Spillover Effects between Medicare Programs

Ira Abraham, Alberto Cappello, and Yunus Semih Coskun

Boston College

April 7, 2026

Abstract

Medicare relies on multiple programs that provide coverage to different groups of beneficiaries. The Medicare Shared Savings Program (MSSP) introduces incentives to reduce unnecessary spending for Traditional Medicare (TM) beneficiaries. Changes induced by the MSSP can spill over to beneficiaries enrolled in Medicare Advantage (MA) and affect MA plans' outcomes. The MSSP can affect MA through two channels. First, a *benchmark channel*: because MA plans payments are tied to county-level TM per-capita spending, MSSP-induced changes in county-level TM spending can mechanically affect MA plan payments. Second, a *provider channel*: because providers treat both TM and MA beneficiaries, MSSP-induced changes in care delivery may directly affect MA enrollees and, in turn, plan expected costs and benefit design.

We measure the plan-level exposure to MSSP activity as the enrollment-weighted average of the county MSSP penetration in the counties that are part of the plan service area. We identify the *provider channel* by controlling for the plan-level payments to net out the mechanical effect that operates through the *benchmark channel.* We leverage quasi-experimental variation created by the 2019 *Pathways to Success* reform of the MSSP, which shifted the strength of MSSP incentives. Combining this policy variation with pre-reform plan shares and MSSP providers penetration, we construct a Bartik-style instrument for plan-level exposure to MSSP activity.

We find evidence that greater MSSP exposure lowers MA plans' expected cost of coverage which, in turn, reduces MA payment to plans. In particular, a 1% increase in MSSP exposure lowers plans' expected costs by 0.34% and lower plans' payments by about 0.20%. Moreover, part of this lower costs are passed through to enrollees through higher rebates. However, we interpret the estimation on rebates more cautiously as the instrument is correlated with pre-2019 trend in rebates. We also find that plan ownership plays an important role: spillovers are not significant among plans sponsored by health systems that participate in the MSSP.

Keywords: Health Economics, Medicare Advantage, Medicare Shared Savings Program, Spillover, Bartik Instrument.

1 Introduction

Historically, Medicare has relied on Fee-for-Service (FFS) payment for the majority of its beneficiaries. Over the past decade, Medicare has increasingly relied on alternative payment models to contain the growing healthcare spending of its beneficiaries. Two of the most important are Medicare Advantage (MA) and the Medicare Shared Savings Program (MSSP). As of 2025, MA covers about half of Medicare beneficiaries, while MSSP covers about one-sixth of the remaining Traditional Medicare (TM) population. Medicare Advantage aims to contain utilization by paying a capitated amount to private plans to cover TM benefits, whereas under the MSSP providers are paid a bonus for keeping average per-capita spending below a benchmark. Although these programs are often studied separately, incentives created by the MSSP can spill over into the MA program.

This paper studies two conceptually distinct channels through which MSSP can affect MA. The first is the *benchmark channel*. Because MA benchmarks are tied to county FFS spending among TM beneficiaries, MSSP-induced changes in TM spending can mechanically alter MA plan payments. The second is the *provider channel*. Since providers treat both TM and MA beneficiaries, practice-wide changes made in response to MSSP incentives —for example, greater care coordination, tighter post-acute management, or substitution away from higher-cost settings— can affect MA beneficiaries and, in turn, MA plans' expected medical costs, plan payments, and benefit design.

The main outcomes of interest are plan-level projected costs, bids, and rebates. Projected costs measure a plan's expected medical cost of covering a standard Medicare beneficiary. Bids represent the plan's submitted revenue requirement for providing standard Medicare benefits and therefore reflect projected medical costs together with administrative costs and a profit margin. In Medicare Advantage, plans are paid the lower of the bid and the county benchmark. When a plan bids below the benchmark, it receives a rebate equal to a fraction of the benchmark-bid difference, which must be returned to enrollees through supplemental benefits or lower cost sharing. We construct these variables by combining CMS bid pricing tools, MA enrollment and plan-characteristics files, county ratebook files, and MSSP public-use files.

The first challenge is measuring the plan-level exposure to MSSP incentives. MSSP providers participate in the MSSP through joint ventures called Accountable Care Organizations (ACOs). In each county, we observe ACO penetration, which is given by the share of TM beneficiaries assigned to ACOs. For each MA plan p in year t, we construct a measure of exposure to MSSP activity, π_{pt}^{ACO}, defined as the enrollment-weighted average of county ACO penetration in the counties that are part of the plan service area — i.e., the set of counties in which the plan operates.

This plan exposure to MSSP, π_{pt}^{ACO}, has several sources of endogeneity. First, MA plans with larger shares in counties with higher ACO penetration might also differ in unobservable quality or the health status of their enrollees. Second, counties with higher ACO penetration might also differ in unobservable healthcare supply characteristics that also affect plan outcomes. Third, plans that are on a different trend in terms of cost per enrollee may sort into markets with higher ACO penetration.

To address these endogeneity threats, we construct a Bartik-style instrument that combines predetermined plan enrollment weights, predetermined county ACO penetration, and quasi-experimental variation in MSSP incentive strength generated by the 2019 *Pathways to Success* MSSP reform. The reform created variation in the strength of the MSSP incentive both across ACO cohorts and within ACO cohorts over time. This generates plausibly exogenous time-series variation in plan exposure to MSSP incentives that interacts with predetermined plan exposure to ACO activity.

π_{pt}^{ACO} can affect MA outcomes both through provider behavior and through benchmark-mediated payment changes. Since we do not have an instrument that creates exogenous variation in the MA benchmark, our empirical strategy does not allow us to separately identify the effect of the benchmark channel. We isolate the provider channel effect by using the plan-level MA benchmark as a control variable and instrumenting π_{pt}^{ACO} with our Bartik-style variable.

Because plan-level MSSP exposure is constructed as the enrollment-weighted average of county ACO penetration across the plan's service area, it is naturally exposed to region-specific time-varying shocks. For example, local changes in provider supply, market structure, coding intensity, or plan demand could simultaneously affect county ACO penetration and MA plan outcomes. To address this concern transparently, our empirical results compare four specifications for each outcome: OLS and IV models with additive core-based statistical area (CBSA) and year fixed effects, and OLS and IV models with CBSA $\times$ year fixed effects. CBSAs group nearby counties into local metropolitan or micropolitan markets and are a natural regional unit in our setting because MA plans often operate across multiple neighboring counties that share provider and demand conditions. The tighter CBSA $\times$ year specification is especially important because it absorbs shocks that are common to all plans operating in the same local market and year.

Another key endogeneity concern is that plans with greater predicted exposure to the reform may already have been on different trends before 2019 in terms of expected cost per beneficiary or other outcome variables. This concern is especially relevant because the Bartik instrument combines policy-driven time variation with predetermined cross-sectional exposure shares. To assess this, we implement an event-study analysis that compares pre- and post-2019 outcome paths for plans with different instrument-induced exposure to the 2019 MSSP reform. The results support the design for projected costs and bids, but are less reassuring for rebates, where the pre-2019 coefficients suggest residual differential trends. For that reason, we place the greatest weight on the projected-cost and bid results and interpret the rebate evidence more cautiously.

Our empirical analysis uses a plan-level panel spanning 2016 to 2023. Within each fixed-effects structure, we estimate two related objects. The first is a benchmark-controlled specification, which we interpret as the *provider channel*: it asks whether MSSP exposure affects MA outcomes after holding fixed the plan-level benchmark. The second omits the plan-level benchmark and is interpreted as a *total effect* object, combining provider-side spillovers with any effects operating through benchmark-linked payments. The comparison between these two panels is informative about whether benchmark-related payment variation matters empirically, but it does not separately identify the benchmark channel in a causal sense.

IV estimates of the impact of MSSP exposure on projected costs are negative in the additive fixed-effects specification, but lose significance under CBSA $\times$ year fixed effects. Because projected costs are only observed through 2021, however, the post-reform period for this outcome is short and heavily affected by the COVID years. For that reason, we also report a targeted specification excluding 2020, the first COVID year. In that specification, under CBSA $\times$ year fixed effects, a 1% increase in MSSP exposure lowers projected costs by about 0.34% in the benchmark-controlled Provider Channel specification and by about 0.35% in the benchmark-uncontrolled Total Effect specification. We interpret this as evidence that greater MSSP exposure lowers plans' expected cost of coverage.

The IV specifications with CBSA $\times$ year fixed effects show that a 1% increase in plan-level MSSP exposure lowers bids by about 0.20%. By contrast, the corresponding OLS coefficients become small and statistically insignificant once CBSA $\times$ year fixed effects are included, suggesting that non-instrumented exposure is contaminated by market-year confounding. The rebate estimates are positive and economically large, which is consistent with pass-through to beneficiaries, but our validation exercise reveals that the Bartik instrument is correlated with pre-2019 trends in rebates. Therefore, we interpret the rebate results as suggestive rather than as causal evidence. Estimated effects on out-of-pocket maxima, premiums, deductibles, and copays are generally imprecise. Thus, the most credible evidence indicates that greater MSSP exposure lowers expected costs and lowers the plan's broad revenue requirement, while the evidence on rebate pass-through is more tentative.

We also examine whether these spillovers depend on organizational structure. Some plans are owned by health systems that operate on both sides of the MSSP–MA linkage: they participate in MSSP through affiliated ACOs while also sponsoring MA plans. For these dual-role systems, reductions in fee-for-service spending increase shared-savings rewards on the MSSP side but can also reduce benchmark-based MA payments to the system's own plans. This creates a trade-off that may weaken incentives to translate MSSP exposure into lower MA costs. We test this heterogeneity by running the main specification with CBSA $\times$ year fixed effects separately for the samples of non-dual-role and dual-role plans.

The heterogeneity results are consistent with this mechanism. Excluding 2020, a 1% increase in MSSP exposure lowers projected costs for non-dual-role plans by about 0.33% in both the Provider Channel and Total Effect specifications, whereas the corresponding dual-role estimates are small in magnitude and statistically indistinguishable from zero. In the bid regressions, a 1% increase in MSSP exposure lowers bids for non-dual-role plans by about 0.11–0.14%, while the corresponding dual-role estimates are much smaller and statistically indistinguishable from zero. The rebate estimates are directionally similar, with larger positive coefficients among non-dual-role plans, but we interpret those heterogeneity patterns cautiously because the validation exercise is less supportive of a causal interpretation. Overall, the evidence suggests that spillovers are attenuated for plans sponsored by health systems that internalize both sides of the MSSP–MA linkage.

This paper makes three contributions. First, it provides new evidence that MSSP and MA are linked not only through payment rules, but also through provider behavior that spills over across Medicare programs. Second, it develops a novel Bartik-style instrument based on the 2019 Pathways reform and uses it to estimate both the overall IV effect of MSSP exposure and the provider channel effect. Third, it shows that organizational structure shapes the extent of these spillovers, with attenuated responses among dual-role health systems.

More broadly, the paper shows that Medicare payment reforms cannot be evaluated in isolation. When providers operate across multiple Medicare programs, incentives targeted at one segment can alter outcomes in another. Understanding the incidence of payment reform therefore requires accounting not only for direct program effects, but also for the institutional and organizational linkages through which those effects propagate.

The remainder of the paper proceeds as follows. Section 2 describes related literature. Section 3 presents the institutional setting and discusses the construction of the measure of plan exposure to MSSP. Section 4 presents the data, describes the key variables of interest, and shows the descriptive statistics. Section 5 discusses the sources of endogeneity and the identification strategy. Section 6 presents the results. Section 7 concludes.

2 Related Literature

This paper connects three strands of the literature: evaluations of the MSSP and its fiscal consequences, spillovers between Medicare programs, and the econometrics of shift-share instrumental variables.

MSSP evaluations. A large body of work evaluates whether MSSP ACOs reduce Medicare spending for their attributed beneficiaries. Early studies found modest gross savings concentrated among physician-led ACOs, driven primarily by reductions in spending on post-acute care, outpatient care in hospital-owned facilities, and acute inpatient admissions (McWilliams et al., 2016, 2018; Colla et al., 2016). However, net savings—after subtracting shared-savings payments—are smaller and sometimes negative (McWilliams et al., 2020). A central concern is that MSSP evaluations may overstate savings because of nonrandom participation and attrition: Markovitz et al. (2019) shows that exit of high-cost clinicians can account for a substantial portion of apparent spending reductions, while Ouayogódé et al. (2021) finds that favorable provider and beneficiary selection attenuates estimated savings. Our paper does not re-evaluate MSSP savings for attributed beneficiaries; instead, we take the ACO activity as given and ask how it spills over to a different segment of Medicare: the

MA market.

Spillovers between Medicare programs. A well-established literature documents spillovers from Medicare Advantage to Traditional Medicare. Baicker et al. (2013) show that higher MA penetration reduces hospital utilization for FFS beneficiaries, and Callison (2016) finds that MA penetration lowers treatment intensity across all local Medicare patients. Chernew et al. (2008) document similar effects on total Medicare expenditures, and Feyman et al. (2021) confirm that these spillovers persist in the post-ACA era. Using a randomized payment reform, Einav et al. (2020) provides experimental evidence that payment-model spillovers extend equally to non-targeted patients.

In the reverse direction—from TM programs to MA—Ryan and Markovitz (2023)[1] study the mechanical fiscal spillover from MSSP to MA benchmarks. Because MA benchmarks are tied to county-level FFS spending and related benchmark-setting inputs, MSSP can mechanically affect federal payments to MA plans through the benchmark formula. Within organizations, Navathe et al. (2018) finds limited spillovers of ACO participation on post-acute care for non-attributed beneficiaries at participating hospitals, and Post et al. (2019) documents modest spillovers from MSSP physician participation to commercial spending.

Most directly related to our paper, Hou et al. (2025) study MSSP spillovers to *non-attributed* FFS beneficiaries using a mover design inspired by Finkelstein et al. (2016). They find that non-attributed beneficiaries who move to areas with greater ACO penetration experience reduced outpatient facility spending and increased physician-services spending, consistent with a shift away from higher-cost facility settings. Their evidence supports the premise that ACO care-coordination practices spill over beyond attributed populations—the same "provider channel" that we study. Our contribution differs in two respects. First, we study spillovers to *MA plans* rather than to non-attributed FFS beneficiaries, examining how provider spillovers translate into changes in plan bids, benefit design, and market participation. Second, we use a shift-share IV design based on the 2019 Pathways to Success reform rather than a mover framework, exploiting quasi-experimental variation in MSSP incentive strength to identify provider-channel effects while conditioning on MA benchmarks.

[1] Ryan and Markovitz (2023) was retracted and replaced in April 2024 after the authors corrected how MSSP incentive payments enter the USPCC used to determine MA benchmarks. The corrected version continues to show that MSSP can mechanically affect MA benchmarks through the benchmark-setting formula, but it changes the sign and magnitude of that fiscal effect relative to the original publication. Our paper is complementary: rather than re-estimating the benchmark-mediated fiscal effect of MSSP on MA payments, we study whether MSSP also spills over to MA through provider behavior, and we assess that provider-side response while conditioning on plan-level benchmarks.

Shift-share instrumental variables. Our identification strategy relies on a Bartik-style (shift-share) instrument. Two complementary frameworks formalize the conditions under which such instruments are valid. Goldsmith-Pinkham et al. (2020) show that the standard Bartik instrument is numerically equivalent to using the exposure shares as individual instruments in a GMM framework. Under their "exogenous exposure" interpretation, validity requires each share to satisfy a conditional independence (parallel-trends) condition, and they propose Rotemberg weight decompositions and pre-trend diagnostics to assess which shares drive identification. Borusyak et al. (2022) develops an alternative "exogenous shifts" framework in which identification follows from the quasi-random assignment of the shifts, while the shares are allowed to be endogenous; the recent practical guide by Borusyak et al. (2025) synthesizes both approaches and provides implementation checklists. Our instrument combines predetermined plan-county enrollment shares with quasi-experimental variation from the Pathways non-rebasing year schedule, following the logic of Goldsmith-Pinkham et al. (2020): we treat the baseline shares as the source of identifying cross-sectional variation and argue for their exogeneity conditional on plan fixed effects, while the Pathways schedule provides time-varying "shocks" whose timing is pinned down by federal policy rules. Section 5 discusses the identifying assumptions and validation exercises in detail.

3 Institutional Background

3.1 The Medicare Shared Savings Program

In Traditional Medicare (TM), the government pays providers an administered fee f_s for each service s. Let Q_{ist} denote the utilization of service s. The total TM spending for beneficiary i in year t can be written as $Y_{it} = \sum_s f_s \cdot Q_{ist}$. The Medicare Shared Savings Program (MSSP), launched in 2012, aims at reducing unnecessary volume of services by changing organizational practices and through improved care coordination (e.g., chronic care management, high-risk registries, transitional care programs, post-acute management, and reductions in duplicate procedures). Groups of providers that meet specified criteria are allowed to form Accountable Care Organizations (ACOs) and share in savings if they reduce the per-capita fee-for-service (FFS) expenditures for their aligned beneficiaries while maintaining or improving quality.

Financial incentives and the ratchet effect. Each ACO faces an expenditure benchmark that is used to establish whether the ACO generated savings in a given year. If realized spending falls sufficiently below the benchmark, the ACO is rewarded by Medicare a fraction of the difference between the benchmark and its per-capita spending. This creates a direct financial return to lowering spending for aligned beneficiaries.

The spending benchmark is a weighted average between the ACO's recent average past spending and the FFS regional average spending. Before 2019, ACOs' participation in the MSSP was organized into agreement periods of three years, and the benchmark is updated every three years to reflect changes in each component. A central challenge of the MSSP design is the *ratchet effect*: successful spending reductions mechanically tighten future benchmarks. Intuitively, if an ACO reduces spending today, rebasing shifts down the benchmark used to measure savings tomorrow, reducing the marginal payoff to sustained savings and weakening dynamic incentives to invest in cost-reducing practice patterns.

Pathways to Success and non-rebasing years. The 2019 *Pathways to Success* reforms represented a major redesign of MSSP intended, in part, to mitigate this ratchet effect. Pathways reorganized participation into five-year agreement periods and introduced *non-rebasing years* within each agreement period. In non-rebasing years, realized spending affects shared savings/losses in the current year but does not tighten the benchmark used in subsequent years. Therefore, ACOs have a stronger marginal incentive to generate savings in those years.

Crucially for this paper, the timing of these non-rebasing years is pinned down by an ACO's entry cohort and the Pathways policy schedule, and not by the ACO's choice or local market conditions. Section 5.3 formalizes this schedule and describes how we use it to build plan-level quasi-experimental exposure measures to MSSP incentives.

3.2 Medicare Advantage

Medicare Advantage (MA) is the capitated alternative to Traditional Medicare. Private plans are paid by Medicare a capitated amount per-beneficiary per month to provide coverage for TM Part A and Part B benefits. Plans are differentiated in terms of provider network and coverage, and negotiate with providers the payment rates for each service. In the rest of this section, we will describe how plan payments are set and how this might create linkages between MSSP and MA.

Plan projected costs. In the MA bid process, plans first form a forward-looking projection of the medical cost of covering the standard Part A and Part B benefit package for a beneficiary of average risk. We refer to this object as *projected cost.* It is an ex ante pricing object rather than realized ex post claims spending, and it summarizes the expected cost of standard Medicare-covered services before administrative expenses and margin are added.

Plan bids, payments, and rebates. Each year, MA plans submit bids that summarize the projected revenue requirement for providing the standard Medicare benefit package. Let PC_{pct} denote projected medical costs, A_{pct} administrative costs, and M_{pct} the gain/loss margin. Conceptually, the plan bid can be written as

$$b_{pct} = PC_{pct} + A_{pct} + M_{pct},$$

where b_{pct} denotes plan p's bid in county c and year t.

Let B_{ct} denote the county benchmark. If $b_{pct} \leq B_{ct}$, the plan's base payment for standard Part A and Part B benefits equals the bid and the plan also receives a rebate equal to a statutory share of the benchmark–bid difference, which must be returned to enrollees through supplemental benefits or lower cost sharing. If $b_{pct} > B_{ct}$, the benchmark caps Medicare's base payment and beneficiaries pay the difference through the basic premium. In stylized form,

$$\text{Base Payment}_{pct} = \min\{b_{pct}, B_{ct}\} \quad \text{and} \quad \text{Rebate}_{pct} = \lambda_{pct}\,(B_{ct} - b_{pct}) \cdot \mathbf{1}\{b_{pct} < B_{ct}\},$$

where λ_{pct} is the statutory rebate share, which depends on the plan's star rating in practice. Plans then choose provider networks and plan benefits (premiums, cost sharing, supplemental benefits), and these choices respond both to the payment environment summarized by B_{ct} and to the underlying projected cost of covering MA enrollees.

Plan-level benchmarks. The MA benchmark in county c and year t depends on the past FFS spending for the TM beneficiaries in county c, and is computed as

$$B_{ct} = q_{ct} \times FFS_{ct-1} \tag{1}$$

where FFS_{ct-1} is the average per-capita TM FFS spending among TM beneficiaries in the county (risk-adjusted and standardized in the benchmark inputs), q_{ct} is the statutory benchmark factor [2]

Because MA plans typically operate in multi-county service areas, the relevant payment environment for plan p in year t is a plan-level benchmark constructed as the enrollment-weighted average of the county benchmarks in its service area. Let A_{pt} denote the set of counties in which plan p enrolls beneficiaries in year t, and let $w_{pc,t}$ be the share of plan p's enrollment in year t residing in county c, so that $\sum_{c \in A_{pt}} w_{pc,t} = 1$. The plan-level benchmark is then

$$B_{pt} \equiv \sum_{c \in A_{pt}} w_{pc,t} \, B_{ct}, \tag{2}$$

where B_{ct} is the county benchmark defined in (1).

3.3 Provider Channel and Benchmark Channel

MA plan outcomes (claims, coverage, generosity, exit) might be affected by changes in ACOs' provider practice behavior through two main channels:

1. **Benchmark channel.** The key institutional linkage to MSSP arises because the MA county benchmark is mechanically tied to county TM FFS spending. Because B_{ct} is proportional to FFS_{ct}, MSSP-induced changes in TM spending translate mechanically into changes in MA benchmarks. A rich literature shows that MA plans then adjust entry, coverage, and benefit generosity in response to the induced benchmark shocks.

2. **Provider channel.** When the same provider treats both ACO-assigned beneficiaries and also MA enrollees, MSSP-induced practice changes designed for the former may also affect the MA beneficiaries. If the MA plans take this into account, MSSP incentives might also impact MA claims and, more generally, MA plan outcomes Y_{pt}, even when the benchmark B_{pt} is held fixed.

As illustrated in Figure 1, any MSSP policy can affect MA plan outcomes through both a provider-side route and a benchmark-related payment route. In this paper, we do not claim to separately identify the magnitude of the benchmark channel. Instead, the empirical strategy estimates two related objects. The specification that conditions on the plan-level benchmark B_{pt} asks whether MSSP exposure affects MA outcomes holding fixed the payment environment faced by the plan; we interpret this as the benchmark-controlled provider-side effect. The specification that omits B_{pt} captures the overall IV effect of MSSP exposure on MA outcomes. Comparing the two is informative about whether payment- environment

[2] There is also a plan quality adjustment, StarBonus_{ct}, that captures the bonus associated with quality ratings.

variation matters for the estimated response to MSSP exposure, but the difference should not be read as a separately identified estimate of the benchmark channel.

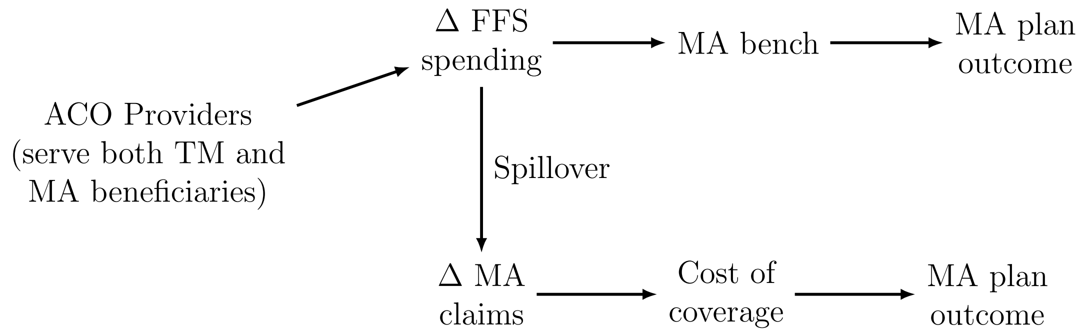

Figure 1: Two channels from MSSP to MA outcomes: benchmark channel and provider channel. The main specification conditions on B_{pt} to isolate the provider channel.

3.4 Dual-Role Health Systems

An important feature of many local markets is that some health systems operate on both sides of the linkage between MSSP and MA. In some cases, the same parent organization sponsors one or more MA contracts while also operating an MSSP ACO that includes its employed physicians and affiliated hospitals. These "dual-role" systems therefore both influence county-level FFS spending through MSSP participation and receive capitated MA payments that are tied to that same FFS spending through equation (1).

This dual role creates a distinctive set of incentives. For an MSSP ACO operated by a system with no MA presence, reductions in FFS_{ct} are unambiguously attractive: they generate shared savings on the TM side and do not directly reduce any internal capitation revenue. By contrast, when a health system sponsors MA contracts in the same county, lower FFS spending mechanically reduces the MA benchmark B_{ct}, thereby shrinking the flow of benchmark-based payments to the system's plans. From the perspective of such a dual-role organization, the marginal benefit of reducing FFS spending therefore reflects a trade-off between: (i) shared savings earned on the MSSP side; and (ii) the reduction in MA revenues associated with lower benchmark payments.

This incentive trade-off implies that provider-channel spillovers may be weaker for health-system-sponsored MA plans than for other plans. The heterogeneity analysis below tests this prediction by allowing the provider-channel effect of MSSP exposure to vary by whether the MA plan is sponsored by a health system (Section 6.1).

4 Data and Descriptive Statistics

The primary analysis sample is a plan-year panel spanning 2013–2023 (coverage years vary by outcome depending on data availability). We construct (i) plan-level outcomes; (ii) county- and plan-level MA benchmarks; (iii) plan-year exposure to MSSP ACOs, π_{pt}^{ACO}; and (iv) descriptive (non-causal) associations between π_{pt}^{ACO} and MA outcomes. The Pathways-based instrument is constructed and discussed in Section 5.3.

4.1 Main data sources and key variables

MSSP alignment and county-level ACO penetration. From CMS MSSP public-use files and attribution/tabulation products, we observe for each county c and year t the number of FFS beneficiaries aligned to each ACO j, denoted N_{jct}. This means N_{jct} beneficiaries are assigned to an ACO based on where they receive their primary care. The total number of FFS beneficiaries in the county, N_{ct}^{FFS}. We denote the ACO j share in county c in year t with $s_{jct} = N_{jct}/N_{ct}^{FFS}$ and compute the county-level MSSP penetration ("ACOs' market share") as the share of FFS beneficiaries aligned to any MSSP ACO:

$$S_{ct} \equiv \sum_{j \in J_c} s_{jct}. \tag{3}$$

where J_c denotes the set of ACOs that have beneficiaries in county c. We also observe ACO entry cohorts, which are used (together with pre-2019 alignment shares) in the shift-share instrument described in Section 5.3.

MA plan variables County benchmarks B_{ct}, statutory benchmark factors q_{ct}, and star-bonus parameters are taken from CMS ratebook files. We construct plan-level benchmarks B_{pt} using plan-county enrollment shares as described in Section 3.2.

Plan identifiers, service-area counties, enrollment, plan type, and average risk scores are drawn from CMS enrollment and plan characteristics files. We construct plan-county enrollment shares $w_{pc,t}$ from plan-county enrollment counts.

Projected costs, bids, and rebates are obtained from CMS bid pricing tools (BPT). The BPT is the actuarial workbook MA organizations submit to CMS each year when they develop annual Medicare Advantage bids. In the BPT files we use for this paper, the relevant variables include projected allowed-cost measures for detailed service categories, credibility weights, blended projections, total medical-expense measures, bids, and rebates. In particular, our projected-cost outcome is built from the total-medical-expenses projected allowed-cost PMPM measure in the BPT. We therefore use these files to recover plans' forward-looking projected medical costs and related pricing objects, even though the specific administrative-cost and margin components are not separately used in our analysis. In our analytic data, projected costs can be consistently recovered only through 2021. This is an important limitation for interpretation: the post-2019 window for projected costs is therefore short and overlaps almost entirely with the COVID period. For that reason, the projected-cost analysis is interpreted cautiously and we also report specifications that exclude 2020.

We focus on the following set of plan features: (i) projected costs; (ii) (log) average bids; (iii) (log) average rebates; (iv) out-of-pocket maximums (OOP max); and (v) an indicator for positive premiums, with additional benefit-design variables in ongoing work. Bid and rebate are aggregated to the plan-year level where necessary. Benefit-design variables (e.g., OOP max, premiums, cost sharing) come from CMS benefit and plan characteristics files.

4.2 Plan-level exposure to ACOs

The empirical analysis is conducted at the *plan-year* level. Mirroring the definition of plan-level benchmark, we define plan p's exposure to MSSP penetration in its service area as the enrollment-weighted average of county ACO penetration:

$$\pi_{pt}^{ACO} \equiv \sum_{c \in A_{pt}} w_{pc,t} \, S_{ct}. \tag{4}$$

Intuitively, π_{pt}^{ACO} is high when a plan's enrollees are concentrated in counties where a larger fraction of FFS beneficiaries are aligned to MSSP ACOs. The value of π_{pt}^{ACO} ranges from 0 to 1. A value of $\pi_{pt}^{ACO} = 1$ implies that all fee-for-service beneficiaries in the plan's service area are attributed to ACOs.[3] A 1% increase in π_{pt}^{ACO} can stem from an increase in ACOs penetration in the plan's service area (holding the plan's enrollment shares fixed), or from an increase in the plan's enrollment shares in those counties where the ACOs penetration is

[3] A TM beneficiary is assigned to an ACO if it receives the majority of its primary care services from providers that are part of the ACO.

larger, or from a combination of both forces.

In the regressions we use $\ln \pi_{pt}^{ACO}$; in practice we either (i) restrict to positive exposure observations or (ii) use a log-plus-constant transformation, $\ln(1 + \pi_{pt}^{ACO})$, in specifications that include years/plans with near-zero exposure.

4.3 Descriptive Statistics

Appendix A reports the summary statistics tables. Table 1 reports summary statistics for the ACO exposure variable π_{pt}^{ACO}, which ranges from 0 to 0.87 with a mean of 0.298 and a standard deviation of 0.146, indicating substantial variation across plan-years. Table 2 shows average plan outcomes across quartiles of ACO exposure for the main analysis sample. In the raw summary statistics, bids are positively correlated with ACO exposure: average bids rise from 790.3 in the lowest-exposure quartile to 829.8 in the highest-exposure quartile. This pattern is unlikely to be causal and instead suggests selection into higher-cost markets: ACOs appear to be more prevalent in geographic areas with richer payment environments. Table 4 reinforces this point, showing that benchmarks are higher in counties with greater ACO exposure. Since benchmarks summarize the MA payment environment and are themselves strongly related to plan bids and benefit design, this descriptive pattern motivates the empirical strategy below. In particular, it suggests that specifications that omit the benchmark may mix provider-side spillovers with positive payment-environment variation, whereas benchmark-controlled specifications are better suited to isolate how outcomes respond to MSSP exposure while holding the payment environment fixed.

Since projected costs are available for a shorter set of years, we report them separately in **??**. That table shows average projected costs across quartiles of ACO exposure in the projected-cost sample. The pattern is less clearly monotone than for bids, but projected costs remain of similar magnitude across quartiles, ranging from 865.8 in the lowest-exposure quartile to 892.5 and 890.7 in the middle quartiles, before declining to 872.0 in the highest-exposure quartile. We therefore treat the projected-cost quartile table as descriptive only and interpret it in light of the smaller sample available for that outcome.

Figure 3 and Figure 4, reported in Appendix A, show county-level ACO penetration (S_{c2018}) and the share of Medicare beneficiaries enrolled in MA in 2018. These maps illustrate substantial geographic heterogeneity in both variables. As discussed above, variation in both county ACO penetration and MA plan enrollment patterns contributes to the variation in our main variable of interest, ACO exposure (π_{pt}^{ACO}).

To illustrate the geographic overlap between the two programs more directly, Figure 2 plots a residualized binscatter of county-level Medicare Advantage penetration against county-level MSSP penetration. MA penetration is measured as the share of Medicare beneficiaries enrolled in Medicare Advantage, while MSSP penetration is measured as the share of Traditional Medicare beneficiaries aligned to an MSSP ACO.

Both variables are residualized after controlling for beneficiary composition, local healthcare supply and utilization, county-level payment conditions, hospital market concentration, and year effects. The figure includes bin-specific 95% confidence intervals together with the fitted linear relationship. The residualized association is small and slightly negative, indicating that once these observable county characteristics are accounted for, counties with greater MSSP penetration do not systematically exhibit higher MA penetration. This pattern is descriptive rather than causal, but it is useful because it suggests that any spillovers we estimate are unlikely to be driven simply by the two programs being disproportionately concentrated in the same observable local markets.

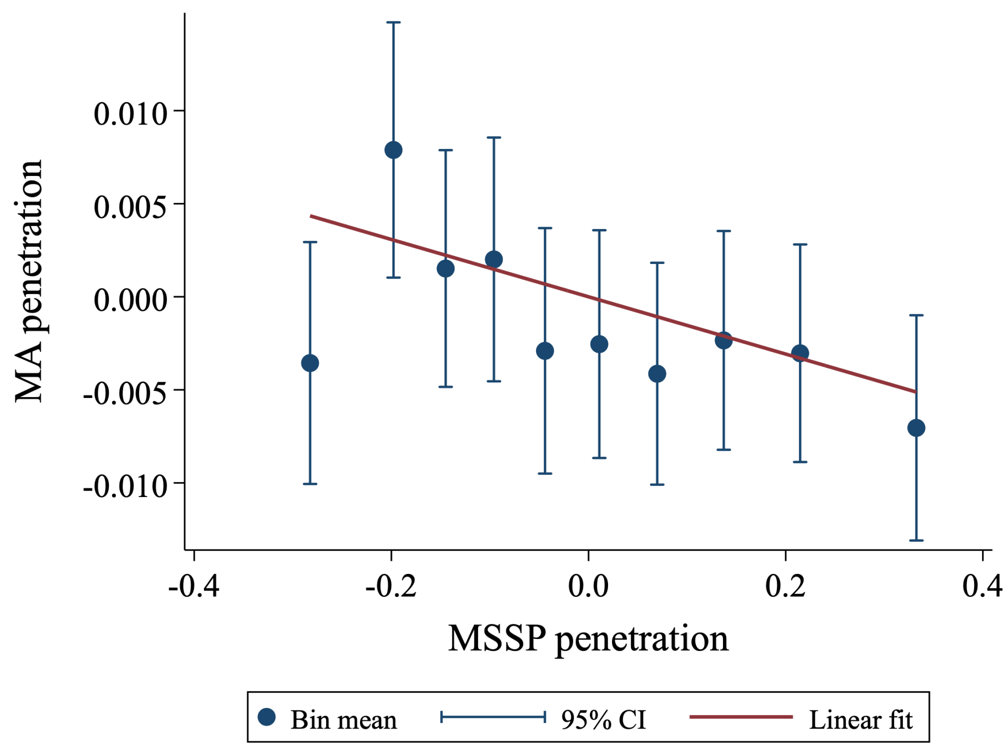

Figure 2: MA Penetration vs. MSSP Penetration

Notes: Residualized association between county MSSP penetration and county Medicare Advantage penetration. The figure reports bin means, 95% confidence intervals, and the fitted linear relationship.

5 Empirical Strategy

This section describes the estimating equations, the sources of endogeneity, the construction of the shift-share instrument, and the identifying assumptions. Our empirical strategy estimates two related causal objects: a benchmark-controlled provider-side effect that conditions on the plan-level payment environment, and an overall IV effect of plan-level MSSP exposure on MA plan outcomes when benchmark-linked payment variation is allowed to move jointly with exposure. The comparison between the two specifications is useful for showing whether payment-environment variation materially affects the estimated response to MSSP exposure, but it does not separately identify the magnitude of the benchmark channel.

5.1 Estimating Equation and Parameter of Interest

Let Y_{pt} denote an outcome for MA plan p in year t (e.g., projected costs, bids, rebates, premiums, or out-of-pocket maxima). We estimate two closely related specifications. We begin with the benchmark-controlled specification:

$$Y_{pt} = \theta^P \ln \pi_{pt}^{ACO} + \rho \ln B_{pt} + X_{pt}'\gamma + \alpha_p + \phi_{c(p)} + \eta_{g(p)} + \lambda_t + \varepsilon_{pt}, \tag{5}$$

where π_{pt}^{ACO} is plan-level exposure to MSSP activity, B_{pt} denotes the plan-level benchmark, X_{pt} includes average risk scores and area-level demographic and healthcare-supply characteristics, α_p denotes plan fixed effects, $\phi_{c(p)}$ contract fixed effects, and $\eta_{g(p)}$ and λ_t denote additive CBSA and year fixed effects. In the main benchmark-controlled specification, the payment environment enters parsimoniously through $\ln B_{pt}$. This keeps the interpretation transparent and aligns the benchmark control across outcomes. In Appendix 7, we show that the main results are robust to more flexible benchmark controls, including a quadratic in $\ln B_{pt}$, year-specific slopes on $\ln B_{pt}$, and benchmark-decile fixed effects.

We then estimate the specification without the benchmark control:

$$Y_{pt} = \theta^O \ln \pi_{pt}^{ACO} + X_{pt}'\gamma + \alpha_p + \phi_{c(p)} + \eta_{g(p)} + \lambda_t + u_{pt}, \tag{6}$$

which we interpret as the total-effect specification.

For each equation, we also report a tighter fixed-effects specification that replaces the additive CBSA and year fixed effects, $\eta_{g(p)} + \lambda_t$, with the full set of CBSA $\times$ year fixed effects, $\kappa_{g(p)\times t}$. This specification absorbs shocks common to all plans operating in the same local market and year. Under additive fixed effects, identification uses both between-market and within-market-year variation in exposure. Under CBSA $\times$ year fixed effects, identification is stricter and comes from differences in MSSP exposure across plans operating in the same market-year.

The coefficient θ^P captures the effect of MSSP exposure conditional on the plan-level payment environment. Because equation (5) holds fixed $\ln B_{pt}$, it isolates the provider channel under the maintained interpretation that remaining variation in outcomes operates through provider-side spillovers rather than the benchmark/payment channel.

The coefficient θ^O captures the overall IV effect of MSSP exposure on the outcome. It measures how MA plan outcomes respond to changes in plan-level ACO exposure when benchmark-mediated payment variation is allowed to move jointly with exposure.

This distinction is central for interpretation. The benchmark-controlled specification isolates the provider channel by holding fixed the plan-level payment environment through which benchmark-mediated effects operate. However, this does not separately identify the benchmark channel itself. In particular, the difference $\theta^O - \theta^P$ should not be interpreted as a separately identified estimate of the benchmark channel.

Under the hypothesis that MSSP-induced changes in care coordination, transitional care, and post-acute management spill over to MA enrollees and reduce the cost of coverage, we expect $\theta^P < 0$ for projected costs and bids, and $\theta^P > 0$ for rebates.

5.2 Endogeneity

The key regressor π_{pt}^{ACO} is the enrollment-weighted average of county-level ACO penetration across the plan's service area:

$$\pi_{pt}^{ACO} = \sum_{c \in A_{pt}} w_{pc,t}\, S_{ct}, \tag{7}$$

where $w_{pc,t}$ is the share of plan p's enrollment in county c in year t, and S_{ct} is county c's realized ACO penetration. Both components are endogenous, and each creates a distinct identification challenge.

Endogenous ACO penetration. County-level ACO penetration S_{ct} can be correlated with local demand conditions, provider market structure, or health system consolidation trends that independently affect MA plan outcomes. For example, counties with more integrated delivery systems may both attract ACOs and offer lower-cost care environments for MA enrollees. If integrated markets have lower MA costs for reasons unrelated to MSSP, OLS would conflate provider spillovers with pre-existing market characteristics, likely biasing $\hat{\theta}$ toward finding cost reductions even if the true provider spillover were zero.

Endogenous plan geography. The enrollment weights $w_{pc,t}$ change over time as plans adjust their service areas, marketing strategies, and provider networks. If plans strategically expand into counties where ACO presence reduces costs (or retreat from counties where tightening benchmarks make operations less profitable), these geographic adjustments can mechanically move π_{pt}^{ACO} in ways that are jointly determined with plan outcomes. This source of endogeneity is especially relevant because plan service-area decisions are made simultaneously with bidding and benefit-design choices.

5.3 Instrument Construction

To address both sources of endogeneity, we construct a shift-share (Bartik) instrument that combines predetermined exposure weights with quasi-experimental time variation from the 2019 Pathways to Success reform.

Source of exogenous variation: the non-rebasing year schedule. As described in Section 3.1, the Pathways reform introduced five-year agreement periods with designated *non-rebasing years*. In these years, spending reductions do not mechanically tighten the ACO's future benchmark, strengthening the marginal incentive to cut costs. Crucially, the timing of non-rebasing years is determined by an ACO's *entry cohort* and the national agreement-period calendar—not by local market conditions or plan decisions. We define:

$$K_{jt} \equiv \mathbf{1}\{\text{ACO } j \text{ is in a non-rebasing year at time } t\}, \tag{8}$$

and treat K_{jt} as a policy-determined indicator. Figure 5 in Appendix C, shows when the non-rebasing years occur for each cohort of ACOs. For the set of ACOs that enter the MSSP prior to 2019, we consider the timing of the non-rebasing years to be exogenous as it is determined by the cohort year and the changes introduced with the 2019 Pathway policy.

Step 1: Baseline plan footprints. Let $t_0 = 2018$ denote the last pre-Pathways year. We fix plan-county enrollment weights at their t_0 values: $w_{pc}^{base} \equiv w_{pc,t_0}$, and set $w_{pc}^{base} = 0$ for counties not served by plan p in t_0. Let A_p^{base} denote the set of counties with $w_{pc}^{base} > 0$. By freezing the plan footprint, we eliminate endogenous geographic adjustments from the instrument.

Step 2: Baseline county-by-ACO alignment shares. Let S_{cj}^{base} denote the share of TM assignable beneficiaries in county c aligned to ACO j at the baseline year t_0. This captures the pre-reform "dose" of each ACO in each county.

Step 3: County-level non-rebasing-year exposure. We construct predicted county-year exposure to anti-ratchet incentives as:

$$P_{ct} \equiv \sum_j S_{cj}^{base} K_{jt}. \tag{9}$$

All time variation in P_{ct} is driven by the Pathways schedule K_{jt}; the cross-sectional pattern is pinned down by the baseline ACO composition of each county. In counties where baseline ACOs happen to belong to entry cohorts whose non-rebasing years arrive early, P_{ct} rises sooner; in counties dominated by later-cohort ACOs, the increase comes later.

Step 4: Plan-level shift-share (Bartik) instrument. We aggregate county exposure to the plan level using baseline enrollment weights:

$$Z_{pt} \equiv \sum_{c \in A_p^{base}} w_{pc}^{base} P_{ct} = \sum_{c \in A_p^{base}} w_{pc}^{base} \left(\sum_j S_{cj}^{base} K_{jt} \right). \tag{10}$$

The instrument Z_{pt} removes both sources of endogeneity identified above: (i) the county ACO composition is frozen at baseline, eliminating confounding from post-reform changes in local ACO penetration; and (ii) the plan's geographic footprint is frozen at baseline, eliminating confounding from strategic service-area adjustments. The only source of time variation is the nationally determined non-rebasing year schedule, which switches on and off mechanically for each ACO entry cohort.

5.4 IV Estimation

We estimate both specifications in Equation 5 and Equation 6 using Z_{pt} (or its log transform) as an excluded instrument for $\ln \pi_{pt}^{ACO}$. For the benchmark-controlled specification, the first stage under additive CBSA and year fixed effects is

$$\ln \pi_{pt}^{ACO} = \mu Z_{pt} + \varrho \ln B_{pt} + X_{pt}' \rho + \tilde{\alpha}_p + \tilde{\phi}_{c(p)} + \tilde{\eta}_{g(p)} + \tilde{\lambda}_t + \eta_{pt}. \tag{11}$$

For the specification without the benchmark control, the corresponding first stage is

$$\ln \pi_{pt}^{ACO} = \tilde{\mu} Z_{pt} + X_{pt}' \tilde{\rho} + \tilde{\alpha}_p + \tilde{\phi}_{c(p)} + \tilde{\eta}_{g(p)} + \tilde{\lambda}_t + \tilde{u}_{pt}. \tag{12}$$

In the tighter specifications, we replace the additive fixed effects $\tilde{\eta}_{g(p)} + \tilde{\lambda}_t$ with CBSA $\times$ year fixed effects. Thus, the first stage is estimated under the same two fixed-effects structures used in the reduced form and second stage.

We expect the coefficient on the instrument to be positive: plans whose baseline service areas load more heavily on ACO cohorts that are in non-rebasing years should face stronger MSSP incentives and therefore higher realized exposure.

Under the identifying assumptions below, the IV coefficient from the specification without $\ln B_{pt}$ identifies the overall IV effect of Pathways-induced variation in MSSP exposure on MA outcomes. The IV coefficient from the benchmark-controlled specification identifies how MA outcomes respond to that same quasi-experimental variation in MSSP exposure while holding fixed the benchmark/payment environment. We interpret the latter as the provider-channel effect.

In all specifications, we control for a rich set of plan-service-area-level characteristics: beneficiary demographics and health status, healthcare supply and FFS utilization, hospital market concentration, and the relevant fixed effects implied by the specification.

5.5 Identifying Assumptions

Identification requires two standard conditions. First, the Pathways-based instrument must be relevant for realized plan-level MSSP exposure; we assess this using the Kleibergen–Paap first-stage F-statistic reported for each specification. Second, conditional on the included controls and fixed effects, the instrument must affect MA outcomes only through its effect on realized MSSP exposure. The remainder of this subsection discusses why that exclusion restriction is plausible in our setting and how we assess it empirically.

Following the logic of Goldsmith-Pinkham et al. (2020), our identifying assumption is that the baseline exposure shares—the plan-county enrollment weights w_{pc}^{base} and county-ACO alignment shares S_{cj}^{base}—are exogenous conditional on the included fixed effects and controls. Because both are measured in the pre-Pathways period and the panel includes plan fixed effects, the key remaining requirement is that the policy-driven component of the instrument be uncorrelated with post-2018 changes in unobserved determinants of MA plan outcomes. This is the relevant exclusion restriction in the presence of possible time-varying beneficiary selection. If plans with high realized π_{pt}^{ACO} also experienced differential post-2018 changes in unobserved enrollee health, coding intensity, or plan demand, OLS could remain biased even after conditioning on plan fixed effects.

We view this exclusion restriction as plausible for two reasons. First, non-rebasing-year timing is mechanically determined by MSSP entry cohort and calendar time, so ACOs cannot choose or manipulate when they enter a non-rebasing year. Second, the Bartik instrument freezes both sides of the shift-share structure at their 2018 values: plan-county enrollment weights and county-ACO alignment shares. The only time-varying component is K_{jt}, which moves mechanically with the Pathways reform and the ACO's entry cohort.

The comparison between additive CBSA and year fixed effects and CBSA $\times$ year fixed effects is also part of the identification strategy. The additive structure removes time-invariant geographic differences and common national shocks, but it does not absorb local market shocks that evolve differently across CBSAs over time. Because π_{pt}^{ACO} is constructed from county ACO penetration in the plan's service area, such market-year shocks are a natural concern. The CBSA $\times$ year specification therefore provides a tighter design: it compares plans with different MSSP exposure within the same local market and year. If estimates that appear under additive fixed effects OLS shrink sharply once CBSA $\times$ year fixed effects are included, that is consistent with market-year confounding rather than causal spillovers. By contrast, IV estimates that remain significant under the tighter fixed-effects structure are more plausibly driven by the policy-based component of exposure. Therefore, the source of identifying variation in the preferred specifications comes from within-market-year differences in predicted MSSP exposure generated by the Bartik instrument, rather than from raw cross-market differences in realized exposure.

Under these assumptions, the benchmark-controlled specification identifies how MA outcomes respond to the same quasi-experimental variation in MSSP exposure while holding fixed the plan-level payment environment. We interpret this as the provider-channel effect. The specification that omits B_{pt} instead identifies a benchmark-uncontrolled total effect of MSSP exposure on MA outcomes. The comparison between the two specifications is informative about whether benchmark-linked payment variation matters empirically, but it does not separately identify the benchmark channel in a causal sense. The difficulty is not only that the same policy-driven variation can move provider-channel incentives and benchmark-linked payments jointly. It is also that the plan-level benchmark is itself a downstream object shaped by local fee-for-service spending and plan geography, so conditioning on it can absorb part of the overall policy response rather than isolate a standalone benchmark effect. Identifying that effect separately would require additional variation that shifts plan benchmarks without also changing provider-channel incentives, or stronger structural assumptions that make the two channels separately recoverable.

Instrument Validation A remaining concern is that even the policy-driven shift-share instrument could proxy for differential pre-existing trends if plans that are more exposed to early Pathways cohorts were already on different trajectories before 2019. This concern is especially relevant for outcomes such as bids, which may respond to persistent differences in market conditions, contracting environments, or insurer strategy that are correlated with the baseline geography embedded in the shift-share design. To assess this directly, we implement a dose-response event-study validation exercise. We construct a time-invariant plan-level dose from post-reform instrument exposure and estimate whether higher-dose plans already exhibited differential pre-trends in each outcome before the 2019 reform. The identifying prediction is that pre-2019 lead coefficients should be close to zero if the instrument is not picking up pre-existing outcome trends.

The event-study plots are presented in section 7. For projected costs and bids, the pre-period coefficients are reassuringly flat, which strengthens the credibility of the projected-cost results despite the short post-reform window. For rebates, however, the event-study plots indicate evidence of differential pre-trends. We therefore do not interpret the bid IV coefficients as clean causal estimates in the same way as the projected-cost estimates. Instead, we view them as suggestive evidence that is directionally consistent with the projected-cost results. The appendix reports the full event-study figures for all outcomes and the benchmark-control robustness exercises.

These assumptions are further assessed in section 7. We report dose-response event studies showing no systematic differential pre-trends prior to the 2019 Pathways reform and the expected post-reform divergence in both plan outcomes and the intermediate exposure measure. We also show that the benchmark-controlled estimates are robust to more flexible specifications for the plan-level benchmark.

6 Results

This section reports both OLS and IV estimates of the effect of plan-level MSSP exposure on Medicare Advantage (MA) plan outcomes. We present as separate table with two panels for each outcome variable. Panel A presents the benchmark-controlled specification and, under our identification assumptions, the estimated coefficient capture the instrument-induce LATE of MSSP exposure through the provider channel. The main specification controls for the plan-level payment environment using $\ln B_{pt}$. Online appendix shows that the conclusions are robust to more flexible benchmark controls. Panel B omits the benchmark control should be interpreted as a *total-effect* object. The comparison between the two specifications is informative about whether benchmark-linked payment variation matters empirically, but it does not separately identify the benchmark channel in a causal sense.

Within each panel, columns (1)–(2) use additive CBSA and year fixed effects and report OLS and IV estimates, respectively, while columns (3)–(4) replace these with CBSA × year fixed effects and again report OLS and IV estimates. A central comparison is between the additive and interacted fixed-effects structures. Additive CBSA and year fixed effects remove time-invariant market differences and common national shocks, but they do not absorb local market shocks that evolve differently across CBSAs over time. By contrast, the CBSA × year specification compares plans facing different MSSP exposure within the same local market and year. For that reason, we place the greatest weight on the IV estimates with CBSA × year fixed effects. When OLS relationships attenuate once CBSA × year fixed effects are included, but IV estimates remain economically meaningful, we interpret that pattern as evidence that realized exposure is contaminated by market-year confounding whereas the Bartik instrument isolates the policy-driven component of exposure.

Projected costs. Table 6 and Table 7 report the projected-cost regressions. This is the most direct outcome for the provider-channel mechanism because projected costs are forward-looking expected medical costs from the bid pricing tools rather than realized ex post spending. Across specifications, the IV coefficients are negative and larger in magnitude than the corresponding OLS coefficients, which remain close to zero and generally imprecise. In the full sample, the negative IV effect is clearest in the additive fixed-effects specifications, while the stricter CBSA × year specification is less precise. We do not view that loss of precision as surprising: projected costs are only observed through 2021 in our analytic data, so the post-reform window is short and heavily overlaps with the COVID period. Reassuringly, once 2020 is excluded, the projected-cost effect becomes statistically meaningful under CBSA × year fixed effects. In that specification, a 1% increase in MSSP exposure lowers projected costs by about 0.34% in the provider-channel specification and by about 0.35% in the total-effect specification. We interpret this as direct evidence that greater MSSP exposure lowers plans' expected cost of covering MA beneficiaries.

Bids. Table 5 reports the log-bid regressions. The most informative comparison is again between OLS and IV across fixed-effects structures. Under the tighter CBSA × year specification, the OLS coefficients become small and statistically insignificant, while the IV coefficients remain negative and economically meaningful. In the benchmark-controlled provider-channel specification, a 1% increase in MSSP exposure lowers bids by about 0.20% in the preferred IV specification with CBSA × year fixed effects. The corresponding total-effect estimate is also negative. This pattern is consistent with the interpretation that non-instrumented exposure is confounded by market-year shocks, whereas the policy-driven IV variation isolates a genuine effect of MSSP exposure on the broad revenue requirement of offering MA coverage.

We also interpret bids cautiously. An MA bid is a broad plan-level object, not a pure medical-claims measure. It combines projected medical costs with administrative costs and margin, so the bid estimates should be read as evidence that MSSP exposure lowers the overall revenue requirement of standard MA coverage rather than as a pure estimate of claims savings alone. The negative bid estimates are also consistent with the projected-cost results, since projected medical costs are a major component of bids.

Rebates. Table 8 reports the log-rebate regressions. The IV coefficients are positive and economically large in both specifications, and under the preferred CBSA × year fixed-effects structure they suggest sizable increases in rebates when MSSP exposure rises. Interpreted mechanically, that pattern is consistent with the idea that lower expected costs may ultimately translate into more generous benefits through the rebate formula. However, the event-study validation is less reassuring for rebates than for projected costs and bids. In particular, the pre-2019 coefficients in the rebate event-study are not as close to zero, suggesting residual differential trends across high-dose and low-dose plans before the Pathways schedule takes effect. We therefore interpret the rebate estimates as suggestive rather than as part of the paper's sharpest causal evidence.

Out-of-pocket maximums. Table 9 reports the log regressions for out-of-pocket maxima. The IV point estimates are negative, but they are imprecisely estimated and statistically indistinguishable from zero. The OLS coefficients are likewise small. We therefore do not view out-of-pocket limits as a margin on which the data provide sharp evidence of spillovers.

Discrete plan features: premiums, deductibles, and copays. The results for positive-premium status, deductibles, PCP copays, and specialist copays are reported in Table 10,

Table 11, Table 12, and Table 13. Across these margins, the IV estimates are generally imprecise and statistically indistinguishable from zero, and the OLS coefficients are likewise small once the tighter fixed-effects structure is imposed. The clearest spillover effects therefore appear on the broad continuous plan margins—projected costs, bids, and rebates—rather than on narrower discrete plan-design margins.

Taken together, the main-results tables show a clear pattern. MSSP exposure has its most credible effects on broad plan-level objects that summarize expected costs and the overall revenue requirement of MA coverage. In IV, projected costs fall and bids fall, especially in the preferred specifications with CBSA × year fixed effects, while the evidence for out-of-pocket maxima and discrete plan-design features is much weaker. The rebate coefficients are positive and economically large, but the event-study validation indicates pre-policy differential trends for that outcome, so we do not place the same weight on rebates as on projected costs and bids. Overall, the strongest evidence is that greater MSSP exposure lowers expected cost of coverage and lowers bids; the evidence on rebate pass-through is more tentative.

6.1 Heterogeneity by Dual-Role Systems

We next examine whether the spillover estimates differ by organizational structure. As discussed in Section 3.4, some health systems operate on both sides of the MSSP–MA linkage: they participate in MSSP through affiliated ACOs while also sponsoring MA plans whose payments are tied to the same local FFS spending through the benchmark formula. For these dual-role systems, reductions in FFS spending generate shared-savings rewards on the TM side but can also reduce benchmark-based MA payments. This creates a trade-off that may weaken the incentive to translate MSSP exposure into lower MA costs, implying smaller spillover responses for health-system-sponsored plans.

To study this heterogeneity, we estimate the same two IV specifications used in the main analysis separately for non-dual-role plans and dual-role plans. Throughout this subsection, we focus on the tighter CBSA × year fixed-effects structure. Within each subsample, the benchmark-controlled provider-channel specification includes $\ln B_{pt}$ exactly as in the main analysis, while the benchmark-uncontrolled total-effect specification omits it. In both cases, $\ln \pi_{pt}^{ACO}$ is instrumented with the same Pathways-based Bartik instrument, estimated separately within each subsample. Online appendix shows that the conclusions are not sensitive to using more flexible benchmark controls.

For each group $g \in \{\text{Non-Dual Role}, \text{Dual Role}\}$, we estimate

$$Y_{pt} = \beta_g \ln \pi_{pt}^{ACO} + \rho_g \ln B_{pt} + X_{pt}' \Gamma_g + \alpha_p + \phi_{c(p)} + \kappa_{g(p)\times t} + \varepsilon_{pt}, \tag{13}$$

where the benchmark term $\ln B_{pt}$ is included only in the provider-channel specification and omitted in the total-effect specification. The coefficient β_g therefore measures the relevant IV effect of MSSP exposure within each organizational form.

The projected-cost results point in the same direction as the main analysis. In Table 15 and especially in Table 16, the non-dual-role estimates are negative in both the provider-channel and total-effect specifications, whereas the dual-role estimates are much smaller and imprecisely estimated. The 2020-excluded specification is particularly informative because projected costs are only observed through 2021 in our data. This pattern suggests that the direct cost-reduction effects of MSSP exposure are concentrated among plans that do not internalize both sides of the MSSP–MA linkage.

The strongest heterogeneity appears in bids. Table 14 shows that, among non-dual-role plans, greater MSSP exposure is associated with lower bids in both the benchmark-controlled provider-channel and total-effect specifications. By contrast, the corresponding bid estimates for dual-role plans are much smaller and statistically indistinguishable from zero. This pattern is consistent with the idea that organizations internalizing both sides of the MSSP–MA linkage have weaker incentives to translate MSSP-related cost reductions into lower MA bids.

The evidence for the remaining outcomes is weaker. In Table 17, the rebate coefficients are positive for both groups and somewhat larger for non-dual-role plans, but they are imprecisely estimated. Moreover, because the main-sample event-study validation is less supportive for rebates than for projected costs and bids, we interpret these rebate heterogeneity patterns cautiously. In Table 18, the estimates are small and imprecise for both groups. The same applies to the binary plan features reported in Table 19, Table 20, Table 21, and Table 22. Overall, the dual-role heterogeneity is clearest for projected costs and bids, while the evidence for rebates, out-of-pocket limits, and discrete plan-design margins is much more limited.

Taken together, the split-sample IV results suggest that provider-channel spillovers from MSSP into MA are strongest for plans that do not belong to organizations internalizing both sides of the MSSP–MA linkage. For dual-role plans, the corresponding estimates are generally smaller and less precise. The clearest heterogeneity therefore appears on projected costs and bids. By contrast, the evidence for rebates, out-of-pocket limits, and discrete plan-design features is more limited, and the rebate results in particular should be interpreted cautiously in light of the validation evidence.

7 Conclusion

This paper studies whether incentives created by the Medicare Shared Savings Program spill over to Medicare Advantage. The central question is whether providers that respond to MSSP incentives also change care delivery in ways that affect MA plans operating in the same markets. We refer to this mechanism as the *provider channel*. Because MA plan payments are mechanically tied to Traditional Medicare spending through the county-level benchmark, MSSP can also affect MA through a *benchmark channel*. Our empirical strategy isolates the causal effect of MSSP exposure through the *provider channel* by using an instrumental-variables design and controlling for the plan-level benchmark.

To study this question, we construct a plan-level measure of MSSP exposure based on the enrollment-weighted ACO penetration in the counties where a plan operates. A central challenge is that this exposure measure is endogenous: plans with higher exposure may differ systematically in enrollee composition, provider environment, or local market conditions. To address this, we use a Bartik-style instrument that combines predetermined exposure shares with quasi-experimental variation generated by the 2019 Pathways to Success MSSP reform. Throughout, we place the greatest weight on IV specifications with CBSA $\times$ year fixed effects, since these absorb shocks common to plans operating in the same local market and year.

The results indicate that greater MSSP exposure lowers MA plans' expected cost of coverage and, in turn, lowers plan payments. These findings provide evidence that MSSP affects MA not only through benchmark-linked payment rules, but also through provider-side spillovers that change the expected cost of serving MA enrollees. By contrast, the evidence on rebates is less conclusive. Although the rebate coefficients are positive and economically large, the event-study validation suggests residual pre-policy differential trends for that outcome, so we interpret the rebate results more cautiously.

Overall, the paper shows that Medicare payment reforms cannot be evaluated in isolation. When providers operate across multiple Medicare programs, incentives targeted at one segment can affect outcomes in another. Our findings suggest that MSSP has meaningful spillover effects on Medicare Advantage through shared providers, and that understanding those effects is important for evaluating the broader incidence of Medicare payment reform.

References

Baicker, K., Chernew, M. E., and Robbins, J. A. (2013). The spillover effects of Medicare managed care: Medicare Advantage and hospital utilization. *Journal of Health Economics*, 32(6):1289–1300.

Borusyak, K., Hull, P., and Jaravel, X. (2022). Quasi-experimental shift-share research designs. *Review of Economic Studies*, 89(1):181–213.

Borusyak, K., Hull, P., and Jaravel, X. (2025). A practical guide to shift-share instruments. *Journal of Economic Perspectives*, 39(1):181–204.

Callison, K. (2016). Medicare managed care spillovers and treatment intensity. *Health Economics*, 25(7):873–887.

Chernew, M., DeCicca, P., and Town, R. (2008). Managed care and medical expenditures of Medicare beneficiaries. *Journal of Health Economics*, 27(6):1451–1461.

Colla, C. H., Lewis, V. A., Kao, L.-S., O'Malley, A. J., Chang, C.-H., and Fisher, E. S. (2016). Association between Medicare accountable care organization implementation and spending among clinically vulnerable beneficiaries. *JAMA Internal Medicine*, 176(8):1167–1175.

Einav, L., Finkelstein, A., Ji, Y., and Mahoney, N. (2020). Randomized trial shows healthcare payment reform has equal-sized spillover effects on patients not targeted by reform. *Proceedings of the National Academy of Sciences*, 117(32):18939–18947.

Feyman, Y., Pizer, S. D., and Frakt, A. B. (2021). The persistence of Medicare Advantage spillovers in the post-Affordable Care Act era. *Health Economics*, 30(2):311–327.

Finkelstein, A., Gentzkow, M., and Williams, H. (2016). Sources of geographic variation in health care: Evidence from patient migration. *Quarterly Journal of Economics*, 131(4):1681–1726.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624.

Hou, Y., Domino, M. E., Lewis, V. A., Gong, Q., Callison, K., and Trogdon, J. G. (2025). Spending changes after moving to areas with greater ACO participation among nonattributed Medicare beneficiaries. *JAMA Network Open*, 8(2):e2458311.

Markovitz, A. A., Hollingsworth, J. M., Ayanian, J. Z., Norton, E. C., Yan, P. L., and Ryan, A. M. (2019). Performance in the Medicare Shared Savings Program after accounting for nonrandom exit: An instrumental variable analysis. *Annals of Internal Medicine*, 171(1):27–36.

McWilliams, J. M., Hatfield, L. A., Chernew, M. E., Landon, B. E., and Schwartz, A. L. (2016). Early performance of accountable care organizations in Medicare. *New England Journal of Medicine*, 374(24):2357–2366.

McWilliams, J. M., Hatfield, L. A., Landon, B. E., and Chernew, M. E. (2020). Savings or selection? Initial spending reductions in the Medicare Shared Savings Program and considerations for reform. *Milbank Quarterly*, 98(3):847–907.

McWilliams, J. M., Hatfield, L. A., Landon, B. E., Hamed, P., and Chernew, M. E. (2018). Medicare spending after 3 years of the Medicare Shared Savings Program. *New England Journal of Medicine*, 379(12):1139–1149.

Navathe, A. S., Bain, A. M., and Werner, R. M. (2018). Do changes in post-acute care use at hospitals participating in an accountable care organization spillover to all Medicare beneficiaries? *Journal of General Internal Medicine*, 33(6):831–838.

Ouayogódé, M. H., Meara, E., Ho, K., Snyder, C. M., and Colla, C. H. (2021). Estimates of ACO savings in the presence of provider and beneficiary selection. *Healthcare*, 9(1):100460.

Post, B., Ryan, A. M., Moloci, N. M., Li, J., Dupree, J. M., and Hollingsworth, J. M. (2019). Physician participation in Medicare accountable care organizations and spillovers in commercial spending. *Medical Care*, 57(4):305–311.

Ryan, A. M. and Markovitz, A. A. (2023). Estimated savings from the Medicare Shared Savings Program. *JAMA Health Forum*, 4(12):e234449. Retracted and replaced April 2024.

Appendix A

Descriptive Statistics

Table 1: Summary statistics of plan-level exposure to MSSP π_{ct}^{ACO}

	Mean	Median	25th Perc.	75th Perc.	Min	Max	SD
π_{ct}^{ACO}	0.298	0.303	0.196	0.400	0.000	0.871	0.146

Table 2: Summary statistics by quartile of π_{ct}^{ACO}

	Q1	Q2	Q3	Q4	Total
Bid	790.3	806.5	812.0	829.8	808.7
	(109.6)	(108.8)	(103.1)	(105.2)	(107.8)
Rebate	91.66	96.85	104.2	104.8	99.01
	(81.38)	(81.49)	(79.12)	(75.42)	(79.72)
Out-of-Pocket Max	5106.4	5553.1	5471.4	5393.9	5375.4
	(1825.2)	(1645.0)	(1647.1)	(1587.8)	(1692.8)
Premium > 0	0.295	0.246	0.225	0.274	0.261
	(0.456)	(0.431)	(0.418)	(0.446)	(0.439)
PCP copay > 0	1.364	1.316	1.296	1.299	1.320
	(0.481)	(0.465)	(0.457)	(0.458)	(0.467)
Specialist Copay > 0	1.289	1.253	1.229	1.185	1.242
	(0.454)	(0.435)	(0.420)	(0.389)	(0.428)
Observations	4783	4545	4216	4008	17552

Notes: The above table shows the mean of each of the outcome variables across the 4 quartiles of plan-level exposure to MSSP. The standard deviations are shown in parentheses.

Table 3: Summary statistics of Projected Cost by quartile of π_{ct}^{ACO}

	Q1	Q2	Q3	Q4	Total
Projected Cost	865.8	892.5	890.7	872.0	880.2
	(285.0)	(280.4)	(286.8)	(257.6)	(278.8)
Observations	4461	4101	4135	3544	16241

Notes: The above table shows the mean of each of the plan projected costs across the 4 quartiles of plan-level exposure to MSSP. The standard deviations are shown in parentheses.

Table 4: Summary statistics of control variables by quartile of π_{ct}^{ACO}

	Q1	Q2	Q3	Q4	Total
Benchmark	884.3	912.3	924.0	944.9	914.9
	(99.10)	(105.6)	(90.66)	(89.63)	(99.26)
Risk Score	1.141	1.156	1.195	1.140	1.158
	(0.387)	(0.383)	(0.417)	(0.346)	(0.385)
Insurer HHI	8.005	7.977	8.007	8.000	7.997
	(0.355)	(0.318)	(0.264)	(0.272)	(0.307)
Hospital HHI	8.294	8.455	8.518	8.617	8.463
	(0.704)	(0.488)	(0.431)	(0.431)	(0.546)
Household Income	64822.8	64147.4	65376.7	66971.8	65271.6
	(17328.1)	(13978.8)	(15196.8)	(14599.8)	(15413.4)
IP Medicare Payment	9.769	9.785	9.777	9.778	9.777
	(0.107)	(0.107)	(0.0868)	(0.0813)	(0.0969)
OP Medicare Payment	2.059	2.063	2.068	2.072	2.065
	(0.0238)	(0.0227)	(0.0210)	(0.0200)	(0.0225)
Pop share above 65	0.153	0.153	0.161	0.167	0.158
	(0.0315)	(0.0354)	(0.0327)	(0.0340)	(0.0339)
Observations	4783	4545	4216	4008	17552

Notes: The above table shows the mean of each of the control variables across the 4 quartiles of plan-level exposure to MSSP. The standard deviations are shown in parentheses.

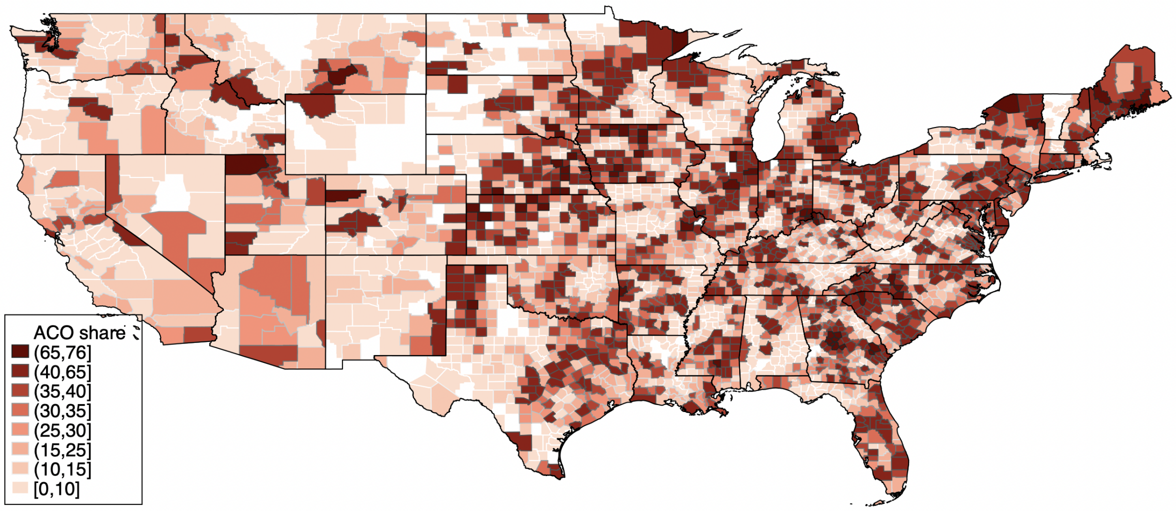

Figure 3: MSSP Penetration in 2018

Notes: This figure shows the MSSP penetration in 2018 across counties. The county MSSP-penetration is calculated as the overall share of TM beneficiaries assigned to any ACO in 2018.

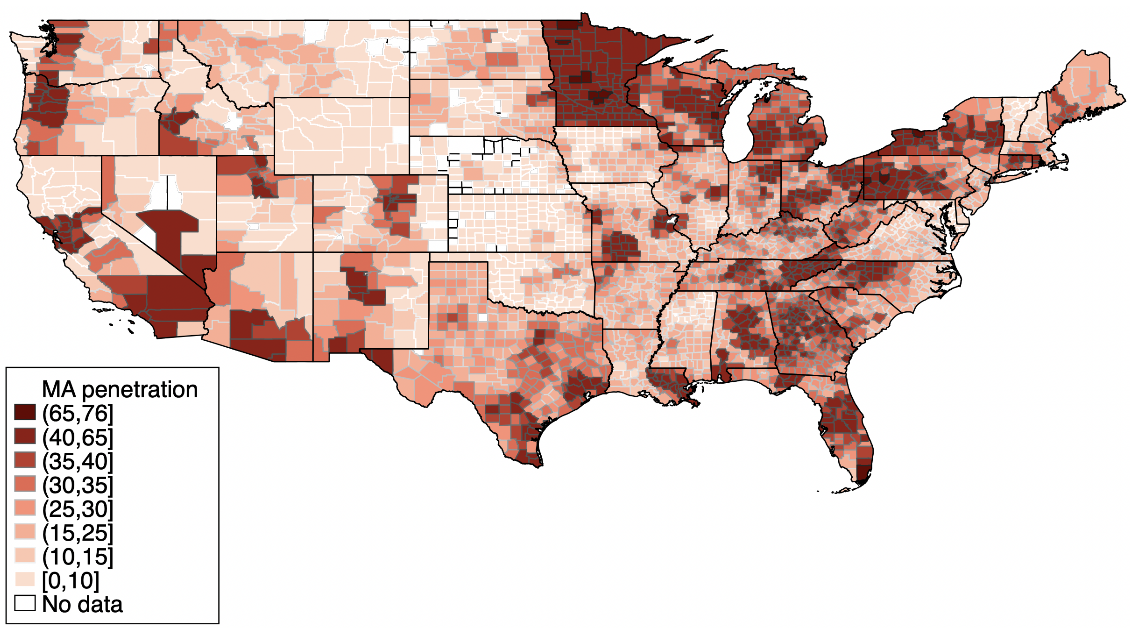

Figure 4: Percentage of Medicare beneficiaries enrolled in MA

Notes: This figure shows the MA penetration in 2018 across counties. The county MA-penetration is calculated as the overall share of MA plans' county share in 2018.

Appendix B

Table 5: Plan bids (log specification)

	Regional + Year FE		Regional × Year FE	
	OLS	IV	OLS	IV
Panel A: Provider Channel				
$\ln \pi^{ACO}$	-0.0336***	-0.145***	-0.0109	-0.202***
	(0.00689)	(0.0315)	(0.0147)	(0.0756)
Plan FE	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	No	No
Year FE	Yes	Yes	No	No
Regional × Year FE	No	No	Yes	Yes
Benchmark	Yes	Yes	Yes	Yes
First stage F		362.4		140.6
Observations	27191	17526	26380	16813
Panel B: Total Effect				
$\ln \pi^{ACO}$	-0.0324***	-0.0941***	-0.00255	-0.221***
	(0.00718)	(0.0323)	(0.0148)	(0.0770)
Plan FE	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	No	No
Year FE	Yes	Yes	No	No
Regional × Year FE	No	No	Yes	Yes
Benchmark	No	No	No	No
First stage F		368.0		137.8
Observations	27191	17526	26380	16813

Notes: The outcome variable in each each specification is the log Bids. Each column reports a plan-year regression. Panel A reports the benchmark-controlled Provider Channel specification, while Panel B reports the benchmark-uncontrolled Total Effect specification. In both panels, columns (1)–(2) use additive core-based statistical area (CBSA) and year fixed effects and report OLS and IV estimates, respectively. Columns (3)–(4) replace these with CBSA × year fixed effects and again report OLS and IV estimates. In Panel A, the benchmark is controlled for using $\ln B_{pt}$; in Panel B, the benchmark is not included. IV columns instrument $\ln \pi_{pt}^{ACO}$ with the Bartik shift-share instrument. All specifications include plan fixed effects and a rich set of plan-service-area controls, including beneficiary demographics and health status, healthcare supply and FFS utilization, and hospital market concentration. Regressions are weighted by log plan enrollment. Standard errors clustered at the plan level are reported in parentheses. Reported first-stage F statistics are the Kleibergen–Paap rk Wald F statistic. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 6: Plan projected costs (log specification)

	Regional + Year FE		Regional × Year FE	
	OLS	IV	OLS	IV
Panel A: Provider Channel				
ln π^{ACO}	-0.00755	-0.118**	-0.00937	-0.192
	(0.0185)	(0.0461)	(0.0422)	(0.148)
Plan FE	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	No	No
Year FE	Yes	Yes	No	No
Regional × Year FE	No	No	Yes	Yes
Benchmark	Yes	Yes	Yes	Yes
First stage F		697.9		186.2
Observations	14202	10352	13681	9888
Panel B: Total Effect				
ln π^{ACO}	-0.00105	-0.0906**	0.00163	-0.193
	(0.0185)	(0.0456)	(0.0423)	(0.148)
Plan FE	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	No	No
Year FE	Yes	Yes	No	No
Regional × Year FE	No	No	Yes	Yes
Benchmark	No	No	No	No
First stage F		718.9		186.7
Observations	14202	10352	13681	9888

Notes: The outcome variable in each each specification is the log Projected Costs. Each column reports a plan-year regression. Panel A reports the benchmark-controlled Provider Channel specification, while Panel B reports the benchmark-uncontrolled Total Effect specification. In both panels, columns (1)–(2) use additive core-based statistical area (CBSA) and year fixed effects and report OLS and IV estimates, respectively. Columns (3)–(4) replace these with CBSA × year fixed effects and again report OLS and IV estimates. In Panel A, the benchmark is controlled for using $\ln B_{pt}$; in Panel B, the benchmark is not included. IV columns instrument $\ln \pi_{pt}^{ACO}$ with the Bartik shift-share instrument. All specifications include plan fixed effects and a rich set of plan-service-area controls, including beneficiary demographics and health status, healthcare supply and FFS utilization, and hospital market concentration. Regressions are weighted by log plan enrollment. Standard errors clustered at the plan level are reported in parentheses. Reported first-stage F statistics are the Kleibergen–Paap rk Wald F statistic. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 7: Projected costs excluding 2020 (log specification)

	Regional + Year FE		Regional × Year FE	
	OLS	IV	OLS	IV
Panel A: Provider Channel				
ln π^{ACO}	-0.0256	-0.128**	-0.0386	-0.343**
	(0.0219)	(0.0549)	(0.0519)	(0.158)
Plan FE	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	No	No
Year FE	Yes	Yes	No	No
Regional × Year FE	No	No	Yes	Yes
Benchmark	Yes	Yes	Yes	Yes
First stage F		523.1		158.6
Observations	8836	7240	8402	6837
Panel B: Total Effect				
ln π^{ACO}	-0.0177	-0.0993*	-0.0295	-0.345**
	(0.0219)	(0.0544)	(0.0520)	(0.159)
Plan FE	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	No	No
Year FE	Yes	Yes	No	No
Regional × Year FE	No	No	Yes	Yes
Benchmark	No	No	No	No
First stage F		551.9		158.9
Observations	8836	7240	8402	6837

Notes: The dependent variable in all specifications is log projected costs. Each column reports a plan-year regression estimated on the sample that excludes 2020, the first COVID year, because projected cost measures may be unusually noisy in that year. Panel A reports the benchmark-controlled Provider Channel specification, while Panel B reports the benchmark-uncontrolled Total Effect specification. In both panels, columns (1)–(2) use additive core-based statistical area (CBSA) and year fixed effects and report OLS and IV estimates, respectively. Columns (3)–(4) replace these with CBSA × year fixed effects and again report OLS and IV estimates. In Panel A, the benchmark is controlled for using $\ln B_{pt}$; in Panel B, the benchmark is not included. IV columns instrument $\ln \pi_{pt}^{ACO}$ with the Bartik shift-share instrument. All specifications include plan fixed effects and a rich set of plan-service-area controls, including beneficiary demographics and health status, healthcare supply and FFS utilization, and hospital market concentration. Regressions are weighted by log plan enrollment. Standard errors clustered at the plan level are reported in parentheses. Reported first-stage F statistics are the Kleibergen–Paap rk Wald F statistic. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 8: Plan rebates (log specification)

	Regional + Year FE		Regional × Year FE	
	OLS	IV	OLS	IV
Panel A: Provider Channel				
ln π^{ACO}	0.251***	1.329***	0.121	2.938***
	(0.0636)	(0.320)	(0.147)	(0.774)
Plan FE	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	No	No
Year FE	Yes	Yes	No	No
Regional × Year FE	No	No	Yes	Yes
Benchmark	Yes	Yes	Yes	Yes
First stage F		299.6		112.9
Observations	24361	15072	23617	14426
Panel B: Total Effect				
ln π^{ACO}	0.259***	1.476***	0.135	2.866***
	(0.0646)	(0.325)	(0.148)	(0.775)
Plan FE	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	No	No
Year FE	Yes	Yes	No	No
Regional × Year FE	No	No	Yes	Yes
Benchmark	No	No	No	No
First stage F		304.9		110.5
Observations	24361	15072	23617	14426

Notes: The outcome variable in each each specification is the log Rebates. Each column reports a plan-year regression. Panel A reports the benchmark-controlled Provider Channel specification, while Panel B reports the benchmark-uncontrolled Total Effect specification. In both panels, columns (1)–(2) use additive core-based statistical area (CBSA) and year fixed effects and report OLS and IV estimates, respectively. Columns (3)–(4) replace these with CBSA × year fixed effects and again report OLS and IV estimates. In Panel A, the benchmark is controlled for using $\ln B_{pt}$; in Panel B, the benchmark is not included. IV columns instrument $\ln \pi_{pt}^{ACO}$ with the Bartik shift-share instrument. All specifications include plan fixed effects and a rich set of plan-service-area controls, including beneficiary demographics and health status, healthcare supply and FFS utilization, and hospital market concentration. Regressions are weighted by log plan enrollment. Standard errors clustered at the plan level are reported in parentheses. Reported first-stage F statistics are the Kleibergen–Paap rk Wald F statistic. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 9: Plan OOP Max (log specification)

	Regional + Year FE		Regional × Year FE	
	OLS	IV	OLS	IV
Panel A: Provider Channel				
ln π^{ACO}	0.0267	-0.329**	0.0319	-0.330
	(0.0404)	(0.150)	(0.0983)	(0.441)
Plan FE	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	No	No
Year FE	Yes	Yes	No	No
Regional × Year FE	No	No	Yes	Yes
Benchmark	Yes	Yes	Yes	Yes
First stage F		362.4		140.6
Observations	27191	17526	26380	16813
Panel B: Total Effect				
ln π^{ACO}	0.0284	-0.250*	0.0408	-0.358
	(0.0407)	(0.151)	(0.0988)	(0.447)
Plan FE	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	No	No
Year FE	Yes	Yes	No	No
Regional × Year FE	No	No	Yes	Yes
Benchmark	No	No	No	No
First stage F		368.0		137.8
Observations	27191	17526	26380	16813

Notes: The outcome variable in each each specification is the log Out-of-Pocket Maximum. Each column reports a plan-year regression. Panel A reports the benchmark-controlled Provider Channel specification, while Panel B reports the benchmark-uncontrolled Total Effect specification. In both panels, columns (1)–(2) use additive core-based statistical area (CBSA) and year fixed effects and report OLS and IV estimates, respectively. Columns (3)–(4) replace these with CBSA × year fixed effects and again report OLS and IV estimates. In Panel A, the benchmark is controlled for using $\ln B_{pt}$; in Panel B, the benchmark is not included. IV columns instrument $\ln \pi_{pt}^{ACO}$ with the Bartik shift-share instrument. All specifications include plan fixed effects and a rich set of plan-service-area controls, including beneficiary demographics and health status, healthcare supply and FFS utilization, and hospital market concentration. Regressions are weighted by log plan enrollment. Standard errors clustered at the plan level are reported in parentheses. Reported first-stage F statistics are the Kleibergen–Paap rk Wald F statistic. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 10: Provider Channel: Positive premium

	Regional + Year FE		Regional × Year FE	
	OLS	IV	OLS	IV
Panel A: Provider Channel				
ln π^{ACO}	-0.117***	-0.550***	0.0307	-0.224
	(0.0343)	(0.158)	(0.0777)	(0.386)
Plan FE	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	No	No
Year FE	Yes	Yes	No	No
Regional × Year FE	No	No	Yes	Yes
Benchmark	Yes	Yes	Yes	Yes
First stage F		317.4		114.8
Observations	24904	15534	24153	14882
Panel B: Total Effect				
ln π^{ACO}	-0.118***	-0.563***	0.0290	-0.213
	(0.0344)	(0.158)	(0.0779)	(0.387)
Plan FE	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	No	No
Year FE	Yes	Yes	No	No
Regional × Year FE	No	No	Yes	Yes
Benchmark	No	No	No	No
First stage F		323.5		112.7
Observations	24904	15534	24153	14882

Notes: The outcome variable in each each specification is $\mathbb{1}$(premium >)0. Panel A reports the benchmark-controlled Provider Channel specification, while Panel B reports the benchmark-uncontrolled Total Effect specification. In both panels, columns (1)–(2) use additive core-based statistical area (CBSA) and year fixed effects and report OLS and IV estimates, respectively. Columns (3)–(4) replace these with CBSA × year fixed effects and again report OLS and IV estimates. In Panel A, the benchmark is controlled for using $\ln B_{pt}$; in Panel B, the benchmark is not included. IV columns instrument $\ln \pi_{pt}^{ACO}$ with the Bartik shift-share instrument. All specifications include plan fixed effects and a rich set of plan-service-area controls, including beneficiary demographics and health status, healthcare supply and FFS utilization, and hospital market concentration. Regressions are weighted by log plan enrollment. Standard errors clustered at the plan level are reported in parentheses. Reported first-stage F statistics are the Kleibergen–Paap rk Wald F statistic. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 11: Provider Channel: Deductibles

	Regional + Year FE		Regional × Year FE	
	OLS	IV	OLS	IV
Panel A: Provider Channel				
ln π^{ACO}	-0.260	0.617	-2.529**	3.426
	(0.217)	(1.418)	(1.212)	(3.166)
Plan FE	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	No	No
Year FE	Yes	Yes	No	No
Regional × Year FE	No	No	Yes	Yes
Benchmark	Yes	Yes	Yes	Yes
First stage F		362.4		140.6
Observations	27191	17526	26380	16813
Panel B: Total Effect				
ln π^{ACO}	-0.261	0.575	-2.518**	3.415
	(0.217)	(1.395)	(1.210)	(3.200)
Plan FE	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	No	No
Year FE	Yes	Yes	No	No
Regional × Year FE	No	No	Yes	Yes
Benchmark	No	No	No	No
First stage F		368.0		137.8
Observations	27191	17526	26380	16813

Notes: The outcome variable in each each specification is the plan Deductibles in level. Panel A reports the benchmark-controlled Provider Channel specification, while Panel B reports the benchmark-uncontrolled Total Effect specification. In both panels, columns (1)–(2) use additive core-based statistical area (CBSA) and year fixed effects and report OLS and IV estimates, respectively. Columns (3)–(4) replace these with CBSA × year fixed effects and again report OLS and IV estimates. In Panel A, the benchmark is controlled for using $\ln B_{pt}$; in Panel B, the benchmark is not included. IV columns instrument $\ln \pi_{pt}^{ACO}$ with the Bartik shift-share instrument. All specifications include plan fixed effects and a rich set of plan-service-area controls, including beneficiary demographics and health status, healthcare supply and FFS utilization, and hospital market concentration. Regressions are weighted by log plan enrollment. Standard errors clustered at the plan level are reported in parentheses. Reported first-stage F statistics are the Kleibergen–Paap rk Wald F statistic. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 12: Provider Channel: PCP copays

	Regional + Year FE		Regional × Year FE	
	OLS	IV	OLS	IV
Panel A: Provider Channel				
ln π^{ACO}	-0.0624**	-0.00158	-0.0559	0.134
	(0.0311)	(0.146)	(0.0681)	(0.351)
Plan FE	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	No	No
Year FE	Yes	Yes	No	No
Regional × Year FE	No	No	Yes	Yes
Benchmark	Yes	Yes	Yes	Yes
First stage F		362.4		140.6
Observations	27191	17526	26380	16813
Panel B: Total Effect				
ln π^{ACO}	-0.0630**	-0.0341	-0.0587	0.143
	(0.0311)	(0.146)	(0.0682)	(0.353)
Plan FE	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	No	No
Year FE	Yes	Yes	No	No
Regional × Year FE	No	No	Yes	Yes
Benchmark	No	No	No	No
First stage F		368.0		137.8
Observations	27191	17526	26380	16813

Notes: The outcome variable in each each specification is $\mathbb{1}(\text{PCP copay} > 0)$. Panel A reports the benchmark-controlled Provider Channel specification, while Panel B reports the benchmark-uncontrolled Total Effect specification. In both panels, columns (1)–(2) use additive core-based statistical area (CBSA) and year fixed effects and report OLS and IV estimates, respectively. Columns (3)–(4) replace these with CBSA × year fixed effects and again report OLS and IV estimates. In Panel A, the benchmark is controlled for using $\ln B_{pt}$; in Panel B, the benchmark is not included. IV columns instrument $\ln \pi^{ACO}_{pt}$ with the Bartik shift-share instrument. All specifications include plan fixed effects and a rich set of plan-service-area controls, including beneficiary demographics and health status, healthcare supply and FFS utilization, and hospital market concentration. Regressions are weighted by log plan enrollment. Standard errors clustered at the plan level are reported in parentheses. Reported first-stage F statistics are the Kleibergen–Paap rk Wald F statistic. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 13: Provider Channel: Specialist copays

	Regional + Year FE		Regional × Year FE	
	OLS	IV	OLS	IV
Panel A: Provider Channel				
ln π^{ACO}	0.0292	0.0297	0.105*	0.483*
	(0.0239)	(0.104)	(0.0545)	(0.289)
Plan FE	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	No	No
Year FE	Yes	Yes	No	No
Regional × Year FE	No	No	Yes	Yes
Benchmark	Yes	Yes	Yes	Yes
First stage F		362.4		140.6
Observations	27191	17526	26380	16813
Panel B: Total Effect				
ln π^{ACO}	0.0289	0.0127	0.104*	0.488*
	(0.0239)	(0.104)	(0.0544)	(0.292)
Plan FE	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	No	No
Year FE	Yes	Yes	No	No
Regional × Year FE	No	No	Yes	Yes
Benchmark	No	No	No	No
First stage F		368.0		137.8
Observations	27191	17526	26380	16813

Notes: The outcome variable in each each specification is $\mathbb{1}$(Specialist copay > 0). Panel A reports the benchmark-controlled Provider Channel specification, while Panel B reports the benchmark uncontrolled Total Effect specification. In both panels, columns (1)–(2) use additive core-based statistical area (CBSA) and year fixed effects and report OLS and IV estimates, respectively. Columns (3)–(4) replace these with CBSA × year fixed effects and again report OLS and IV estimates. In Panel A, the benchmark is controlled for using $\ln B_{pt}$; in Panel B, the benchmark is not included. IV columns instrument $\ln \pi_{pt}^{ACO}$ with the Bartik shift-share instrument. All specifications include plan fixed effects and a rich set of plan-service-area controls, including beneficiary demographics and health status, healthcare supply and FFS utilization, and hospital market concentration. Regressions are weighted by log plan enrollment. Standard errors clustered at the plan level are reported in parentheses. Reported first-stage F statistics are the Kleibergen–Paap rk Wald F statistic. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 14: Dual-role heterogeneity: log bids

	Sample: Non-Dual Role		Sample: Dual Role	
	OLS	IV	OLS	IV
Panel A: Provider Channel				
ln π^{ACO}	-0.0169	-0.114	0.0389	-0.0538
	(0.0169)	(0.0742)	(0.0388)	(0.195)
Plan FE	Yes	Yes	Yes	Yes
CBSA × Year FE	Yes	Yes	Yes	Yes
Benchmark	Yes	Yes	Yes	Yes
First stage F		139.8		4.238
Observations	22452	13947	3573	2597
Panel B: Total Effect				
ln π^{ACO}	-0.00957	-0.136*	0.0412	-0.0502
	(0.0172)	(0.0754)	(0.0382)	(0.202)
Plan FE	Yes	Yes	Yes	Yes
CBSA × Year FE	Yes	Yes	Yes	Yes
Benchmark	No	No	No	No
First stage F		137.4		4.227
Observations	22452	13947	3573	2597

Notes: The outcome variable in each each specification is the log Bids. Each column reports a plan-year regression estimated separately by organizational form. Columns (1)–(2) use the sample of non-dual-role plans, while columns (3)–(4) use the sample of dual-role plans. Panel A reports the benchmark-controlled Provider Channel specification, while Panel B reports the benchmark-uncontrolled Total Effect specification. In both panels, odd-numbered columns report OLS estimates and even-numbered columns report IV estimates. All specifications include plan fixed effects and CBSA × year fixed effects. In Panel A, the benchmark is controlled for using $\ln B_{pt}$; in Panel B, the benchmark is not included. IV columns instrument $\ln \pi_{pt}^{ACO}$ with the Bartik shift-share instrument. All specifications include a rich set of plan-service-area controls, including beneficiary demographics and health status, healthcare supply and FFS utilization, and hospital market concentration. Regressions are weighted by log plan enrollment. Standard errors clustered at the plan level are reported in parentheses. Reported first-stage F statistics are the Kleibergen–Paap rk Wald F statistic. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 15: Dual-role heterogeneity: log projected costs

	Sample: Non-Dual Role		Sample: Dual Role	
	OLS	IV	OLS	IV
Panel A: Provider Channel				
ln π^{ACO}	-0.00170	-0.198	0.137	0.574
	(0.0482)	(0.159)	(0.125)	(0.695)
Plan FE	Yes	Yes	Yes	Yes
CBSA × Year FE	Yes	Yes	Yes	Yes
Benchmark	Yes	Yes	Yes	Yes
First stage F		150.6		17.90
Observations	11784	8360	1576	1267
Panel B: Total Effect				
ln π^{ACO}	0.0105	-0.200	0.137	0.572
	(0.0483)	(0.159)	(0.124)	(0.701)
Plan FE	Yes	Yes	Yes	Yes
CBSA × Year FE	Yes	Yes	Yes	Yes
Benchmark	No	No	No	No
First stage F		150.7		17.30
Observations	11784	8360	1576	1267

Notes: The outcome variable in each each specification is log Projected Costs. Each column reports a plan-year regression estimated separately by organizational form. Columns (1)–(2) use the sample of non-dual-role plans, while columns (3)–(4) use the sample of dual-role plans. Panel A reports the benchmark-controlled Provider Channel specification, while Panel B reports the benchmark-uncontrolled Total Effect specification. In both panels, odd-numbered columns report OLS estimates and even-numbered columns report IV estimates. All specifications include plan fixed effects and CBSA × year fixed effects. In Panel A, the benchmark is controlled for using $\ln B_{pt}$; in Panel B, the benchmark is not included. IV columns instrument $\ln \pi_{pt}^{ACO}$ with the Bartik shift-share instrument. All specifications include a rich set of plan-service-area controls, including beneficiary demographics and health status, healthcare supply and FFS utilization, and hospital market concentration. Regressions are weighted by log plan enrollment. Standard errors clustered at the plan level are reported in parentheses. Reported first-stage F statistics are the Kleibergen–Paap rk Wald F statistic. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 16: Dual-role heterogeneity: log projected costs excluding 2020

	Sample: Non-Dual Role		Sample: Dual Role	
	OLS	IV	OLS	IV
Panel A: Provider Channel				
ln π^{ACO}	-0.0160	-0.331**	0.120	0.206
	(0.0592)	(0.165)	(0.151)	(0.657)
Plan FE	Yes	Yes	Yes	Yes
CBSA × Year FE	Yes	Yes	Yes	Yes
Benchmark	Yes	Yes	Yes	Yes
First stage F		138.3		18.98
Observations	7199	5769	1178	1062
Panel B: Total Effect				
ln π^{ACO}	-0.00723	-0.333**	0.113	0.206
	(0.0592)	(0.166)	(0.151)	(0.683)
Plan FE	Yes	Yes	Yes	Yes
CBSA × Year FE	Yes	Yes	Yes	Yes
Benchmark	No	No	No	No
First stage F		139.0		18.09
Observations	7199	5769	1178	1062

Notes: The dependent variable in all specifications is log projected costs. Each column reports a plan-year regression estimated on the sample that excludes 2020, the first COVID year, because projected cost measures may be unusually noisy in that year. Each column reports a plan-year regression estimated separately by organizational form. Columns (1)–(2) use the sample of non-dual-role plans, while columns (3)–(4) use the sample of dual-role plans. Panel A reports the benchmark-controlled Provider Channel specification, while Panel B reports the benchmark-uncontrolled Total Effect specification. In both panels, odd-numbered columns report OLS estimates and even-numbered columns report IV estimates. All specifications include plan fixed effects and CBSA × year fixed effects. In Panel A, the benchmark is controlled for using $\ln B_{pt}$; in Panel B, the benchmark is not included. IV columns instrument $\ln \pi_{pt}^{ACO}$ with the Bartik shift-share instrument. All specifications include a rich set of plan-service-area controls, including beneficiary demographics and health status, healthcare supply and FFS utilization, and hospital market concentration. Regressions are weighted by log plan enrollment. Standard errors clustered at the plan level are reported in parentheses. Reported first-stage F statistics are the Kleibergen–Paap rk Wald F statistic. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 17: Dual-role heterogeneity: log rebates

	Sample: Non-Dual Role		Sample: Dual Role	
	OLS	IV	OLS	IV
Panel A: Provider Channel				
ln π^{ACO}	0.0640	2.233***	0.319	8.338
	(0.171)	(0.743)	(0.374)	(8.024)
Plan FE	Yes	Yes	Yes	Yes
CBSA × Year FE	Yes	Yes	Yes	Yes
Benchmark	Yes	Yes	Yes	Yes
First stage F		107.8		1.288
Observations	20382	12189	2904	1983
Panel B: Total Effect				
ln π^{ACO}	0.0768	2.163***	0.302	7.884
	(0.171)	(0.744)	(0.376)	(7.558)
Plan FE	Yes	Yes	Yes	Yes
CBSA × Year FE	Yes	Yes	Yes	Yes
Benchmark	No	No	No	No
First stage F		105.7		1.346
Observations	20382	12189	2904	1983

Notes: The outcome variable in each each specification is log Rebates. Each column reports a plan-year regression estimated separately by organizational form. Columns (1)–(2) use the sample of non-dual-role plans, while columns (3)–(4) use the sample of dual-role plans. Panel A reports the benchmark-controlled Provider Channel specification, while Panel B reports the benchmark-uncontrolled Total Effect specification. In both panels, odd-numbered columns report OLS estimates and even-numbered columns report IV estimates. All specifications include plan fixed effects and CBSA × year fixed effects. In Panel A, the benchmark is controlled for using $\ln B_{pt}$; in Panel B, the benchmark is not included. IV columns instrument $\ln \pi_{pt}^{ACO}$ with the Bartik shift-share instrument. All specifications include a rich set of plan-service-area controls, including beneficiary demographics and health status, healthcare supply and FFS utilization, and hospital market concentration. Regressions are weighted by log plan enrollment. Standard errors clustered at the plan level are reported in parentheses. Reported first-stage F statistics are the Kleibergen–Paap rk Wald F statistic. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 18: Dual-role heterogeneity: log OOP Max

	Sample: Non-Dual Role		Sample: Dual Role	
	OLS	IV	OLS	IV
Panel A: Provider Channel				
$\ln \pi^{ACO}$	0.0118	0.186	0.588	-4.151*
	(0.103)	(0.486)	(0.435)	(2.155)
Plan FE	Yes	Yes	Yes	Yes
CBSA × Year FE	Yes	Yes	Yes	Yes
Benchmark	Yes	Yes	Yes	Yes
First stage F		139.8		4.238
Observations	22452	13947	3573	2597
Panel B: Total Effect				
$\ln \pi^{ACO}$	0.0188	0.159	0.591	-4.143*
	(0.104)	(0.493)	(0.437)	(2.168)
Plan FE	Yes	Yes	Yes	Yes
CBSA × Year FE	Yes	Yes	Yes	Yes
Benchmark	No	No	No	No
First stage F		137.4		4.227
Observations	22452	13947	3573	2597

Notes: The outcome variable in each each specification is log Out-of-Pocket Max. Each column reports a plan-year regression estimated separately by organizational form. Columns (1)–(2) use the sample of non-dual-role plans, while columns (3)–(4) use the sample of dual-role plans. Panel A reports the benchmark-controlled Provider Channel specification, while Panel B reports the benchmark-uncontrolled Total Effect specification. In both panels, odd-numbered columns report OLS estimates and even-numbered columns report IV estimates. All specifications include plan fixed effects and CBSA × year fixed effects. In Panel A, the benchmark is controlled for using $\ln B_{pt}$; in Panel B, the benchmark is not included. IV columns instrument $\ln \pi^{ACO}_{pt}$ with the Bartik shift-share instrument. All specifications include a rich set of plan-service-area controls, including beneficiary demographics and health status, healthcare supply and FFS utilization, and hospital market concentration. Regressions are weighted by log plan enrollment. Standard errors clustered at the plan level are reported in parentheses. Reported first-stage F statistics are the Kleibergen–Paap rk Wald F statistic. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 19: Dual-role heterogeneity: Positive premium

	Sample: Non-Dual Role		Sample: Dual Role	
	OLS	IV	OLS	IV
Panel A: Provider Channel				
$\ln \pi^{ACO}$	0.0587	-0.195	-0.0358	-1.494
	(0.0834)	(0.378)	(0.200)	(1.463)
Plan FE	Yes	Yes	Yes	Yes
CBSA × Year FE	Yes	Yes	Yes	Yes
Benchmark	Yes	Yes	Yes	Yes
First stage F		114.7		2.228
Observations	20823	12565	2975	2052
Panel B: Total Effect				
$\ln \pi^{ACO}$	0.0568	-0.181	-0.0425	-1.659
	(0.0836)	(0.379)	(0.199)	(1.447)
Plan FE	Yes	Yes	Yes	Yes
CBSA × Year FE	Yes	Yes	Yes	Yes
Benchmark	No	No	No	No
First stage F		113.0		2.323
Observations	20823	12565	2975	2052

Notes: The outcome variable in each each specification is $\mathbb{1}(\text{premium} > 0)$. Each column reports a plan-year regression estimated separately by organizational form. Columns (1)–(2) use the sample of non-dual-role plans, while columns (3)–(4) use the sample of dual-role plans. Panel A reports the benchmark-controlled Provider Channel specification, while Panel B reports the benchmark-uncontrolled Total Effect specification. In both panels, odd-numbered columns report OLS estimates and even-numbered columns report IV estimates. All specifications include plan fixed effects and CBSA × year fixed effects. In Panel A, the benchmark is controlled for using $\ln B_{pt}$; in Panel B, the benchmark is not included. IV columns instrument $\ln \pi_{pt}^{ACO}$ with the Bartik shift-share instrument. All specifications include a rich set of plan-service-area controls, including beneficiary demographics and health status, healthcare supply and FFS utilization, and hospital market concentration. Regressions are weighted by log plan enrollment. Standard errors clustered at the plan level are reported in parentheses. Reported first-stage F statistics are the Kleibergen–Paap rk Wald F statistic. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 20: Dual-role heterogeneity: Deductible

	Sample: Non-Dual Role		Sample: Dual Role	
	OLS	IV	OLS	IV
Panel A: Provider Channel				
$\ln \pi^{ACO}$	-1.784*	2.037	-4.885	4.232
	(0.973)	(1.985)	(5.023)	(7.402)
Plan FE	Yes	Yes	Yes	Yes
CBSA × Year FE	Yes	Yes	Yes	Yes
Benchmark	Yes	Yes	Yes	Yes
First stage F		139.8		4.238
Observations	22452	13947	3573	2597
Panel B: Total Effect				
$\ln \pi^{ACO}$	-1.771*	2.030	-4.942	4.144
	(0.966)	(2.027)	(5.094)	(7.041)
Plan FE	Yes	Yes	Yes	Yes
CBSA × Year FE	Yes	Yes	Yes	Yes
Benchmark	No	No	No	No
First stage F		137.4		4.227
Observations	22452	13947	3573	2597

Notes: The outcome variable in each each specification is the plan Deductibles in level. Each column reports a plan-year regression estimated separately by organizational form. Columns (1)–(2) use the sample of non-dual-role plans, while columns (3)–(4) use the sample of dual-role plans. Panel A reports the benchmark-controlled Provider Channel specification, while Panel B reports the benchmark-uncontrolled Total Effect specification. In both panels, odd-numbered columns report OLS estimates and even-numbered columns report IV estimates. All specifications include plan fixed effects and CBSA × year fixed effects. In Panel A, the benchmark is controlled for using $\ln B_{pt}$; in Panel B, the benchmark is not included. IV columns instrument $\ln \pi_{pt}^{ACO}$ with the Bartik shift-share instrument. All specifications include a rich set of plan-service-area controls, including beneficiary demographics and health status, healthcare supply and FFS utilization, and hospital market concentration. Regressions are weighted by log plan enrollment. Standard errors clustered at the plan level are reported in parentheses. Reported first-stage F statistics are the Kleibergen–Paap rk Wald F statistic. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 21: Dual-role heterogeneity: PCP copay

	Sample: Non-Dual Role		Sample: Dual Role	
	OLS	IV	OLS	IV
Panel A: Provider Channel				
$\ln \pi^{ACO}$	-0.128**	-0.157	0.0872	-4.160
	(0.0610)	(0.253)	(0.278)	(2.661)
Plan FE	Yes	Yes	Yes	Yes
CBSA × Year FE	Yes	Yes	Yes	Yes
Benchmark	Yes	Yes	Yes	Yes
First stage F		139.8		4.238
Observations	22452	13947	3573	2597
Panel B: Total Effect				
$\ln \pi^{ACO}$	-0.129**	-0.150	0.0830	-4.165
	(0.0610)	(0.254)	(0.279)	(2.656)
Plan FE	Yes	Yes	Yes	Yes
CBSA × Year FE	Yes	Yes	Yes	Yes
Benchmark	No	No	No	No
First stage F		137.4		4.227
Observations	22452	13947	3573	2597

Notes: The outcome variable in each each specification is $\mathbb{1}(\text{PCP copay} > 0)$. Each column reports a plan-year regression estimated separately by organizational form. Columns (1)–(2) use the sample of non-dual-role plans, while columns (3)–(4) use the sample of dual-role plans. Panel A reports the benchmark-controlled Provider Channel specification, while Panel B reports the benchmark-uncontrolled Total Effect specification. In both panels, odd-numbered columns report OLS estimates and even-numbered columns report IV estimates. All specifications include plan fixed effects and CBSA × year fixed effects. In Panel A, the benchmark is controlled for using $\ln B_{pt}$; in Panel B, the benchmark is not included. IV columns instrument $\ln \pi_{pt}^{ACO}$ with the Bartik shift-share instrument. All specifications include a rich set of plan-service-area controls, including beneficiary demographics and health status, healthcare supply and FFS utilization, and hospital market concentration. Regressions are weighted by log plan enrollment. Standard errors clustered at the plan level are reported in parentheses. Reported first-stage F statistics are the Kleibergen–Paap rk Wald F statistic. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 22: Dual-role heterogeneity: Specialist copay

	Sample: Non-Dual Role		Sample: Dual Role	
	OLS	IV	OLS	IV
Panel A: Provider Channel				
$\ln \pi^{ACO}$	0.115*	0.398	0.0187	-1.194
	(0.0656)	(0.341)	(0.105)	(0.855)
Plan FE	Yes	Yes	Yes	Yes
CBSA × Year FE	Yes	Yes	Yes	Yes
Benchmark	Yes	Yes	Yes	Yes
First stage F		139.8		4.238
Observations	22452	13947	3573	2597
Panel B: Total Effect				
$\ln \pi^{ACO}$	0.115*	0.404	0.0199	-1.191
	(0.0654)	(0.343)	(0.105)	(0.859)
Plan FE	Yes	Yes	Yes	Yes
CBSA × Year FE	Yes	Yes	Yes	Yes
Benchmark	No	No	No	No
First stage F		137.4		4.227
Observations	22452	13947	3573	2597

Notes: The outcome variable in each each specification is $\mathbb{1}(\text{Specialist copay} > 0)$. Each column reports a plan-year regression estimated separately by organizational form. Columns (1)–(2) use the sample of non-dual-role plans, while columns (3)–(4) use the sample of dual-role plans. Panel A reports the benchmark-controlled Provider Channel specification, while Panel B reports the benchmark-uncontrolled Total Effect specification. In both panels, odd-numbered columns report OLS estimates and even-numbered columns report IV estimates. All specifications include plan fixed effects and CBSA × year fixed effects. In Panel A, the benchmark is controlled for using $\ln B_{pt}$; in Panel B, the benchmark is not included. IV columns instrument $\ln \pi_{pt}^{ACO}$ with the Bartik shift-share instrument. All specifications include a rich set of plan-service-area controls, including beneficiary demographics and health status, healthcare supply and FFS utilization, and hospital market concentration. Regressions are weighted by log plan enrollment. Standard errors clustered at the plan level are reported in parentheses. Reported first-stage F statistics are the Kleibergen–Paap rk Wald F statistic. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Appendix C

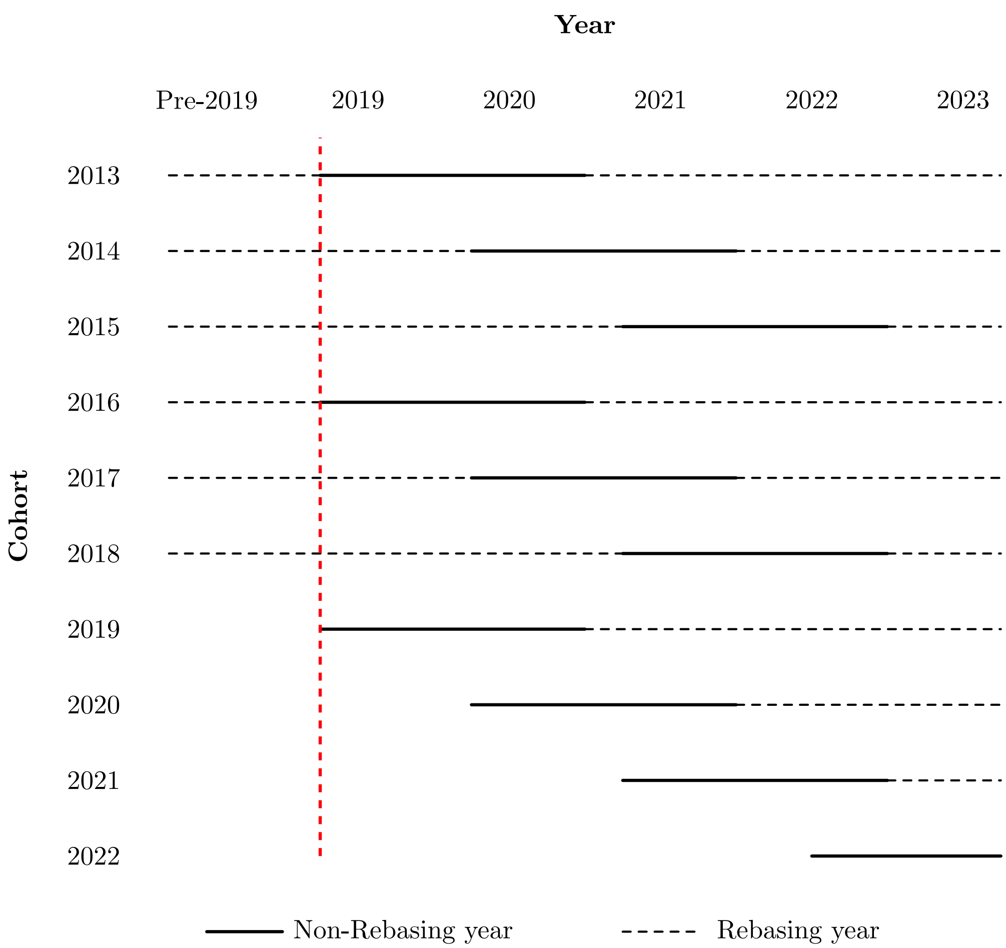

Figure 5: This figure shows when the non-rebasing years occur for each entry cohort.

Appendix D

Instrument Validation and Robustness

We implement several diagnostic exercises to probe the validity of the identification strategy. Throughout, the excluded instrument is the Pathways-based shift-share variable Z_{pt} defined in (10). The most direct test of the identifying assumptions is whether the instrument predicts changes in MA outcomes *before* the Pathways reform took effect. If the instrument is valid, there should be no systematic relationship between baseline Pathways exposure and pre-reform plan outcomes, since non-rebasing year incentives did not exist before 2019.

To implement this test, we summarize each plan's exposure with a time-invariant "dose" constructed as the post-2018 average of the time-varying instrument:

$$D_p \equiv \frac{1}{T_{\text{post}}} \sum_{t \in \text{post}} Z_{pt}. \tag{14}$$

We then estimate a dose-response event study:

$$Y_{pt} = \alpha_p + \lambda_t + \sum_{k \neq -1} \beta_k \Big(\mathbf{1}\{t - 2019 = k\} \times D_p \Big) + X'_{pt}\gamma + \varepsilon_{pt}, \tag{15}$$

where $k = -1$ (i.e., 2018) is the omitted baseline year. The identifying prediction is that *lead* coefficients are approximately zero: $\beta_k \approx 0$ for $k \leq -2$.

Intuition. To see concretely why this test is informative, consider two MA plans, Plan A and Plan B, that are identical in 2018: they serve similar counties, face the same benchmarks, and offer comparable benefits. The only difference is that Plan A's service area happens to be concentrated in counties where the dominant ACOs entered MSSP in 2013, while Plan B's counties have ACOs that entered in 2016. Under the Pathways schedule, the 2013-cohort ACOs reach their first non-rebasing years before the 2016-cohort ACOs do, so after 2019 Plan A receives a higher "dose" D_p than Plan B.

The event study asks: were Plans A and B already on different *outcome trajectories before 2019*, when neither cohort faced non-rebasing year incentives? If the pre-2019 coefficients $\beta_{-2}, \beta_{-3}, \ldots$ are indistinguishable from zero, it suggests that the two plans were evolving in parallel before the Pathways schedule activated. Any post-2019 divergence in bids, rebates, or benefits can then be more credibly attributed to the differential non-rebasing year exposure, rather than to pre-existing differences between markets where early versus late ACO cohorts operate.

Event Study Plots

Figure 6 to Figure 13 report the dose-response event-study estimates described above. We plot the estimated coefficients β_k for $k \in \{-4, 3\}$, where 2018 ($k = -1$) is the omitted base year because it immediately precedes the 2019 Pathways reform.

The main lesson from these plots is that the validation exercise is most supportive for projected costs and bids. In both cases, the lead coefficients are generally close to zero and do not display a systematic pre-2019 pattern, which is consistent with the identifying prediction that high-dose and low-dose plans were not already on differential trends before the reform. After 2019, the coefficients move in the expected direction, with higher reform exposure associated with lower projected costs and lower bids. For projected costs, this pattern is especially informative because the main analysis sample is short and overlaps with the COVID period. The results from the event-study plots, therefore, strengthen the interpretation that the negative projected-cost estimates reflect reform-induced changes rather than pre-existing differences across plans.

For rebates, the event-study evidence is less reassuring. The pre-2019 coefficients are not as close to zero and suggest residual differential trends across high-dose and low-dose plans even before the Pathways schedule takes effect. Although the post-2019 rebate coefficients are directionally consistent with greater pass-through to beneficiaries, the presence of non-flat pre-trends means that we do not interpret the rebate IV estimates as clean causal evidence in the same way as the projected-cost and bid estimates.

For the remaining outcomes—out-of-pocket maxima, positive-premium status, deductibles, and physician copays—the event-study coefficients are generally noisy and do not reveal a clear post-reform pattern. This is consistent with the main regression results, where the corresponding estimates are typically imprecise. Overall, the event-study plots reinforce the view that the most credible evidence in the paper concerns projected costs and bids, while the evidence for rebates is more tentative and the evidence for the other benefit-design margins is limited.

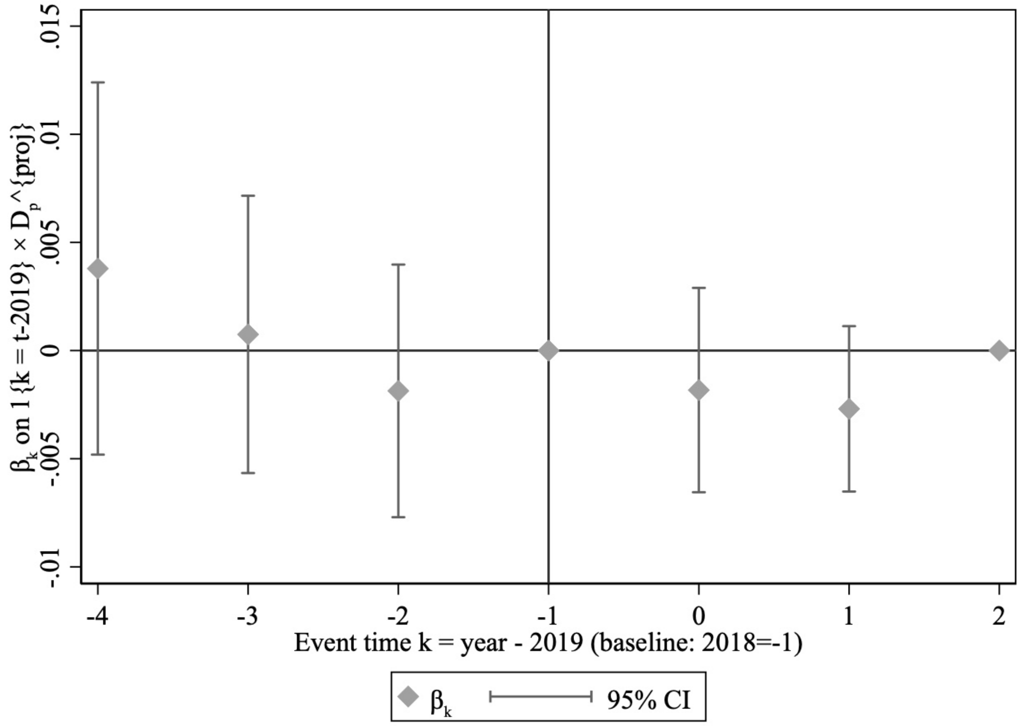

Figure 6: Event-study: log projected cost.

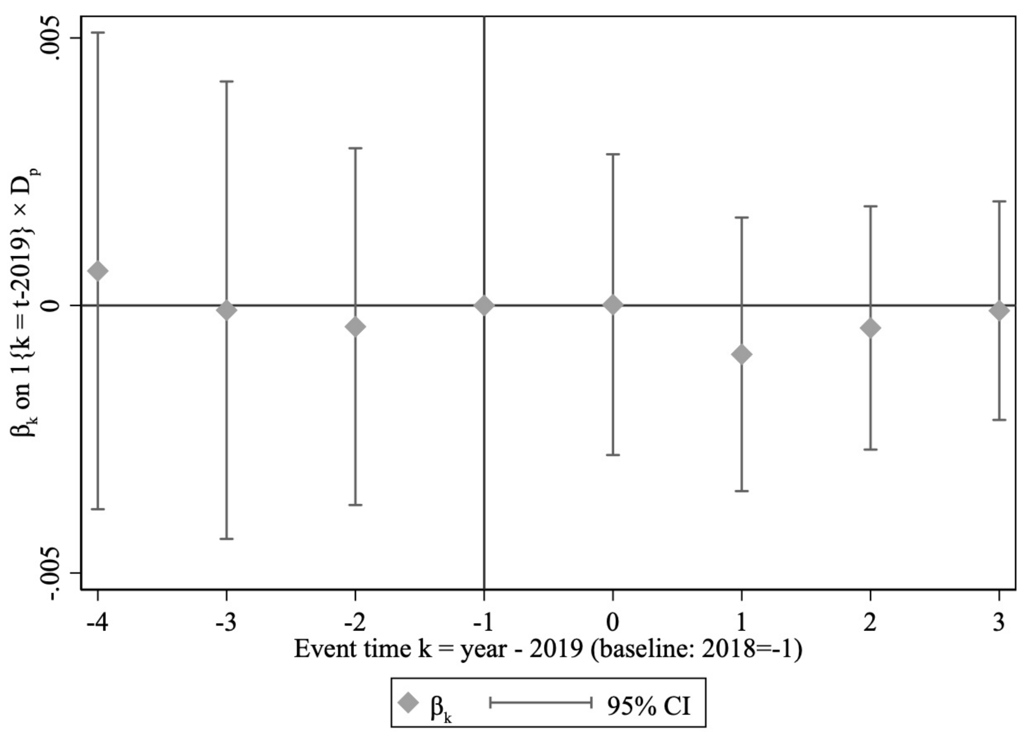

Figure 7: Event-study: log average bid.

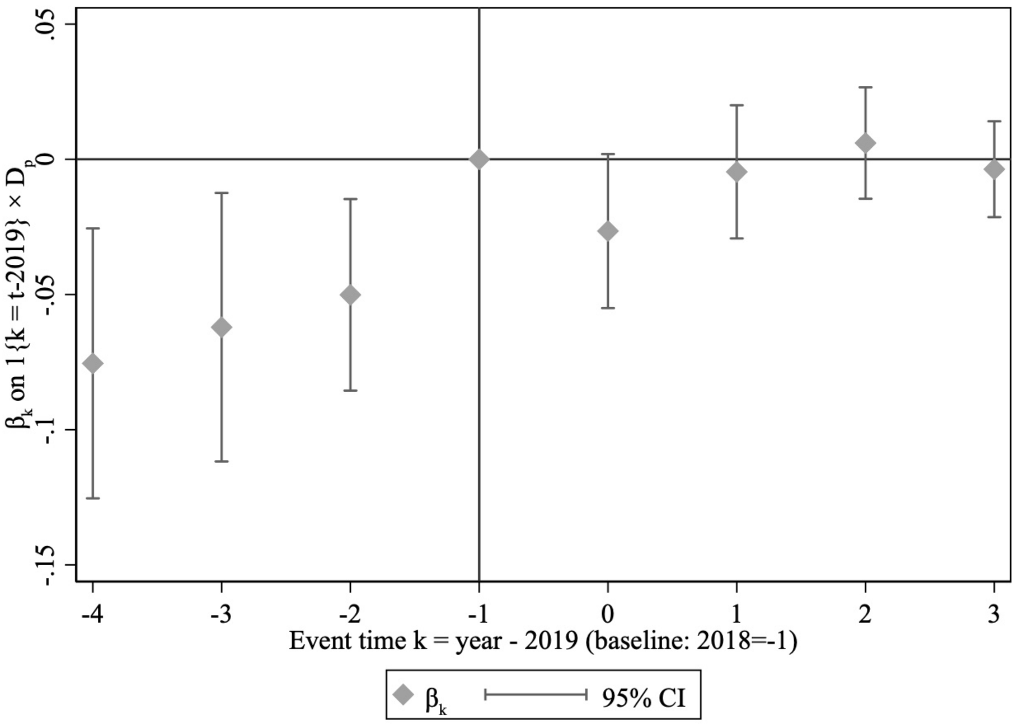

Figure 8: Event-study: log average rebate.

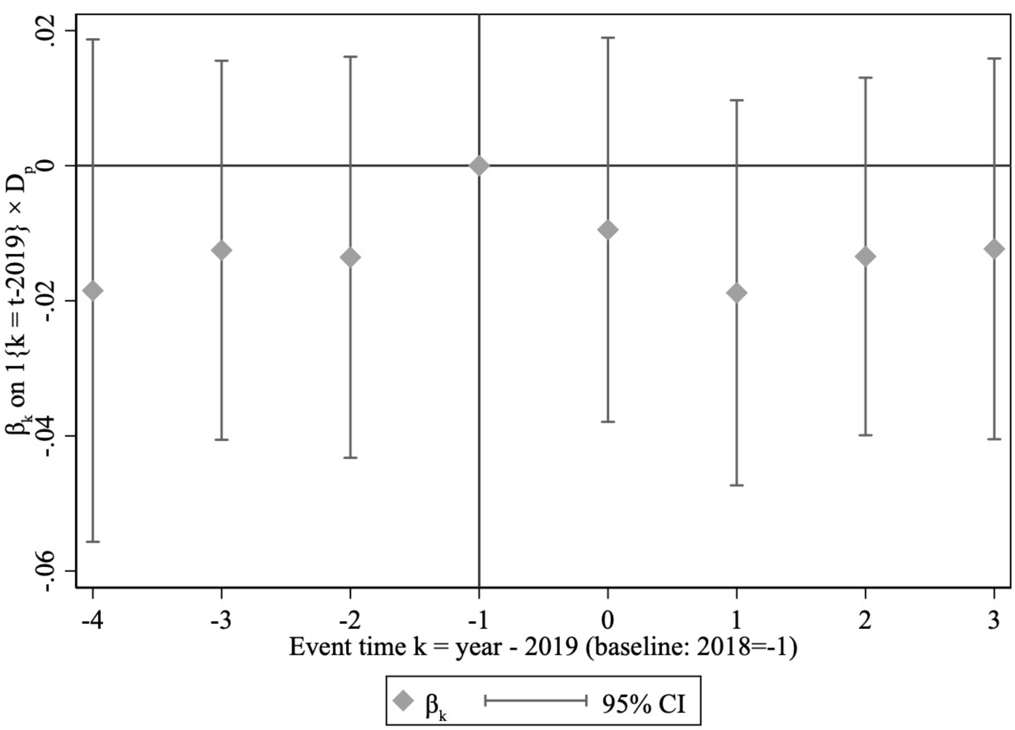

Figure 9: Event-study: log out-of-pocket maximum and exit.

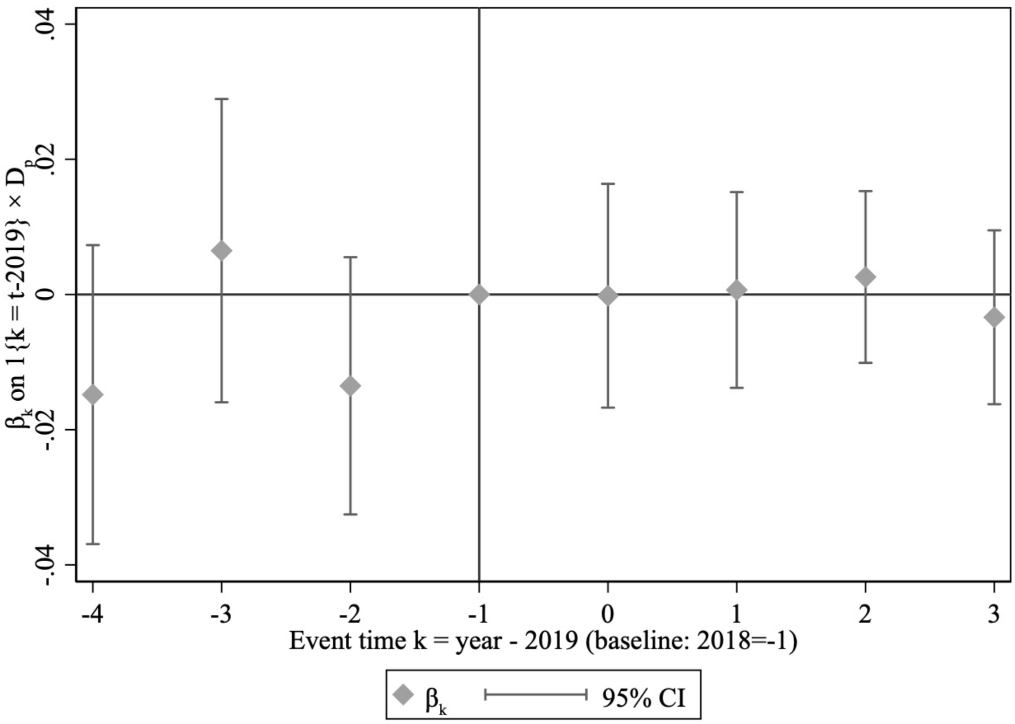

Figure 10: Event-study: Positive Premium indicator

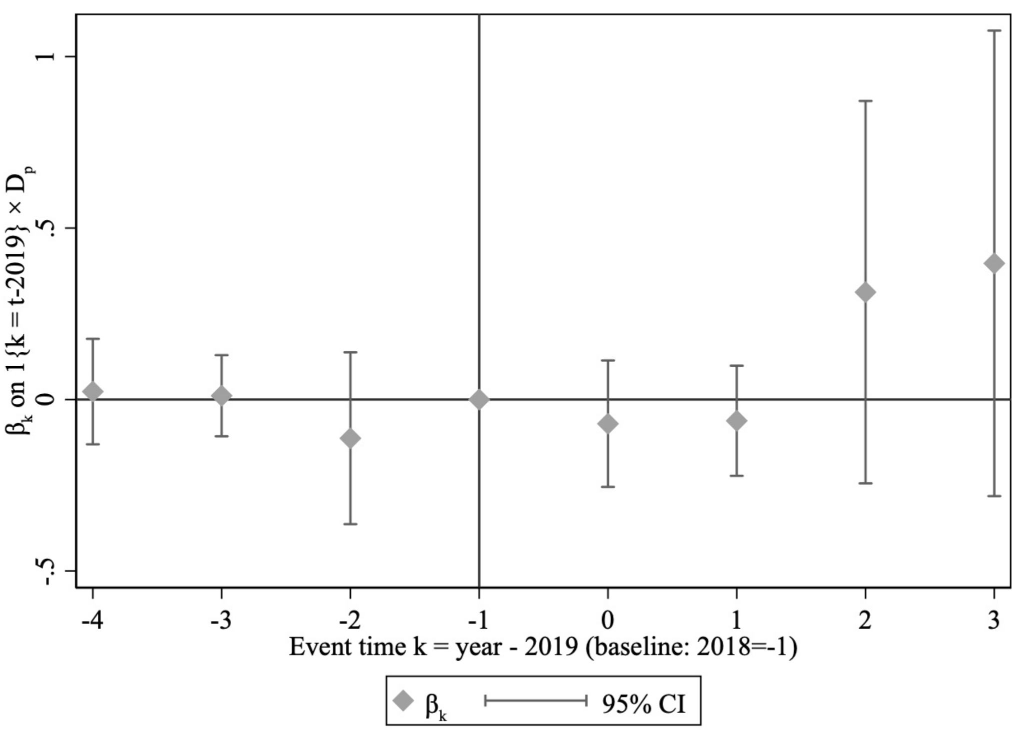

Figure 11: Event-study: Deductibles.

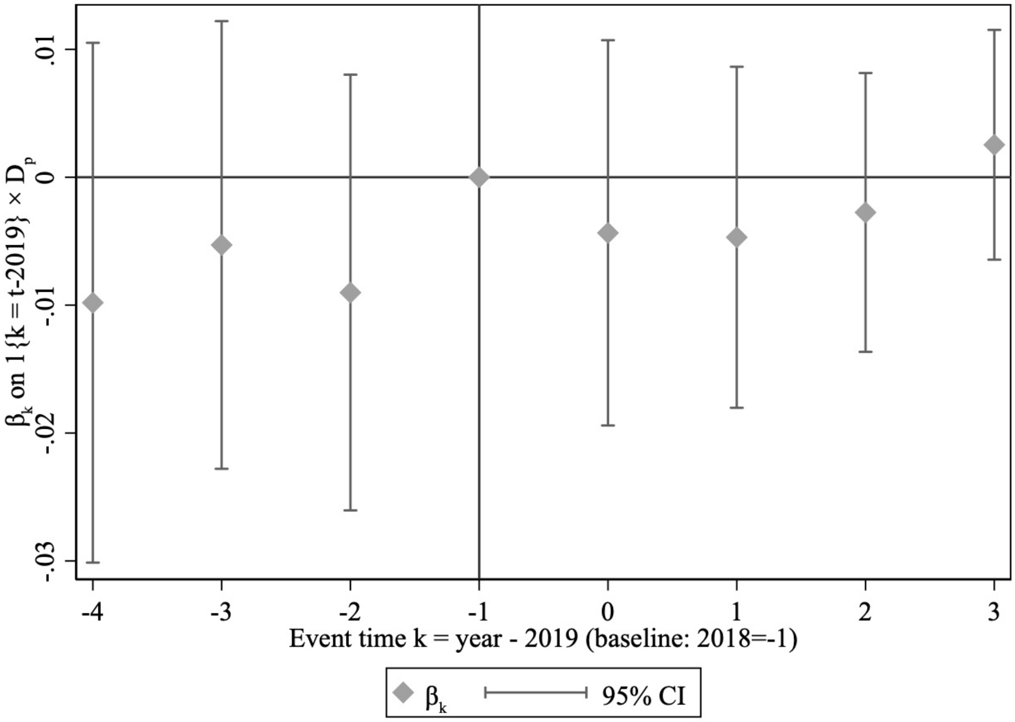

Figure 12: Event-study: PCP copay indicator.

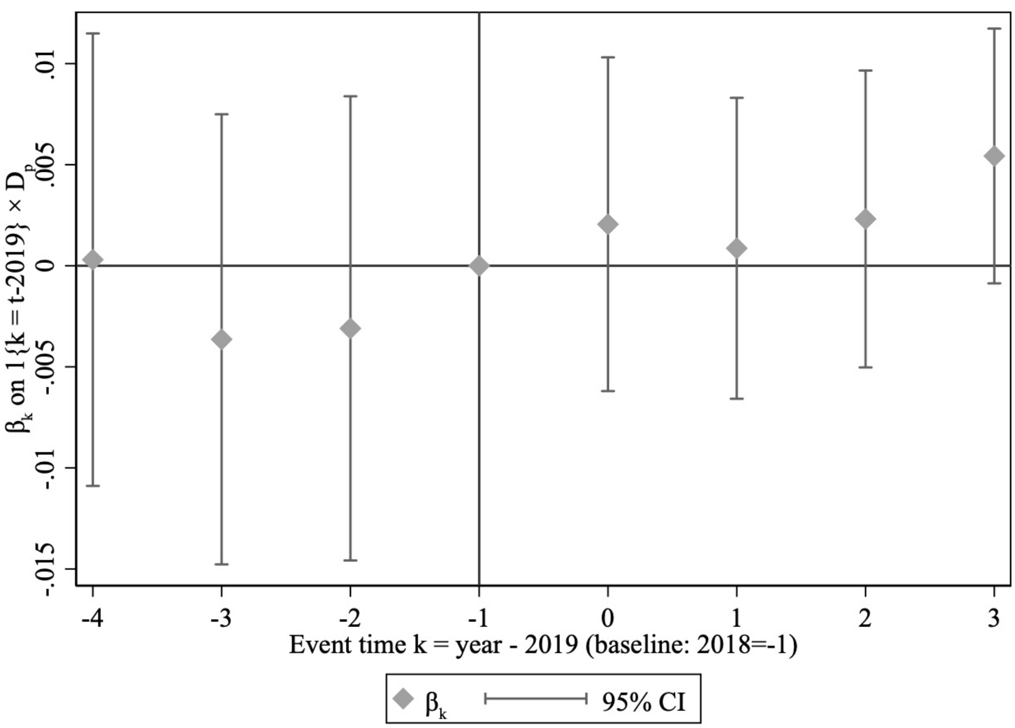

Figure 13: Event-study: Specialist copay indicator indicator.

Benchmark Robustness

Even if the instrument has some effect on benchmarks, we can assess whether the IV estimate of the provider channel is sensitive to how flexibly we control for B_{pt}. In our benchmark-robustness specification, we replace $\ln B_{pt}$ with a full set of benchmark-decile $\times$ year fixed effects. This is a highly flexible control for the payment environment: it effectively compares plans within narrow benchmark bins in each year and eliminates concerns that the main results are driven by functional-form misspecification in the benchmark control. The resulting estimates are very similar to those obtained in the baseline specification with a linear $\ln B_{pt}$ control, indicating that our conclusions are not sensitive to how the benchmark is modeled. The results from this specification are reported in the Online Appendix.