whose benchmark increased (Panel A) and decreased (Panel B). ATT = average treatment effect on the treated. Reference year is t = −1. Bars show 95% confidence intervals based on standard errors clustered by county.

Figure A.4: Event Study Estimates of Effect of Benchmark Threshold Crossing on Average MA Bid



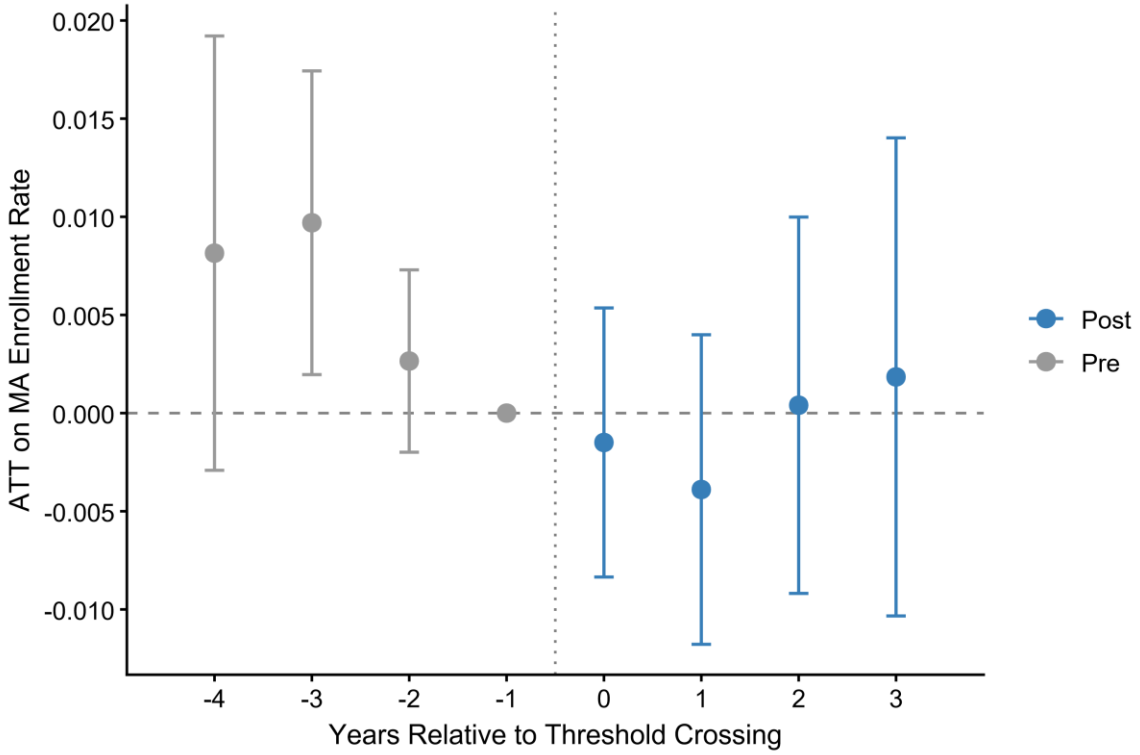
Note: Sun and Abraham (2021) staggered event-study estimates, estimated separately for counties whose benchmark increased (Panel A) and decreased (Panel B). ATT = average treatment effect on the treated. Reference year is t = −1. Bars show 95% confidence intervals based on standard errors clustered by county.

Figure A.5: Event Study Estimates of Effect of Benchmark Threshold Crossing on MA Benchmark



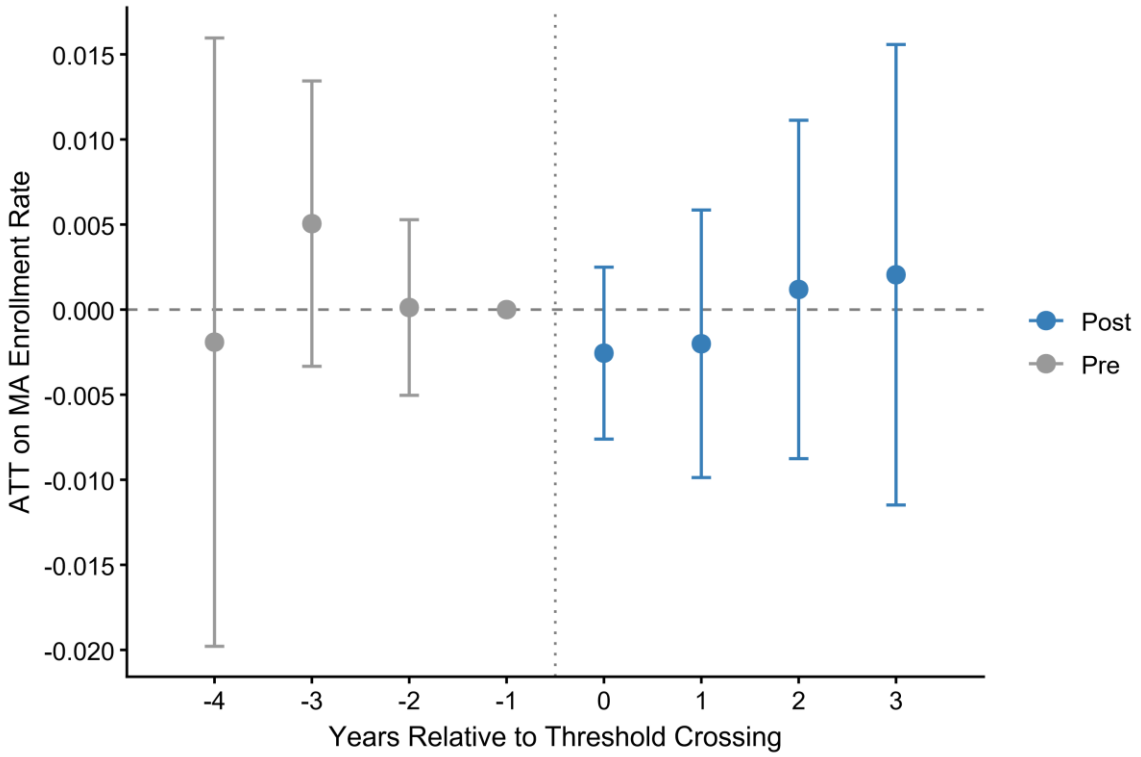
Note: Sun and Abraham (2021) staggered event-study estimates, estimated separately for counties whose benchmark increased (Panel A) and decreased (Panel B). ATT = average treatment effect on the treated. Reference year is t = −1. Bars show 95% confidence intervals based on standard errors clustered by county.

Figure A.6: Placebo Test: Effect of Quartile Increases in Capped Benchmark Counties on MA Enrollment Rate



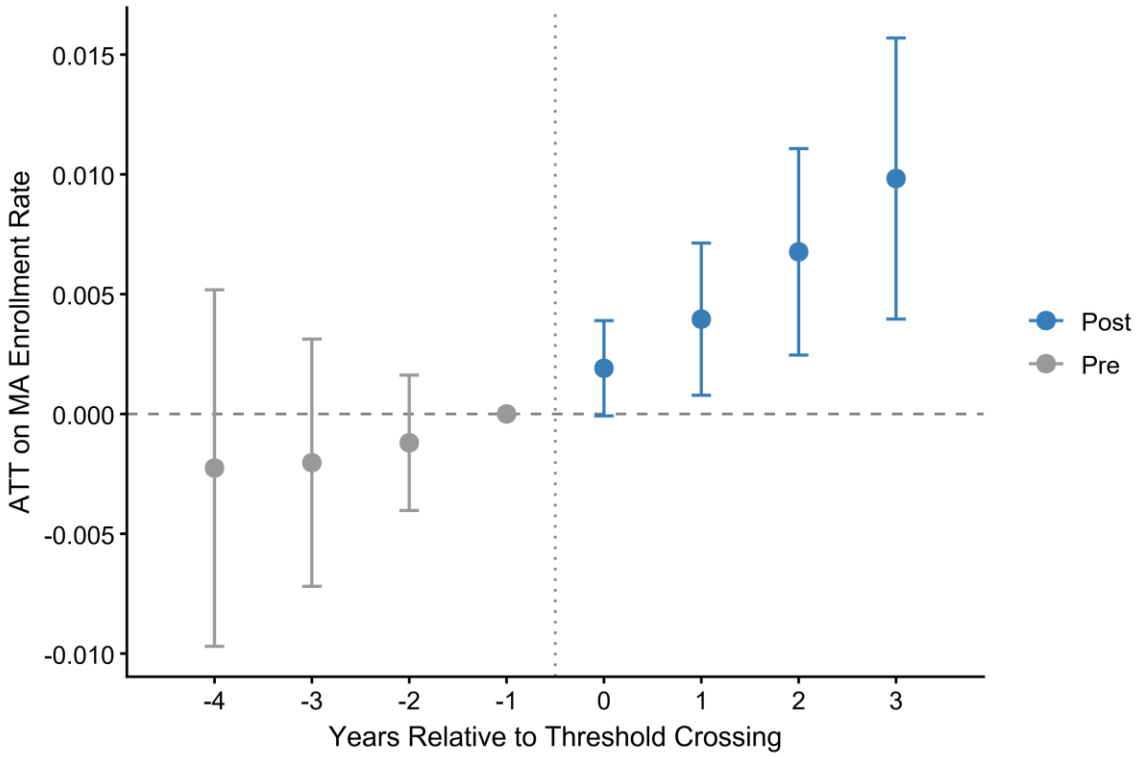
Note: Sun and Abraham (2021) staggered event-study estimates for counties whose benchmark was capped, so that moving to a higher-percentage quartile did not raise payment; comparison counties are also capped. ATT = average treatment effect on the treated. Reference year is t = −1. Bars show 95% confidence intervals based on standard errors clustered by county.

Figure A.7: Placebo Testing Effect of FFS Tercile Crossing on MA Enrollment Rate



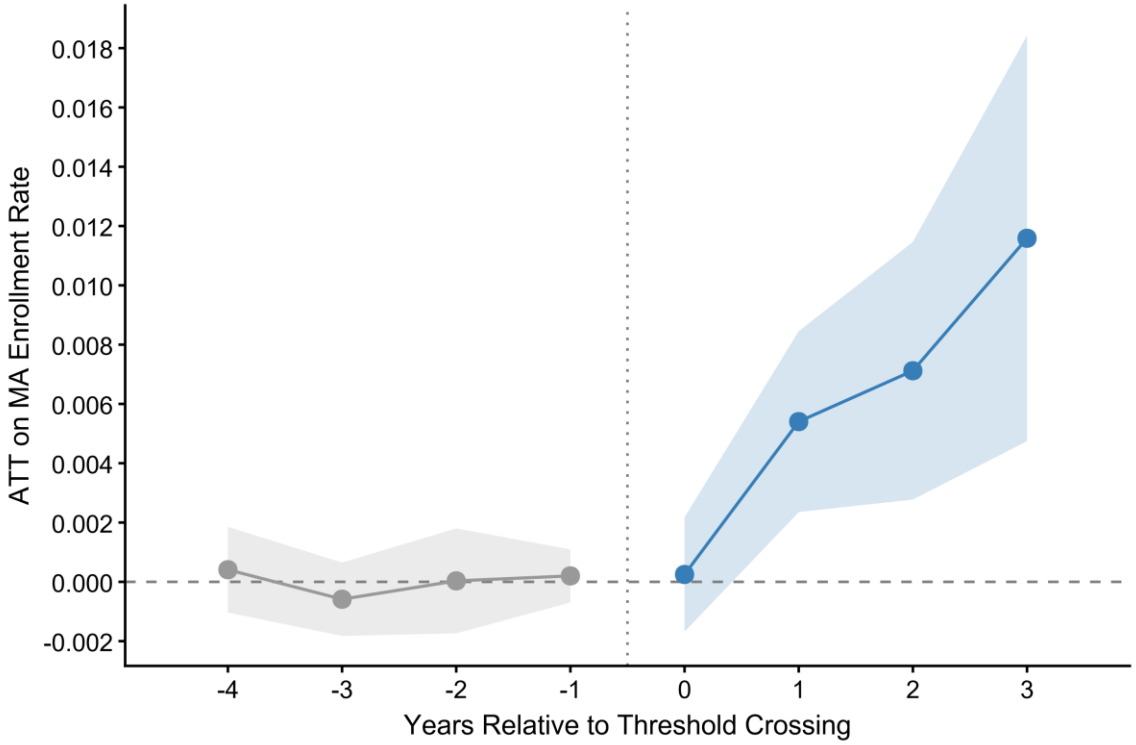
Note: Sun and Abraham (2021) staggered event-study estimates in which treatment is crossing a tercile of the county fee-for-service (FFS) spending distribution, which carries no change in payment; counties that crossed a benchmark quartile are excluded. Estimates pool crossings in both directions and are signed in the direction of the implied benchmark change. ATT = average treatment effect on the treated. Reference year is t = −1. Bars show 95% confidence intervals based on standard errors clustered by county.

Figure A.8: Event Study Estimates of Simulated Benchmark Crossing on MA Enrollment Rate



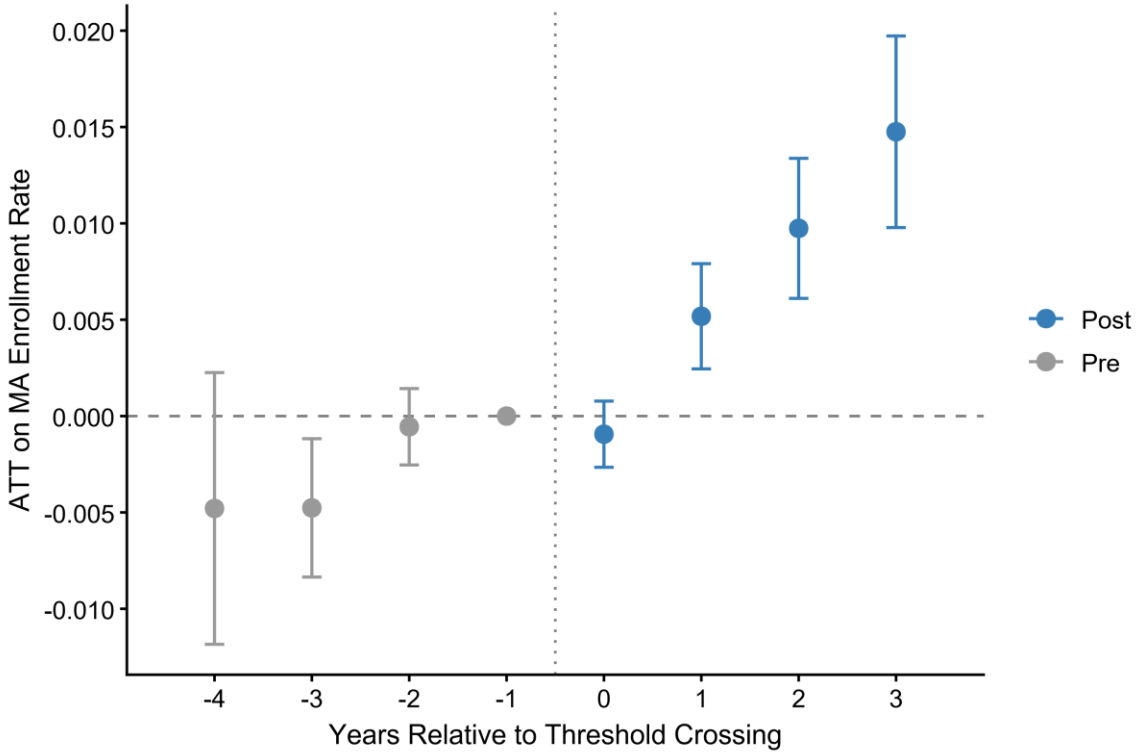
Note: Sun and Abraham (2021) staggered event-study estimates in which treatment is a simulated quartile crossing that holds each county's fee-for-service (FFS) spending at its prior-year level, so crossings arise only from shifts in national quartile cutoffs. Benchmark increases only. ATT = average treatment effect on the treated. Reference year is t = −1. Bars show 95% confidence intervals based on standard errors clustered by county.

Figure A.9: Synthetic Control Event Study: Effect of Benchmark Threshold Crossing on MA Enrollment Rate



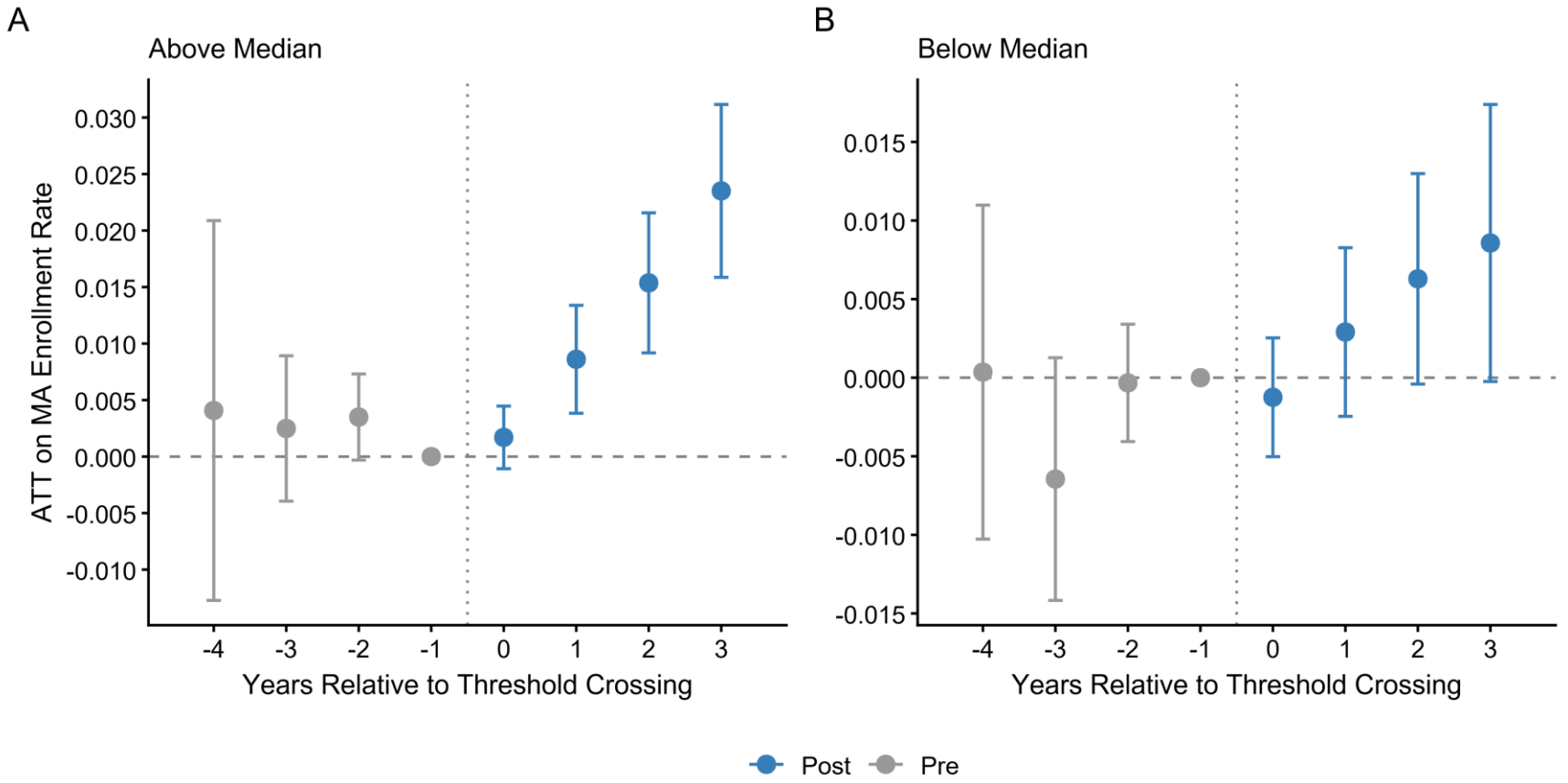
Note: Augmented synthetic control estimates for staggered adoption (Ben-Michael et al., 2021); benchmark decreases are sign-flipped and stacked with increases. ATT = average treatment effect on the treated. Shaded areas show 95% confidence intervals based on jackknife standard errors.

Figure A.10: Event Study Estimates Controlling for Projected FFS: Effect of Crossing on MA Enrollment Rate



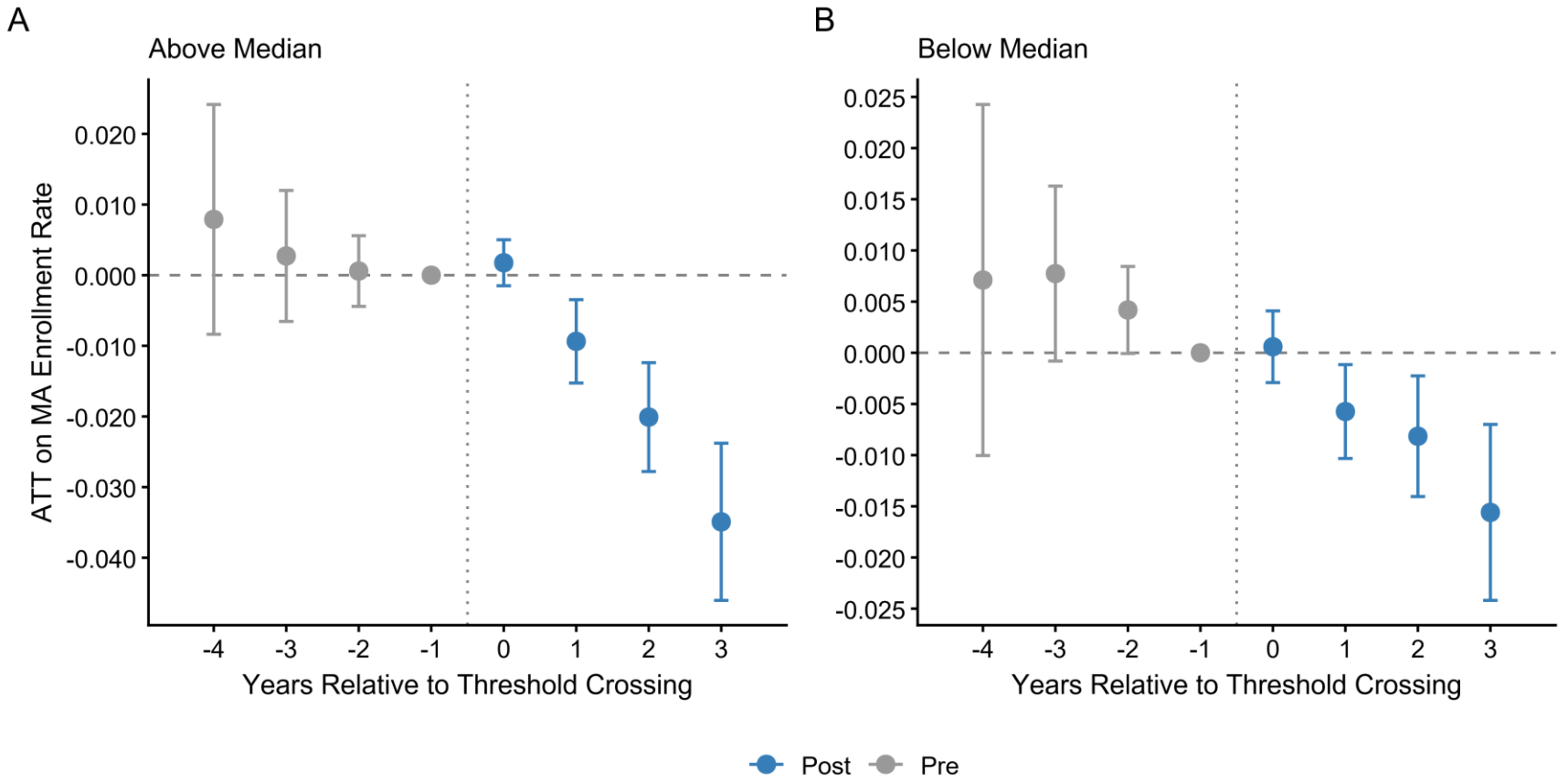
Note: Sun and Abraham (2021) staggered event-study estimates controlling for projected fee-for-service (FFS) spending, based on each county's 2007–2015 spending trend, in place of the FFS spending used to set benchmarks. Estimates pool benchmark increases and decreases and are signed in the direction of the benchmark change. ATT = average treatment effect on the treated. Reference year is t = −1. Bars show 95% confidence intervals based on standard errors clustered by county.

Figure A.11: Dose-Response Estimates of Effect of Benchmark Increases on MA Enrollment Rate



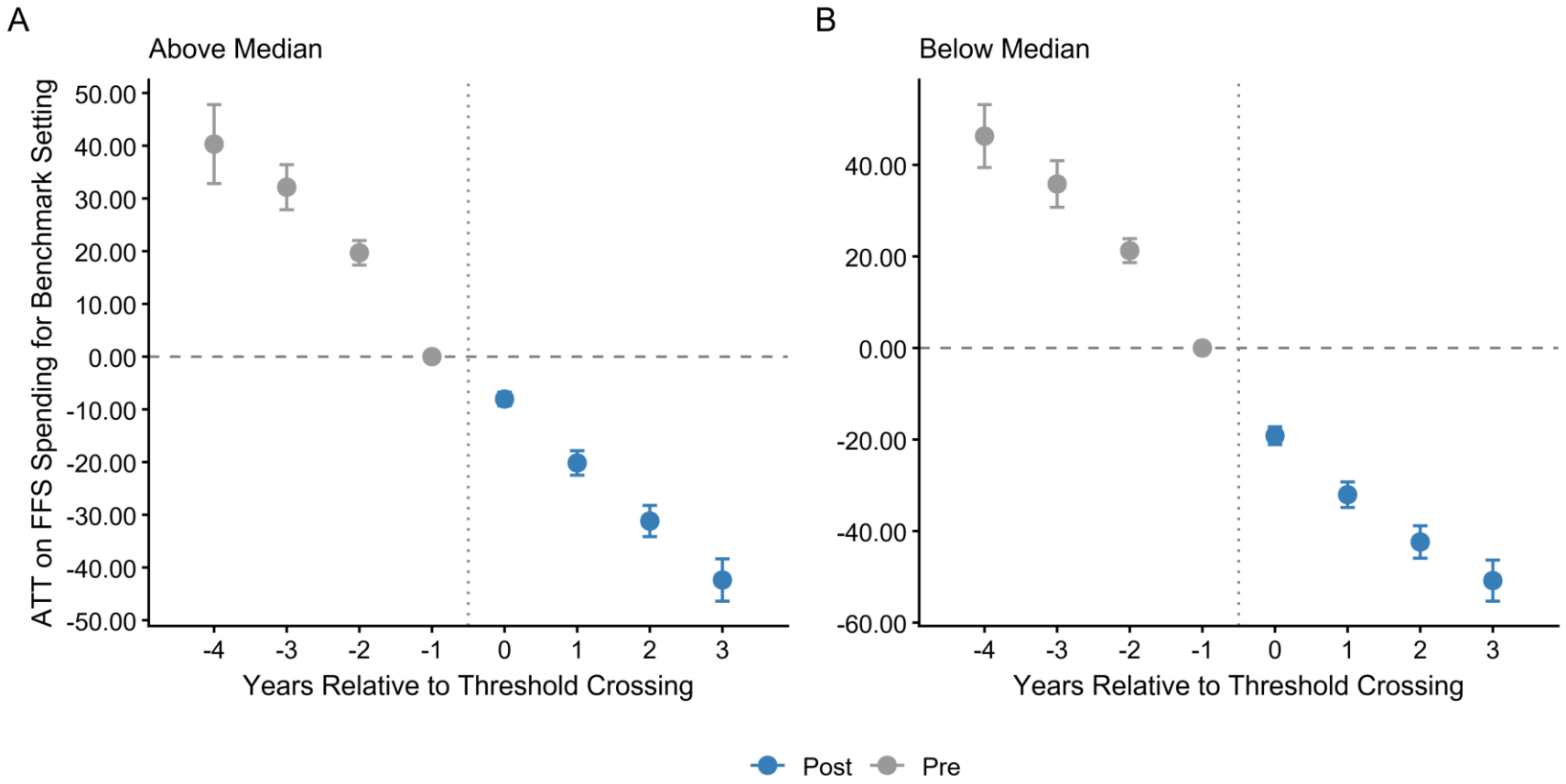
Note: Sun and Abraham (2021) staggered event-study estimates for counties whose benchmark increased, estimated separately for counties with above-median (Panel A) and below-median (Panel B) benchmark changes at crossing, relative to the median within each crossing cohort. ATT = average treatment effect on the treated. Reference year is t = −1. Bars show 95% confidence intervals based on standard errors clustered by county.

Figure A.12: Dose-Response Estimates of Effect of Benchmark Decreases on MA Enrollment Rate



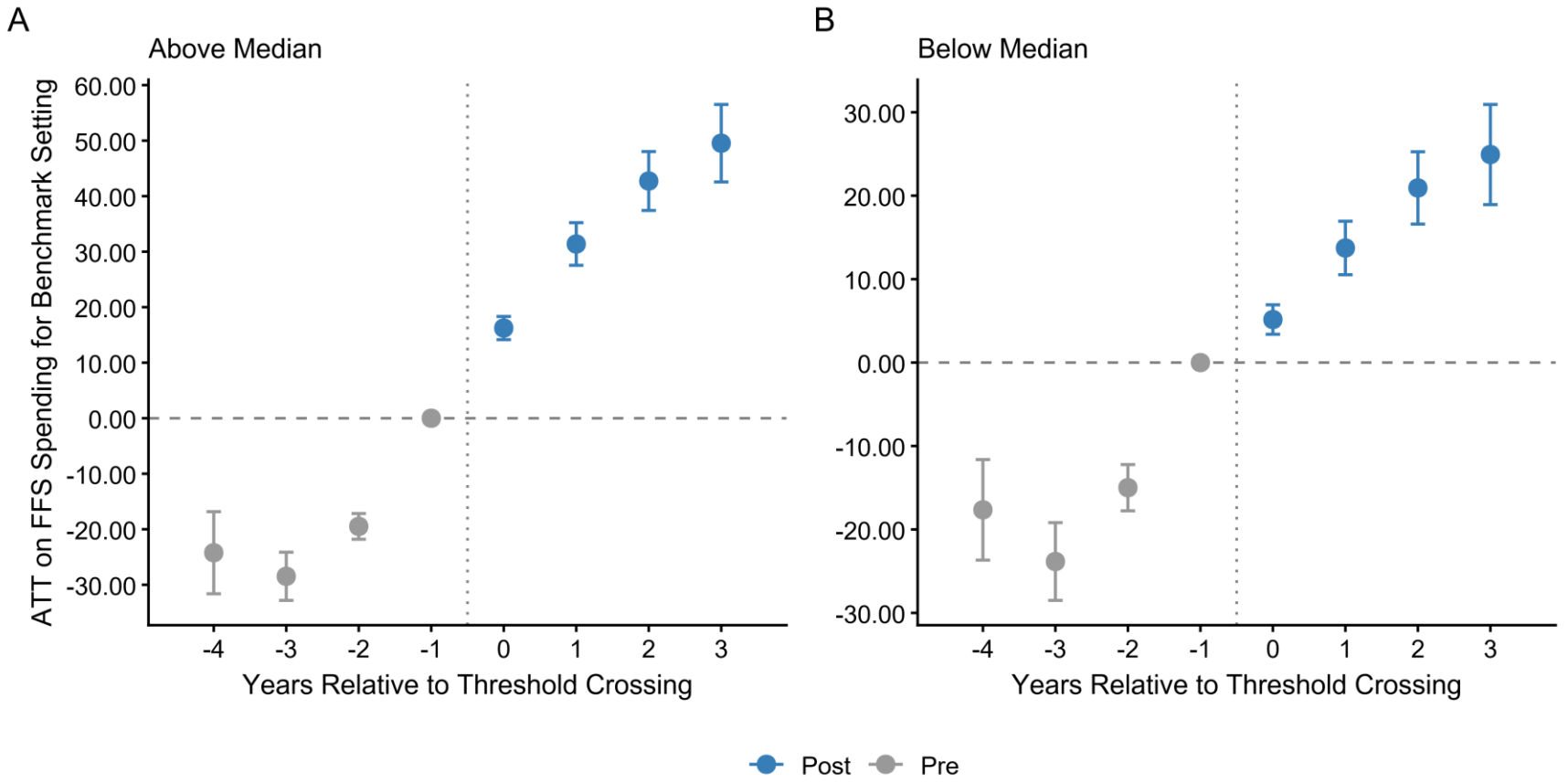
Note: Sun and Abraham (2021) staggered event-study estimates for counties whose benchmark decreased, estimated separately for counties with above-median (Panel A) and below-median (Panel B) benchmark changes at crossing, relative to the median within each crossing cohort. ATT = average treatment effect on the treated. Reference year is t = −1. Bars show 95% confidence intervals based on standard errors clustered by county.

Figure A.13: Dose-Response Estimates of Effect of Benchmark Increases on FFS Spending for Benchmark Setting



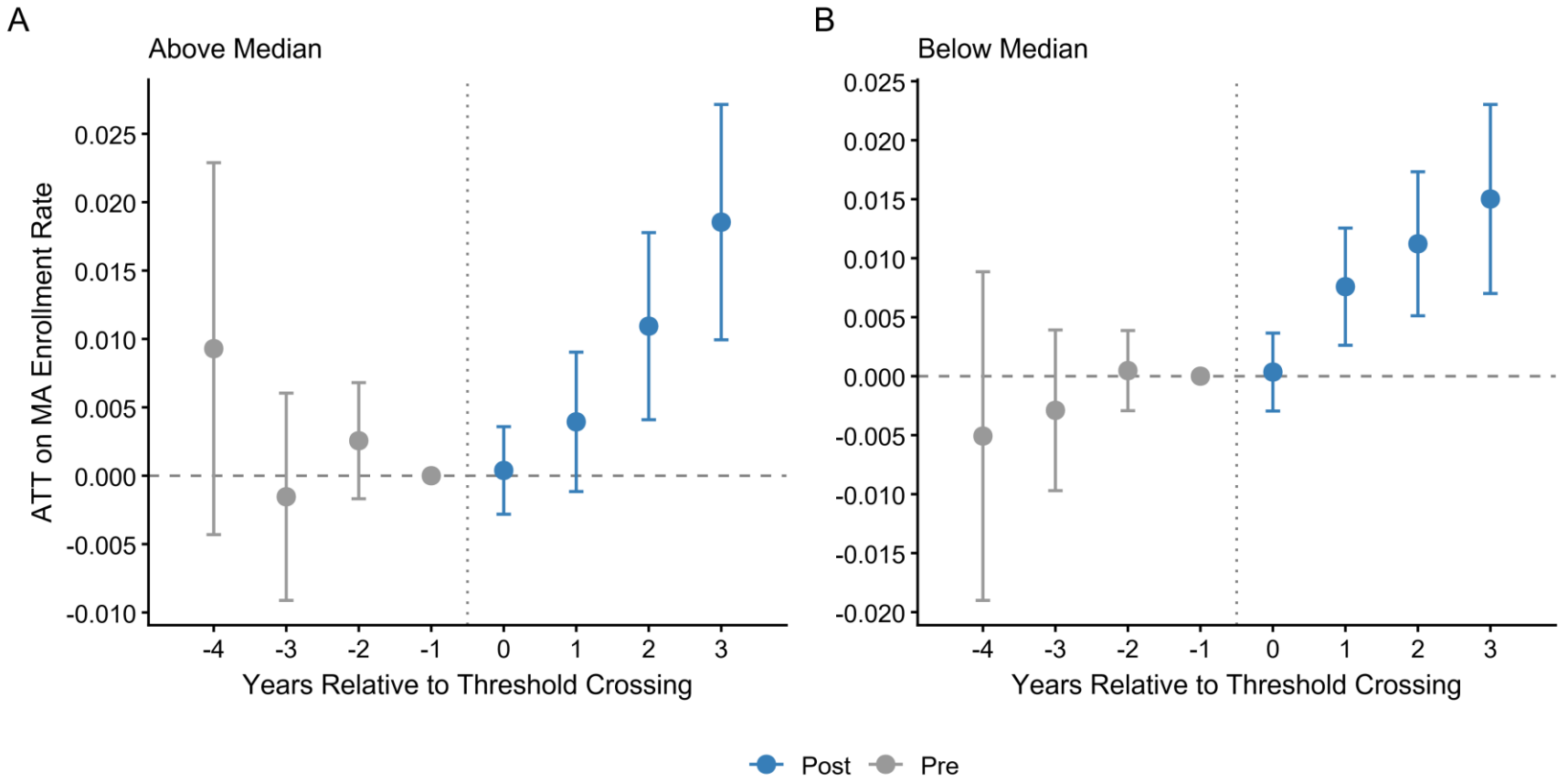
Note: Sun and Abraham (2021) staggered event-study estimates for counties whose benchmark increased, estimated separately for counties with above-median (Panel A) and below-median (Panel B) benchmark changes at crossing, relative to the median within each crossing cohort. Models do not control for fee-for-service (FFS) spending. ATT = average treatment effect on the treated. Reference year is t = −1. Bars show 95% confidence intervals based on standard errors clustered by county.

Figure A.14: Dose-Response Estimates of Effect of Benchmark Decreases on FFS Spending for Benchmark Setting



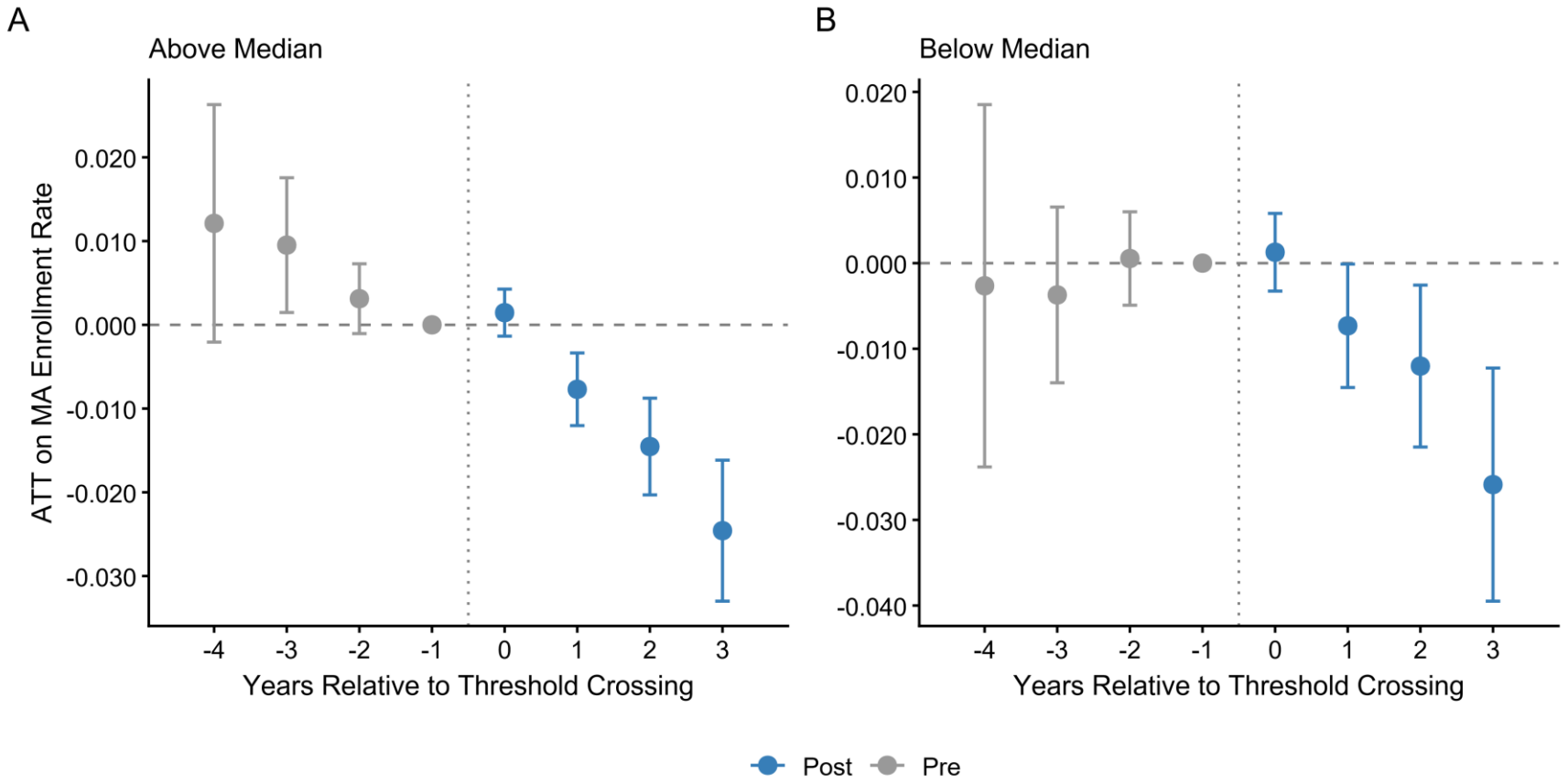
Note: Sun and Abraham (2021) staggered event-study estimates for counties whose benchmark decreased, estimated separately for counties with above-median (Panel A) and below-median (Panel B) benchmark changes at crossing, relative to the median within each crossing cohort. Models do not control for fee-for-service (FFS) spending. ATT = average treatment effect on the treated. Reference year is t = −1. Bars show 95% confidence intervals based on standard errors clustered by county.

Figure A.15: Estimates of Effect of Benchmark Increases by County HHI on MA Enrollment Rate



Note: Sun and Abraham (2021) staggered event-study estimates for counties whose benchmark increased, estimated separately for counties above (Panel A) and below (Panel B) the 2015 median county Herfindahl-Hirschman Index (HHI) across MA parent organizations. ATT = average treatment effect on the treated. Reference year is t = −1. Bars show 95% confidence intervals based on standard errors clustered by county.

Figure A.16: Estimates of Effect of Benchmark Decreases by County HHI on MA Enrollment Rate



Note: Sun and Abraham (2021) staggered event-study estimates for counties whose benchmark decreased, estimated separately for counties above (Panel A) and below (Panel B) the 2015 median county Herfindahl-Hirschman Index (HHI) across MA parent organizations. ATT = average treatment effect on the treated. Reference year is t = −1. Bars show 95% confidence intervals based on standard errors clustered by county.

Table A.1: Hazard Model for County-Level Quartile Crossing Resulting from Lagged MA Enrollment Penetration.

Direction	Coefficient on lagged MA penetration	SE	p-value	Crossing events	County-years
Increasers	0.121	0.136	0.371	270	4,889
Decreasers	-0.197	0.139	0.158	231	4,723

Note: Discrete-time hazard model (linear probability) of a county's first quartile crossing in year t, estimated on county-years at risk of crossing, 2016–19, separately for crossings that raised (Increasers) and lowered (Decreasers) the benchmark. Lagged MA penetration is the county MA enrollment share in year $t - 1$. Models control for lagged fee-for-service (FFS) spending and its square, county population, median household income, physicians and hospital beds per 100,000 residents, and county and year fixed effects. Standard errors are clustered by county.