9-2024

Strategic cross-subsidization in healthcare capitation programs: Evidence from Medicare Advantage

Zhaowei SHE
Singapore Management University, zwshe@smu.edu.sg

Turgay AYER

Bilal GOKPINAR

Danny R. HUGHESD

Strategic Cross Subsidization in Healthcare Capitation Programs: Evidence from Medicare Advantage

Zhaowei She
Lee Kong Chian School of Business, Singapore Management University, Singapore

Turgay Ayer
H. Milton Stewart School of Industrial & Systems Engineering, Georgia Institute of Technology, Atlanta, GA

Bilal Gokpinar
UCL School of Management, University College London, London, United Kingdom

Danny Hughes
School of Economics, Georgia Institute of Technology, Atlanta, GA

Problem definition: This study identifies a resource misallocation problem in Medicare Advantage (MA), the U.S.'s largest healthcare capitation program, which may result in discrepancies between patients' health status and the healthcare resources allocated to them.

Methodology/results: Utilizing a large commercial insurance database with claims from over 2 million MA enrollees, this research investigates the allocation of MA capitation payments. The findings reveal that MA inadvertently encourages health plans to reallocate portions of the capitation payments intended for one group of patients to another group of patients, a practice known as strategic cross subsidization. By exploiting an exogenous policy shock on MA capitation payments through a Difference-in-Difference (DID) design, we identify, for the first time in the literature, this cross subsidization practice. Additionally, we show that the strategic cross subsidization practice is associated with the risk selection problem in MA, where low-risk patients are more likely to enroll in MA compared to high-risk patients.

Managerial implications: This research uncovers a previously undocumented healthcare resource misallocation problem, i.e., strategic cross subsidization. U.S. law explicitly prohibits this practice due to its exacerbated effect on the undesired risk selection problem and healthcare inequality in capitation programs. To alleviate the identified strategic cross subsidization behavior, increased transparency in MA claims data and heightened attention from healthcare payers and policymakers are needed.

---

1. Introduction

Capitation has been highlighted as one of the most promising healthcare payment models and is increasingly adopted by healthcare payers around the world (James and Poulsen 2016, Lai 2022). In particular, over the last decade, the Centers for Medicare & Medicaid Services (CMS), the largest healthcare payer in the U.S., has regarded capitation as one of its targeted payment models (Rajkumar et al. 2014, Centers for Medicare & Medicaid Services 2015). CMS oversees the largest healthcare insurance program in the U.S., i.e. Medicare. The current Medicare program consists of two components 1) Traditional Medicare (Part A and Part B) which operates under traditional volume-based payment models and 2) Medicare Advantage (Part C) which is a capitation program. Over the past 10 years, the size of Medicare Advantage (MA) has more than doubled, from 13 million enrollees in 2012 to 28 million in 2022, corresponding to 48% of the total Medicare population (Freed et al. 2022).

CMS is a pure payer in MA and completely outsources the provision and management of healthcare services for its Medicare Advantage enrolled patients to private MA health plans. Specifically, CMS pays MA health plans a risk-adjusted capitation rate for each Medicare beneficiary enrolled, and authorizes the health plans to use the capitation payments to manage the overall health of their enrollees. For example, UnitedHealthcare, as a Medicare Advantage plan provider, oversees the overall provision of healthcare services to Medicare Advantage patients enrolled in its MA health plans and in return receives a fixed risk-adjusted payment per patient from CMS.

The allocation of capitation payments and healthcare resources in MA is determined by MA health plans, instead of CMS. By design, MA health plans have flexibility in operations management decisions, such as which providers to contract with or at what level to offer various direct or non-direct healthcare services. For example, some MA health plans offer free or discounted gym memberships, which are not offered in Traditional Medicare (TM) (Cooper and Trivedi 2012). Furthermore, as MA health plans are not required to submit their claims data to CMS, the operations of these MA health plans remain a black box to the public. This lack of transparency in MA health plans' operations, as argued in a recent paper published by the flagship journal of the American Medical Association, raises questions about "the relative value Medicare Advantage provides to beneficiaries" (Brennan et al. 2018).

The goal of this paper is to open the black box of MA health plans' operations and empirically investigate how MA health plans actually allocate capitation payments among Medicare beneficiaries. Due to the limited accessibility of MA claims data (c.f. Brennan et al. (2018)), little is known about how capitation payments are spent at the patient level in MA. In that regard, some early findings published in leading medical journals motivate this paper. For example, it has been reported that MA health plans tend to offer better coverage in terms of preventive care accessibility (e.g., mammographic screening) and supplementary benefits (e.g., gym memberships) that are mainly appealing to healthier Medicare beneficiaries (Cooper and Trivedi 2012, Hung et al. 2016). In contrast, the quality of nursing home care in MA, a service mostly utilized by sicker patients, is reported to be lower compared with TM (Meyers et al. 2018). Nevertheless, it remains unclear whether these isolated observations have generalizable implications for the whole MA program. In other words, is there a systematic mismatch between a patient's health status and the healthcare resources she gets from MA in general? This is the primary research question we set to explore in this study.

To answer this question, we utilize a large commercial insurance database containing claims from more than 2 million MA enrollees. Specifically, this database enables us to derive two critical variables to study MA's healthcare resource allocation problem. First, we are able to quantify MA health plans' actual healthcare spending on each patient, which indicates how healthcare resources are actually allocated in MA. Second, based on the patient health status information, we are able to calculate the healthcare costs of each patient as estimated by CMS. This corresponds to the expected amount of healthcare spendings CMS would have spent on a patient if this patient was in TM, conditional on this patient's age, gender, and preexisting conditions (Brown et al. 2014). With these two variables, we can quantify the *spending-cost difference* of each MA enrollee in this dataset.[1] This primary outcome variable of our study enables us to investigate whether a systematic discrepancy exists between MA health plans' actual healthcare spending and the estimated healthcare cost for MA patients.

Our first hypothesis examines whether there exists an inverse relationship between the sickness levels of MA patients and the spending-cost differences of such patients. That is, as an MA patient became sicker, it is more likely that they will find their MA health plans insufficient for their needs, as indicated by declining spending-cost differences. This hypothesis is supported by an observation made by a delegate of the American Medical Association (AMA),

> "if your health is good, maybe these plans represent value for some patients, like providing gym memberships ... But that can change in the blink of an eye ... with a stroke or an accident or some acute medical condition and they need a rehabilitation stay. Then, services are restricted so much that they can't recover adequately from the stroke ..." (MedPage Today 2020)

This observation aligns with extensive medical literature indicating that MA health plans tend to overspend on low-risk patients and underspend on high-risk patients(Cooper and Trivedi 2012, Hung et al. 2016, Meyers et al. 2018, Rahman et al. 2015).

Our second hypothesis seeks to identify the potential cause of the disparities in spending on low-risk and high-risk patients by MA health plans. Specifically, we hypothesize that this discrepancy is partly driven by the strategic cross subsidization behavior of MA health plans, where they strategically reallocate a portion of the capitation payment intended for high-risk patients to spend

[1] We note that this spending-cost difference presented in this paper is not a measure of the difference in healthcare benefits between MA and TM. Particularly, the TM cost estimates provided by CMS are only used to benchmark against the actual MA spending.

on low-risk patients. This hypothesis reflects the reality that MA health plans, in order to offer additional benefits such as fitness memberships, often need to curtail some benefits to "free up funds". Moreover, when faced with a trade-off between investing in different benefits, MA health plans tend to opt for "lower-cost benefits that may broadly attract more enrollees" at the expense of "benefits targeted to high-cost patients" who are fewer in number (Skopec et al. 2019b). This cross subsidization incentive of MA health plans is also theoretically validated in recent literature (c.f. She et al. (2022)).

Nevertheless, an alternative hypothesis may also explain MA health plans' overspending and underspending behaviors. Specifically, the MA risk adjustment formula tends to overestimate the healthcare costs of low-risk patients and underestimate those of high-risk patients, leading to overpayment for the former and underpayment for the latter (Glazer and McGuire 2000, Brown et al. 2014). Therefore, the observed spending trends might directly stem from these payment inaccuracies. To address these alternative explanations, we leverage a policy shock from the Affordable Care Act (ACA), which incrementally reduced MA capitation payments over time. Notably, this policy shock imposed deeper capitation payment cuts for high-risk patients compared to their low-risk counterparts, exacerbating the potential underpayments for high-risk patients. Our Difference-in-Difference (DID) design employs this institutional detail to empirically demonstrate that exacerbated underpayment for high-risk patients did not result in corresponding underspending on these patients. Thus, the spending trends observed within MA, namely overspending on low-risk patients and underspending on high-risk patients, cannot be attributed exclusively to overpayment for the former group and underpayment for the latter. As such, this identification strategy, detailed in §6.2.1, allows us to separate our cross subsidization hypothesis from the competing one empirically.

Our third hypothesis delves into the implications of strategic cross subsidization, investigating its possible link to the risk selection problem in MA. While direct patient selection is prohibited in MA, MA health plans may subtly skew the provision of services to attract healthier patients and under-provide other services to deter sicker patients. In fact, despite the rapid increase in overall enrollment in MA, a growing number of sicker Medicare beneficiaries are observed to switch from MA to TM (Li et al. 2018, Rahman et al. 2015, Meyers et al. 2018, 2019, Morrisey et al. 2013, Jacobson et al. 2019b, 2015). This patient self-selection phenomenon, widely referred to as risk selection, is considered a primary issue inhibiting the market's smooth functioning (Geruso and Layton 2017). Existing literature primarily attributes the risk selection problem to payment inaccuracies in MA (Brown et al. 2014, Glazer and McGuire 2000). That is, high-risk patients have a higher likelihood of leaving MA due to persistent underpayment. From this perspective, the reductions in MA capitation payments induced by ACA would exacerbate the risk selection problem in MA, as it further intensifies the underpayment problem for high-risk patients. However, in the presence of strategic cross subsidization, further underpayments for high-risk patients do not exacerbate the underspending on these patients but may mitigate overspending on their low-risk counterparts. Therefore, the cuts in ACA capitation payments primarily diminish the latitude for MA health plans to engage in strategic cross subsidization, thereby potentially alleviating risk selection within MA. To empirically examine these competing propositions, our third hypothesis scrutinizes the impact of the ACA policy shock on risk selection within MA, and finds additional support for our proposed cross subsidization mechanism.

Overall, our paper makes two important contributions to healthcare management and operations literature with significant implications for health policy and practice. First, we provide, to our knowledge, the first empirical evidence of a previously unknown resource misallocation problem in Medicare Advantage, and identify strategic cross subsidization as the underlying mechanism for this problem. While previous operations management literature has examined various resource misallocation problems in healthcare operations, these tend to focus on local settings such as certain emergency departments, nursing homes, or hospitals (Ibanez et al. 2018, Lu and Lu 2017, Kc and Terwiesch 2011). Our work investigates a broad but unrecognized resource misallocation problem in health plans. Second, our paper underscores a contentious practice in MA, thereby drawing attention to the necessity for CMS to champion greater transparency in MA claims data. Cross subsidization is explicitly prohibited in MA.[2] Consequently, if such practice prevails, CMS is legally obligated to confront it. However, owing to the fact that most MA health plans are not mandated to submit their claims data to CMS, the agency often lacks comprehensive insight into MA health plans' operations, which creates a barrier to overseeing this market properly. Therefore, an increase in transparency in MA claims data is imperative to enable CMS to more effectively administer the MA program. Indeed, the Health Subcommittee of the U.S. House Committee on Energy and Commerce has recently proposed legislation with the aim of enhancing transparency within MA, striving to ensure that policymakers have a lucid comprehension of "how these private plans are utilizing taxpayer dollars" (Committee for a Responsible Federal Budget 2022).

The remainder of this paper proceeds as follows. We review the relevant literature in §2 and develop our research hypotheses in §3. In §4, we present the institutional background for Medicare Advantage (MA) with an emphasis on relevant policy changes regarding MA capitation payments in the Affordable Care Act (ACA). In §5, we introduce our data sources and construct the variables used in our econometric models. In §6, we develop the econometric model to test our main

[2] "Payments from any rate cell must not cross-subsidize or be cross-subsidized by payments for any other rate cell" (U.S. Congress 2016).

hypothesis. In §7, we empirically assess the impact of MA health plans' strategic cross subsidization practice on the risk selection problem in MA. We conduct various robustness checks in §8 to examine our main findings. We summarize our findings and conclude in §9.

2. Literature Review

Capitation payment models have long been among healthcare payment reforms' most debated issues. On the one hand, healthcare payers (e.g. CMS) have been among the main advocates for them (James and Poulsen 2016, Rajkumar et al. 2014). It is believed that capitation models would improve the efficiency and value of healthcare delivery "by linking the financial incentives for providers to the total cost and/or quality of care they provide" (Centers for Medicare & Medicaid Services 2015). On the other hand, there has been a discrepancy between the expectations and the actual performance of capitation programs in practice. Specifically, several studies noted that MA, the largest capitation program in the U.S., has lower accessibility to long-term and post-acute care compared to the standard Medicare coverage.[3] While this low service accessibility in MA may not translate into worse healthcare outcomes, e.g. 30-day readmission rate (c.f. Agarwal et al. (2021)), evidence exists that high-risk patients have a worse experience of care in MA compared to other Medicare programs (Keenan et al. 2009, Halpern et al. 2017, Martino et al. 2016).[4] Furthermore, when getting sicker, MA patients tend to switch back to Traditional Medicare (c.f. Li et al. (2018), Ankuda et al. (2020), Meyers et al. (2019)). Healthcare inequality is a main concern of capitation programs as it is in direct contradiction with the purpose of capitation payment models (Rice et al. 1999).

Past literature mostly attributes the healthcare inequality in capitation programs to the misspecification of risk adjustment formulas. Specifically, Glazer and McGuire (2000) theoretically showed that when risk adjustment is imperfect, healthier patients tend to be overcompensated in MA while sicker patients tend to be undercompensated in MA. As such, MA health plans would have incentives to attract the overcompensated patients and to deter the undercompensated patients, which results in a disparity of service provisions. In fact, Brown et al. (2014) estimated that the annual healthcare costs of each MA enrollee were on average $317 lower than their capitation rates after risk selection, and further empirically demonstrated that imperfect risk adjustment can lead to risk selection in MA.

[3] Compared to Medicare beneficiaries in TM, it is clear that MA enrollees have lower accessibility to high-quality nursing homes (c.f. Meyers et al. (2018)), home health services (c.f. Schwartz et al. (2019)), post-acute care facilities (c.f. Figueroa et al. (2020)) and cardiologists (c.f. Huesch (2010)).

[4] For example, Halpern et al. (2017) studied the experience of patients with end-stage cancer and found that MA patients had less access to care and a worse care experience compared to their counterparts in TM.

Findings in this literature, however, have not been corroborated by subsequent empirical studies on MA. In particular, Newhouse et al. (2013) found that MA health plans did not over-provide high-profit healthcare services and under-provide low-profit healthcare services in practice. Therefore, they argued that even though the current MA risk adjustment formula is misspecified and provides incentives for MA health plans to risk select patients, "there is no evidence here that plans act on this incentive". In addition, a recent policy report documents a positive association between the number of chronic conditions an MA enrollee has and the potential overpayment for this MA enrollee (Jacobson et al. 2019b). As such, the imperfect risk adjustment explanation (c.f. Glazer and McGuire (2000), Brown et al. (2014)), although theoretically plausible, is empirically not well-supported.

To better explain the observed healthcare inequality and risk selection in MA, She et al. (2022) proposed a new theory based on cross subsidization practice. That is, if MA health plans can cross subsidize capitation payments across patients, they would "cherry-pick" patients not only according to how profitable these patients are, but also how costly it is to attract these patients. Therefore, the discrepancy between Brown et al. (2014) and Newhouse et al. (2013) can be reconciled if cross subsidization practice exists. In fact, the available empirical evidence suggests that it is easier and less costly for MA health plans to attract low-risk patients compared to high-risk patients (c.f. Aizawa and Kim (2018), Cooper and Trivedi (2012), Meyers et al. (2018), Han and Lavetti (2017)). However, while She et al. (2022) offer an appealing theory, there exists no work in the literature that empirically documents such cross subsidization practice in healthcare capitation programs. Building on this notion and expanding this literature, our paper provides the first empirical evidence of strategic cross subsidization in MA, and explains how it is associated with inequality of healthcare access and risk selection in MA.

This study also sheds light on recent debates regarding overpayment in MA (c.f. Ginsburg and Lieberman (2022)). A recent report to Congress delivered by the Medicare Payment Advisory Commission (2022) shows that CMS has consistently overpaid MA plans during the past 20 years compared to the amount TM would have spent on the same beneficiaries. There are many factors contributing to this overpayment, including the statutory overpayment due to the "minimum update rule" (c.f. U.S. Congress (2008)), upcoding incentives in MA (c.f. Geruso and Layton (2020)), and risk selection in MA (c.f. Brown et al. (2014)). To reduce this overpayment, ACA required CMS to start removing the statutory overpayment in MA from 2012. As a result, the overpayment in MA as a percentage of TM spending has dropped more than 50%, or $68 billion, in the past decade (Medicare Payment Advisory Commission 2022, Look 2019). Nevertheless, a recent study shows that MA enrollees' access to or affordability of care was not significantly affected by this ACA payment cut (Skopec et al. 2019a). This paper further analyzes this policy shock by examining MA plans' changes in spending patterns after the payment reduction and explains why accessibility and affordability of care in MA were not significantly affected. Particularly, we find that deeper capitation payment cuts for high-risk patients unexpectedly lead to more substantial spending reductions for low-risk, rather than high-risk, MA enrollees because of the strategic cross subsidization mechanism.

Our paper has important implications for the emerging operations management literature on healthcare payment reforms. Specifically, the existing literature on population-based payment models shows that failure to accurately estimate the certainty equivalent healthcare costs when determining capitation payments can lead to patient selection (Adida et al. 2016, Ata et al. 2013). Our findings complement and extend this line of research by empirically demonstrating that cross subsidization incentives can also lead to patient selection in population-based payment models such as capitation. Furthermore, by empirically identifying the cross subsidization practice, our work helps provide a more complete understanding of healthcare payment models based on past cost estimates (e.g., Erhun et al. (2015), Savva et al. (2018)) and shows that healthcare costs can be endogenous to healthcare payment models.

Our work also contributes to the empirical operations management literature on cherry-picking behavior in healthcare settings. Patterson et al. (2016) studied patient waiting time in emergency departments, and found that resident physicians tended to pick up patients with less complex complaints and avoid patients with more complex complaints. Similarly, Ibanez et al. (2018) empirically showed that physicians preferred to prioritize tasks with shorter expected processing time, albeit it leads to slower completion time on average. In addition, KC et al. (2020) demonstrated that physicians cherry-pick tasks "because of both fatigue and the sense of progress individuals get from task completion". Furthermore, Kc and Terwiesch (2011) provided evidence showing that these cherry-picking behavior were also found at the hospital level. Our paper extends this literature by providing empirical evidence of cherry-picking behavior at the health plan level.

Finally, our paper is broadly related to the empirical operations literature that focuses on identifying strategic behaviors in healthcare operations. Bastani et al. (2018) found that hospitals strategically respond to quality improvement regulations by "upcoding" their patients. Chen and Savva (2018) empirically documented that hospitals strategically utilized observation beds to avoid penalties from the Hospital Readmissions Reduction Program. Lu and Lu (2017) found that regulations limiting nurse overtime unintentionally reduce service quality as nursing homes strategically replace permanent nurses with contracted nurses to circumvent regulations. Our study enhances this literature and identifies the cross subsidization behaviors of MA health plans as a strategic response to the MA capitation payment model.

3. Hypotheses Development

The objective of this study is to empirically investigate how MA health plans actually allocate capitation payments among Medicare beneficiaries. Specifically, we are interested in whether there is a potential misallocation problem in MA in that MA health plans systematically overspend on low-risk patients and underspend on high-risk patients. To formally answer these questions, we first develop three research hypotheses in this section.

The first research hypothesis aims to explore if there is an inverse relation between an MA enrollee's health status and the healthcare resources she gets from MA. It is extensively documented in the medical literature that, compared to the standard Medicare coverage, MA health plans tend to provide more supplementary benefits (e.g., gym membership) and better preventive care (e.g., high accessibility to mammographic screening) but worse nursing home or post-acute care (e.g., lower star ratings in MA nursing homes or post-acute facilities) (Cooper and Trivedi 2012, Hung et al. 2016, Meyers et al. 2018, Rahman et al. 2015). Clearly, as an MA patient became sicker, they would utilize less preventive care but more nursing home care or post-acute care. This skewed provision of services essentially results in overspending on patients with low risk and underspending on high-risk patients. We thus hypothesize that:

Hypothesis 1: *There is an inverse relationship between the sickness levels of patients and the spending-cost differences of such patients.*

Here, spending-cost difference refers to the difference between MA health plans' actual healthcare spending on a patient and the estimated healthcare cost of this patient.

The main hypothesis of this paper aims to understand whether MA health plans reallocate parts of the capitation payments from high-risk patients to cross subsidize low-risk patients, a practice we refer to as strategic cross subsidization. In particular, when faced with a decision to invest in various benefits, MA health plans often favor lower-cost benefits that appeal to a broader range of low-risk enrollees over high-cost benefits primarily utilized by a limited number of high-risk enrollees (Skopec et al. 2019b). As an illustration, incorporating fitness memberships as a supplementary benefit costs MA health plans less than \$22 per enrollee per month (c.f. Medicare Payment Advisory Commission (2019)), yet this inclusion attracts beneficiaries exhibiting "significantly better general health" (Cooper and Trivedi 2012). Additionally, an experienced health plan executive reinforced in our discussion that MA health plans could effectively "free up funds" by reducing spending on high-risk patients. He illustrated this with a clear example: a 10% cut on a high-risk patient's \$1,000 per beneficiary per month (PBPM) capitation yields \$100 savings, double the amount saved from the same 10% cut on a low-risk patient's \$500 PBPM capitation. As a matter of fact, due to the lack of transparency in MA claim data, MA health plans have a lot of flexibility with the reallocation of resources. A recent report by the Office of Inspector General (OIG) for the United States Department of Health and Human Services found that MA health plans tend to "deny beneficiary access to services and deny payments to providers in an attempt to increase profits" (c.f. Department of Health and Human Services, Office of Inspector General (2022)). To empirically identify this potential misallocation problem in MA, we test whether the inverse relationship studied in Hypothesis 1 is caused by the strategic cross subsidization practice of MA health plans.

Main Hypothesis (Hypothesis 2): *There exists strategic cross subsidization practice in MA, which causes the inverse relationship characterized in Hypothesis 1.*

To provide further support for the strategic cross subsidization hypothesis, we conducted an empirical investigation into its effects on risk selection within MA. As outlined in the introduction, the ACA policy change resulted in more significant capitation payment reductions for high-risk patients compared to their low-risk counterparts. Based on the prevailing risk selection theory (e.g., Brown et al. (2014)), this underpayment for high-risk patients would exacerbate risk selection in MA, leading to a higher likelihood of high-risk patients leaving MA and reverting to TM. Nevertheless, according to our strategic cross subsidization hypothesis, these more substantial capitation payment cuts for high-risk patients would lead to greater spending reductions for low-risk rather than high-risk MA enrollees. Consequently, the ACA policy change would alleviate the risk selection issue within MA. This potential influence of strategic cross subsidization on MA is encapsulated in the following hypothesis:

Hypothesis 3: *The ACA capitation payment reductions mitigated risk selection in MA.*

4. Institutional Background of Medicare Advantage (MA)

The first part of this section provides a general introduction to the Medicare Advantage program, with emphasis on its capitation payment model (§4.1). The second part discusses the Affordable Care Act (ACA) as a shock to this capitation payment model in MA (§4.2).

4.1. Medicare Advantage and Its Capitation Payment Model

Medicare Advantage (Medicare Part C) is an insurance program that provides coverage to Medicare beneficiaries, i.e. disabled or senior Americans (at least aged 65), for their healthcare services. It was formally established by the Balanced Budget Act of 1997 as an option for Medicare beneficiaries to receive Traditional Medicare benefits (Medicare Part A and B) through private health plans. More specifically, CMS outsources the inpatient hospital and outpatient services to health plans in MA, and reimburses these services through risk-adjusted capitation payments. By 2019, 34% or 22 million Medicare beneficiaries have enrolled in Medicare Advantage, which makes MA the largest capitation program in the U.S. (Jacobson et al. 2019a).

As a payer in MA, CMS's main responsibility is to use capitation payments to incentivize MA health plans to deliver similar healthcare services to MA enrollees as those in Traditional Medicare. Specifically, the Social Security Act requires CMS to set the MA capitation rates so that the benefits delivered in MA are actuarially equivalent to those in Traditional Medicare (Medicare Part A and Part B). To this aim, CMS first estimates the risk scores ($RiskScore_{i,t}$) of each Medicare beneficiary i based on individual risk factors such as age, gender, and pre-existing conditions at time t using the CMS-HCC risk adjustment formula (Centers for Medicare & Medicaid Services 2017). In addition, CMS calculates the benchmark payments ($Benchmark_{c,t}$) for each average (standard risk) enrollee of MA health plans at county c. As such, to provide equivalent healthcare services as those in Medicare Part A and Part B to an MA enrollee i living in county c at year t, the capitation payment an MA health plan receives from CMS is

$$CapitationRate_{i,t} = RiskScore_{i,t} \times Benchmark_{c,t}, \tag{1}$$

which is adjusted yearly.[5] Other details of MA capitation payments are provided in Appendix A.

4.2. The Affordable Care Act (ACA) Reform of MA Capitation Payments

The Affordable Care Act (ACA), which was signed into law in 2010, required CMS to reduce the benchmark payment ($Benchmark_{c,t}$) in MA capitation rates (1). The main motivation for this reduction was to reduce the overpayment MA health plans received for providing healthcare services to MA beneficiaries before the 2010 ACA reform.[6]

These benchmark payment reductions were not uniform across U.S. counties. Specifically, counties with lower local area per capita costs of Traditional Medicare compared to the national per capita costs of Traditional Medicare would have larger projected benchmark reductions. Figure 2 in Appendix B provides a heatmap to visualize the projected benchmark reductions for each U.S. county. These benchmark payment reductions range from around \$0 PBPM in some counties to around \$300 PBPM in other counties. In other words, some counties have disproportionately large

[5] See Appendix A.3 for further explanations.

[6] It was estimated that, in 2009, the per Medicare beneficiary reimbursement in MA was 13% higher compared to the corresponding reimbursement in Traditional Medicare (Part A and Part B) (Biles et al. 2009). These overpayments were mainly due to the "minimum update rule" (c.f. U.S. Congress (2008)), which required CMS to set the pre-ACA benchmark payments equal to the greater of the percentage increase in national per capita costs of Traditional Medicare ("minimum update rate") and the local area per capita costs of Traditional Medicare. To reduce overpayment in MA, the ACA benchmark payments were adjusted to only reflect the local area per capita costs of Traditional Medicare but not the "minimum update rate".

	Two Year Phase-in Group		Four Year Phase-in Group		Six Year Phase-in Group	
Year	Pre-ACA	ACA	Pre-ACA	ACA	Pre-ACA	ACA
2012	1/2	1/2	3/4	1/4	5/6	1/6
2013	0	1	1/2	1/2	2/3	1/3
2014	0	1	1/4	3/4	1/2	1/2
2015	0	1	0	1	1/3	2/3
2016	0	1	0	1	1/6	5/6
2017	0	1	0	1	0	1

Table 1 The Phase-in Factor of ACA capitation Rates of County c at Year t ($PhaseInFactor_{c,t}$): Counties across the U.S. phased into the ACA benchmark in three groups. The Two Year Phase-in Group: These counties phased in $\frac{1}{2}$ of the ACA capitation rates in 2012, and fully phased in afterwards. The Four Year Phase-in Group: These counties phased in $\frac{1}{4}$ of the ACA capitation rates in 2012, phased in $\frac{1}{2}$ of the ACA capitation rates in 2013, phased in $\frac{3}{4}$ of the ACA capitation rates in 2014, and fully phased in afterwards. The Six Year Phase-in Group: These counties phased in $\frac{1}{6}$ of the ACA capitation rates in 2012, phased in $\frac{1}{3}$ of the ACA capitation rates in 2013, phased in $\frac{1}{2}$ of the ACA capitation rates in 2014, phased in $\frac{2}{3}$ of the ACA capitation rates in 2015, phased in $\frac{5}{6}$ of the ACA capitation rates in 2016, and fully phased in afterwards (Centers for Medicare & Medicaid Services 2011).

benchmark payment reductions compared to other counties. The ACA recognized these differential impacts of the benchmark payment reductions on different counties, and designed a phase-in scheme to smooth out the transition process.

To smooth the transition from pre-ACA to ACA benchmark payments, ACA specified a 6-year transition period (2012-2017) to gradually phase in these payment reductions. Specifically, ACA divided counties across the U.S. into three groups, namely the Two, Four, and Six Year Phase-in Groups. Figure 3 in Appendix B plots the geographic distribution of counties in different Phase-in Groups. In the transition period (2012-2017), ACA benchmark payments were phased into different Phase-in Groups in a staggered manner, according to the phase-in factor ($PhaseInFactor_{c,t}$) listed in Table 1. More precisely, the benchmark payments for each MA health plan at county c during years $t \in \{2012, \ldots, 2017\}$ were calculated as $Benchmark_{c,t} = ACA\_Benchmark_{c,t} \times PhaseInFactor_{c,t} + PreACA\_Benchmark_{c,t} \times (1 - PhaseInFactor_{c,t})$, where $PhaseInFactor_{c,t} \in [0, 1]$ as specified in Table 1. In other words, the actual MA benchmark payments in the transition period were a weighted average of the ACA benchmark payments ($ACA\_Benchmark_{c,t}$) and the pre-ACA benchmark payments ($PreACA\_Benchmark_{c,t}$), where counties c with higher "weight" ($PhaseInFactor_{c,t}$) would phase into the ACA benchmark earlier.

To see how this phase-in group assignment smooths the transition from the pre-ACA benchmark to the ACA benchmark, we shall explain how the $PhaseInFactor_{c,t}$ presented in Table 1 was determined. Specifically, $PhaseInFactor_{c,t}$ was calculated based on its projected difference between the ACA and Pre-ACA benchmark payments, i.e. $PreACA\_Benchmark_{c,t} - ACA\_Benchmark_{c,t}$.[7]

[7] A detailed illustration is given in Appendix A on how to calculate this projected benchmark difference based on the data and official documents listed on the CMS website.

The idea is that counties with larger projected benchmark cuts would have lower phase-in factors ($PhaseInFactor_{c,t}$) and thus longer phase-in periods. More precisely, counties with projected benchmark payment cut less than $30 PBPM were assigned to the two-year phase-in group, while counties with at least $30 PBPM projected benchmark cuts were assigned to groups with longer, i.e. four or six-year, phase-in periods. This phase-in group assignment rule ensures that no counties would experience disproportionately large annual benchmark cuts during the transition period,[8] and thus smooths the transition process. In §6.2.2, we describe in detail how we exploit this Phase-in Group assignment rule to implement our DID design.

5. Data Sources and Variables' Construction

Data used in this study are mainly from two sources. To test Hypotheses 1 and 2, we use individual patient-level data from Optum's de-identified Clinformatics® Data Mart (CDM) between 2009-2015. This study period contains years before (2009-2011) and after (2012-2015) the ACA capitation payment cuts in which the same CMS-HCC risk adjustment formula was used in MA. Therefore, this is a suitable time frame to employ our main identification strategy, i.e. Difference-in-Difference (DID). To test Hypothesis 3, we use county-level data from the public use files on CMS websites in the same study period. §5.1 and §5.2 introduce these two data sources and explain the construction of relevant variables from these datasets.

5.1. Individual-level Data

The individual-level data used in this study comes from CDM, a commercial insurance claims database that contains medical claims from more than 150 million people in the U.S. (Wallace et al. 2014). In this study, we focus on a subsample of this database consisting of Medicare beneficiaries who enrolled in MA for at least one year from 2009 to 2015 and were at least age 65 during the time of enrollment. We exclude those individuals who had moved to another county during a year within the study period.[9] As the healthcare cost estimate of an MA enrollee depends on the county in which this enrollee lived, i.e. Estimated $\text{Cost}_{i,t} := RiskScore_{i,t} \times \text{FFS Rate}_{c,t}$, it is not possible to make a reliable estimation of this variable if a patient lived in more than one county in year t. Furthermore, we exclude MA enrollees with risk scores above 7.5, i.e. the top 0.1 percentile sickest patient in MA, because these patients typically experienced medical complications that were not representative of the MA population.

[8] For example, a county with $40 PBPM projected benchmark cut would be assigned to the four-year phase-in group, and thus only had $\$40\times\frac{1}{4}=\$10$ PBPM benchmark cut in 2012. In contrast, a county with $20 PBPM projected benchmark cut would also have $\$20\times\frac{1}{2}=\$10$ PBPM benchmark cut in 2012, because it is in the two-year phase-in group.

[9] These "movers" accounted for about 8% of the total Medicare beneficiaries in the study sample.

The final sample contains MA claims data from 3,882,330 distinct Medicare beneficiaries living in 2,787 counties or county equivalent in 50 states and DC in the U.S. across 7 years (2009-2015). The individual-level dataset is used to construct variables to test Hypotheses 1 and 2. These variables are illustrated as follows. $RiskScore_{i,t}$ is the risk score of Medicare beneficiary i in MA at the beginning of year t. This score is calculated using the established CMS-HCC risk adjustment formula, which requires inputs of this Medicare beneficiary's age, gender, and preexisting conditions at the end of year $t-1$. (Centers for Medicare & Medicaid Services 2017). The CDM provides these individual characteristics of Medicare beneficiary i to calculate his/her risk score. MA $\text{Spending}_{i,t}$ is the total amount an MA health plan spent on an MA beneficiary i at year t. Specifically, this variable is calculated by summing up all MA medical claims generated by patient i and paid by her MA health plan in year t.[10] FFS $\text{Rate}_{c,t}$ is the estimated Traditional Medicare spending on an average (standard risk) MA beneficiary in county c during year t. This variable is calculated based on the 2009-2015 county-level Traditional Medicare data publicly available on the CMS's website (c.f. Centers for Medicare & Medicaid Services (2019)). Estimated $\text{Cost}_{i,t}$ is the estimated amount CMS would spend on patient i in year t if this patient was in TM during year t. Specifically, this variable is calculated as Estimated $\text{Cost}_{i,t} = RiskScore_{i,t} \times$ FFS $\text{Rate}_{c,t}$ based on the capitation payment formula (1), with the MA benchmark payment $Benchmark_{c,t}$ replaced by the TM benchmark spending FFS $\text{Rate}_{c,t}$. $\text{Age}_{i,t}$ is the age of MA enrollee i in year t. $Treatment_{i,t}$ is a binary variable indicating whether the county in which Medicare beneficiary i lived was in the treatment group (=1) or the control group (=0) at year t. The assignment rule for treatment and control groups is discussed in §6.2.2.

In Appendix B, we present tables (See Tables 9 and 10) summarizing county-level characteristics (e.g. enrollment size and quality star ratings of different MA health plan types) as well as individual-level characteristics (e.g. risk scores and healthcare spending) of CDM throughout our study period (2009-2015).

5.2. County-level Data

All data used to conduct the county-level analysis in this study are publicly available on the CMS website. Specifically, we use county-level MA data, Traditional Medicare data and health-plan-level MA data from Centers for Medicare & Medicaid Services (2019). We focus on the analysis

[10] Optum does not provide the actual price paid by MA health plans for each service, and only has the amount charged by physicians and hospitals for each service. Nevertheless, it offers a standardized algorithm to convert these charged prices to actual paid prices. For example, the professional service (e.g. physician visits, surgery, lab tests, imaging) charges are standardized using the resource-based relative value scale (RBRVS) approach, a method applied by CMS to price services provided in Traditional Medicare. We use these standardized charged prices to create a consistent proxy for the actual expense.

of non-special needs Health Maintenance Organizations (HMOs) in the main text as they are the dominant health plan type in this market (Song et al. 2013). We also present an analysis for both HMOs and Preferred Provider Organizations (PPOs) in Appendix C.5. The final sample in this dataset contains observations from 2,968 counties or county equivalent in 50 states and DC in the U.S. across 7 years (2009-2015).

The county-level dataset is used to construct variables to test Hypothesis 3. Specifically, $RiskScore_{c,t}$ is the average risk score of MA beneficiaries in county c of year t. $Treatment_{c,t}$ is a binary variable indicating whether county c was in the treatment group (=1) or the control group (=0) at year t. The assignment rule for treatment and control groups is discussed in §6.2.2. $Star_{c,t}$ is the average star rating of health plans in county c at year t. $McPop_{c,t}$ is the population size of Medicare beneficiaries, including both TM and MA, in county c of year t. $Rebate_{c,t}$ is the average rebate amount for MA health plans to provide supplementary benefits in county c of year t. $RUCC_{c,2010}$ is the Rural-Urban Continuum Code (ranging from 1-9) which classifies the rural status of each U.S. county in 2010 (United States Department of Agriculture 2013). Here, the higher the $RUCC_{c,2010}$, the more rural county c is. $\text{Income}_{c,2010}$ is the median household income of county c at year 2010.

6. Empirical Evidence of Strategic Cross Subsidization

This section provides empirical evidence of strategic cross subsidization in Medicare Advantage. Specifically, §6.1 develops the econometric model to test Hypothesis 1, empirically showing an inverse relationship between the sickness levels of patients and the spending-cost differences of such patients. §6.2 illustrates our empirical strategy to test the main hypothesis (i.e. Hypothesis 2), and presents the econometric model based on this strategy. Particularly, we empirically show that the inverse relationship established in Hypothesis 1 is caused by the strategic cross subsidization practice of MA health plans.

6.1. Hypothesis 1

Hypothesis 1 states that there is an inverse relationship between the sickness levels of patients and the spending-cost differences of such patients. To empirically test this hypothesis, we construct the econometric model in §6.1.1, and discuss the test results in §6.1.2.

6.1.1. The Econometric Model to Test Hypothesis 1

To test Hypothesis 1, it suffices to examine whether there is a negative association between Medicare beneficiaries' risk scores ($RiskScore_{i,t}$) and their spending-cost differences ($\text{MA Spending}_{i,t} - \text{Estimated Cost}_{i,t}$). CMS assigns a risk score, $RiskScore_{i,t}$, to each Medicare beneficiary i in MA at the beginning of each year t. Following the existing literature on risk selection (c.f. Brown et al. (2014), Newhouse et al. (2013)), we use the risk score of a patient as a proxy for the sickness level of this patient. The spending-cost difference of an MA health plan enrollee i at year t is given by the difference between MA health plans' spending on patient i at year t and the estimated healthcare cost of this patient at year t, i.e. MA $\text{Spending}_{i,t} - \text{Estimated Cost}_{i,t}$.

The corresponding econometric model is thus specified as:

$$\text{MA Spending}_{i,t} - \text{Estimated Cost}_{i,t} = \beta_0 + \beta_{RS} RiskScore_{i,t} + \alpha_i + \gamma_t + \epsilon_{i,t}, \tag{2}$$

where α_i and γ_t are the individual and time fixed effects, respectively. Individual fixed effects control for the service utilization patterns of different Medicare beneficiaries, whereas the time fixed effect takes into account any pure temporal changes in the study period.

If Hypothesis 1 is true, we should have $\beta_{RS} < 0$. That is, as an MA health plan enrollee gets sicker, the MA health plan's spending on this patient does not increase as fast as the estimated healthcare cost of this patient.

β_{RS}	-9.97***
Observations	12,937,959
R^2	0.533
Adjusted R^2	0.330

Note: Dependent variable is in $1,000 per year unit. All standard errors are clustered at county level. *p<0.1; **p<0.05; ***p<0.01

Table 2 Estimation Results of Econometric Model (2)

6.1.2. Results for Hypothesis 1

We estimate the econometric model (2) using two-way fixed effect estimators. The estimation results presented in Table 2 confirm Hypothesis 1. That is, as an MA beneficiary becomes sicker, the healthcare spending on this MA beneficiary does not increase as fast as her estimated healthcare cost does, i.e. $\beta_{RS} < 0$. Specifically, a one-unit increase of a patient's risk score is associated with a more than $9,000 decrease in the relative annual MA health plan spending on this patient compared to her estimated healthcare cost, i.e. spending-cost difference. As such, we establish an inverse relationship between the sickness of MA beneficiaries and their spending-cost differences.

To better understand this inverse relationship, we conduct several case studies to examine the changes in spending-cost differences and risk scores of MA enrollees after they develop certain medical conditions. As shown in Table 3, after being diagnosed with renal diseases, an MA enrollee

	Renal Diseases		Cancer	
	Before	After	Before	After
Spending-Cost Difference	\$5,843.59	-\$686.18	\$7,731.53	\$2,703.56
Risk Score	1.18	2.05	1.09	1.76

Table 3 Case Study: spending-cost differences and risk scores of MA enrollees before and after they develop renal diseases (HCC130, HCC131) and Cancer (HCC7-HCC10)

would on average have a 0.87 points increase in risk scores. However, her spending-cost difference would on average decrease from \$5,843.59 to -\$686.18. That is, compared to healthier MA enrollees without renal diseases, MA health plans actually spend less on sicker MA enrollees with renal diseases relative to their estimated healthcare cost. This finding is consistent with the literature (c.f. Li et al. (2018)) that MA health plans are less likely to satisfy the care needs of incident renal disease patients with newly initiated dialysis. Table 3 reports similar results also for MA enrollees who have developed cancer. Furthermore, Table 15 and 16 in Appendix B find similar results in conditions including Diabetes, Cardiovascular diseases, Inflammatory Bowel Disease (IBD), Cerebrovascular diseases, Chronic Hepatitis and HIV/AIDS. These findings further support the econometric model's main estimation results (2).

6.2. Main Hypothesis (Hypothesis 2)

In order to get a causal interpretation of the association in (2), we need an exogenous shock to this association which does not affect how $RiskScore_{i,t}$ was estimated. In other words, if $RiskScore_{i,t}$ was estimated in the same way before and after this shock, any changes in this association due to this shock can be identified as the causal impact. §6.2.1 discusses that the ACA capitation rate cuts provided such an exogenous shock and thus can serve as our source of identification. §6.2.2 discusses the assignment of treatment and control groups in this quasi-natural experiment. Finally, we develop the econometric model in §6.2.3 and present the estimation results in §6.2.4.

6.2.1. Identifying Strategic Cross Subsidization through ACA Capitation Rate Cuts

As discussed in §4.2, the ACA reform changed MA capitation payments by reducing their benchmark payments. Specifically, we note that the ACA capitation rate cuts have an asymmetric impact on MA patients in different risk groups in that patients with higher risk scores experienced deeper cuts in their capitation payments ($CapitationRate_{i,t}$). The reason is that the ACA capitation rate cuts only affect $CapitationRate_{i,t} := RiskScore_{i,t} \times Benchmark_{c,t}$ (c.f. (1)) through the benchmark payments ($Benchmark_{c,t}$). As such, the same benchmark payment cuts would result in deeper capitation payment cuts for MA patients with higher risk scores ($RiskScore_{i,t}$). Therefore, the

ACA capitation rate cuts essentially exogenously reduce greater amount of capitation payments from the high-risk patients than from the low-risk patients.

Based on this observation, we can empirically identify the strategic cross subsidization practice of MA health plans by examining how the observed inverse relationship between risk scores and spending-cost differences would change after the ACA shock. As illustrated in Figure 1a, if there were no strategic cross subsidization, the ACA shock, with its deeper payment cut for high-risk patients, would simply amplify the inverse relationship between risk scores and spending-cost differences. In contrast, if the inverse relationship between risk scores and spending-cost differences were caused by the strategic cross subsidization practice of MA health plans, the ACA shock would mitigate this inverse relationship. The reason is as follows. With strategic cross subsidization, we argue that MA health plans were using the capitation payments collected from the high-risk patients to cross subsidize the low-risk patients (i.e., the yellow shaded regions in Figure 1b). Therefore after the ACA shock which exogenously brings deeper capitation payment cuts for high-risk patients, there will be less money available for MA health plans to cross subsidize low-risk patients. If this is the case, strategic cross subsidization will subside as in Figure 1b.

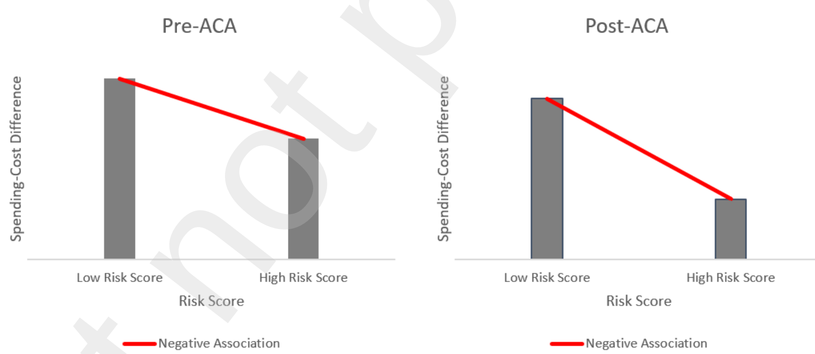

Figure 1a **Identification Strategy Based on ACA Cut (No Cross Subsidization):** If the inverse relationship between risk scores and spending-cost differences were not caused by strategic cross subsidization, the ACA shock would not mitigate this inverse relationship. Here, the inverse relationships in the Pre-ACA and Post-ACA figures correspond to β_{RS} and $\beta_{RS} + \delta_{DID}$ in (3), respectively.

Hence, to test the main hypothesis (i.e., Hypothesis 2), it suffices to check whether the inverse relationship between risk scores and spending-cost differences was alleviated post the ACA shock. If further underpayment for high-risk patients indeed results in amplified underspending on these patients, we would anticipate an intensified inverse relationship following the ACA shock, as illustrated in Figure 1a. Alternatively, if we observe a pattern akin to Figure 1b, it suggests that

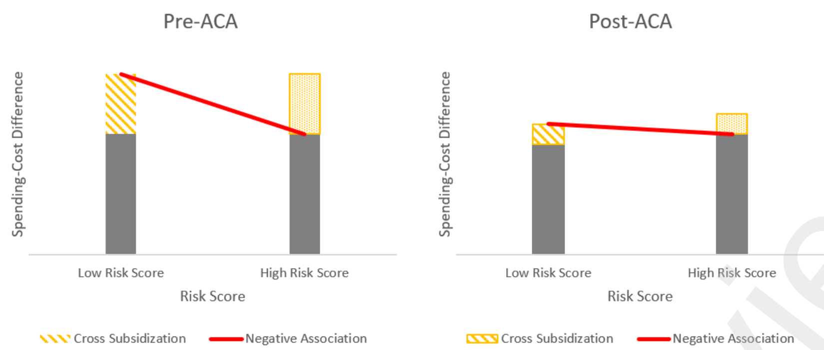

Figure 1b **Identification Strategy Based on ACA Cut (Cross Subsidization):** If the inverse relationship between risk scores and spending-cost differences were caused by the strategic cross subsidization practice of MA health plans, the ACA shock would mitigate this inverse relationship. Here, the inverse relationships in the Pre-ACA and Post-ACA figures correspond to β_{RS} and $\beta_{RS} + \delta_{DID}$ in (3), respectively.

exacerbated underpayment for high-risk patients does not instigate further underspending on these patients, but instead mitigates overspending on their low-risk counterparts. Consequently, the heightened spending-cost differences or overspending on low-risk patients and diminished spending-cost differences or underspending on high-risk patients cannot be solely ascribed to overpayment for the former and underpayment for the latter. A cross subsidization mechanism, therefore, becomes essential to explain the inverse relationship between risk scores and spending-cost differences.

6.2.2. Treatment Assignment

To effectively isolate the impact of ACA capitation rate cuts on the inverse relationship between risk scores and spending-cost differences from other confounding factors, we need a near-random assignment of capitation rate cuts across U.S. counties. The staggered roll-out of these capitation rate cuts across U.S. counties provides a quasi-natural experiment to separate treatment and control groups. Specifically, as discussed in §4.2, each U.S. county was assigned to different phase-in groups in our study period. Furthermore, these group assignments were based on the projected difference between the Pre-ACA benchmarks and ACA benchmarks of each county. These two institutional details provide us with the quasi-random regional variations of ACA capitation rate cuts for treatment assignment.

First, the phase-in group assignment of each U.S. county (c.f. Table 1) enables us to implement the Difference-in-Difference (DID) design. Specifically, during the transition period, counties with shorter phase-in time had higher percentage capitation rate cuts compared to counties with longer phase-in time. For example, 50% of the ACA capitation payment cuts became effective in 2012 for counties in the Two Year Phase-in Group. In contrast, for counties in the Four Year and Six Year Phase-in Groups, only 25% and 17% of the ACA capitation payment cuts became effective in 2012.

Therefore, in principle, counties with shorter phase-in time should be assigned to the treatment group, while counties with longer phase-in time should be assigned to the control group.

Second, phase-in time for counties was determined based on their projected difference between the Pre-ACA benchmarks and ACA benchmarks, and whether this value fell above or below an arbitrary cut-off point. In particular, counties with projected benchmark differences below \$30 PBPM were assigned to the Two Year Phase-in Group, while counties with projected benchmark differences of at least \$30 PBPM were assigned to the Four Year and Six Year Phase-in Groups. As such, this \$30 PBPM provides a clean cut-off value to implement a "local randomized experiment" as defined by Lee and Lemieux (2010).[11] That is, by focusing on the counties with benchmark differences just above or just below the \$30 cut-off value, we have a study sample where the ACA capitation payment cuts were assigned independent of other temporal changes. The reason is that except for the ACA capitation payment cuts, no other temporal changes would affect our study sample via the \$30 cut-off value. Therefore, we use this \$30 cut-off value to create a quasi-random assignment of the ACA capitation payment cuts across otherwise similar U.S. counties.

In our DID analysis, counties with projected benchmark differences slightly below the \$30 cut-off value are assigned to the treatment group, while counties with projected benchmark differences slightly above the \$30 cut-off value are assigned to the control group. Specifically, we test our main hypothesis in three different neighborhood specifications around the \$30 cut-off value (i.e. [\$29,\$31], [\$28,\$32], [\$27,\$33]), and assess whether they provide consistent treatment effect estimates. As an illustration, Figure 4 in Appendix B plots the treatment-control group assignment in the [\$27,\$33] neighborhood specification. Lastly, we remark that the phase-in group assignment and phase-in time were announced in February 2011 based on the projected benchmark difference from the 2010 data (Centers for Medicare & Medicaid Services 2011). Therefore, the treatment path was predetermined before 2012.

6.2.3. The Econometric Model to Test the Main Hypothesis

The econometric model to test our main hypothesis is specified as:

$$\begin{aligned} & \text{MA Spending}_{i,t} - \text{Estimated Cost}_{i,t} \\ = & \beta_0 + \delta_{DID} RiskScore_{i,t} \times Treatment_{i,t} + \beta_1 Treatment_{i,t} \\ + & \beta_{RS} RiskScore_{i,t} + \alpha_i + \gamma_t + \epsilon_{i,t}, \end{aligned} \tag{3}$$

[11] For counties with projected benchmark difference slightly below \$30 PBPM, they experienced $50\% - 25\% = 25\%$ deeper capitation payment cuts than counties with projected benchmark difference slightly above \$30 PBPM.

where the treatment $Treatment_{i,t}$ is assigned as described in §6.2.2. In particular, since treatment and control groups are assigned at the county-level, the individual identifiers i become county-patient specific.[12] In this specification, if our main hypothesis was true, then $\delta_{DID} > 0$. This is because the inverse relationship between risk scores and spending-cost differences (β_{RS}) estimated in (2) would be mitigated by the ACA capitation rate cuts, i.e. $\beta_{RS} < \beta_{RS} + \delta_{DID}$.

Consistent with prior research employing similar identification strategies (c.f. Verhoogen (2008), Flach (2016), Irarrazabal et al. (2015)), the econometric model (3) can be systematically formulated in three steps. First, we identify the inverse relationship between risk scores and spending-cost differences in the control, i.e., $\beta_{RS(Control)}$. That is, for patients i in the control counties or in the control period (t before the ACA shock), we estimate

$$\begin{aligned}&\text{MA Spending}_{i,t} - \text{Estimated Cost}_{i,t}\\&=\beta_{0(Control)} + \beta_{RS(Control)} RiskScore_{i,t} + \alpha_i + \gamma_t + \epsilon_{i,t}.\end{aligned}$$

Second, we identify the inverse relationship between risk scores and spending-cost differences in the treatment, i.e., $\beta_{RS(Treatment)}$. That is, for patients i in the treatment counties and in the treatment period (t after the ACA shock), we estimate

$$\begin{aligned}&\text{MA Spending}_{i,t} - \text{Estimated Cost}_{i,t}\\&=\beta_{0(Treatment)} + \beta_{RS(Treatment)} RiskScore_{i,t} + \alpha_i + \gamma_t + \epsilon_{i,t}.\end{aligned}$$

Third, we identify the change in the inverse relationships between control and treatment, i.e.

$$\delta_{DID} := \beta_{RS(Treatment)} - \beta_{RS(Control)}. \tag{4}$$

Mathematically, it is clear that the parameter of interest in (3), i.e., δ_{DID}, can be identified as (4). Econometrically, Verhoogen (2008) noted that this identification strategy "would amount to a familiar triple-differences strategy" if $RiskScore_{i,t}$ was a binary variable. As such, the econometric model (3) can be structured as a triple-difference design with continuous heterogeneous treatment effects, as summarized in Table 8 of Appendix B.

6.2.4. Results for the Main Hypothesis

The econometric model (3) is estimated using two-way fixed effect estimators. Table 4 presents the estimation results from different neighborhood specifications around the $30 cut-off value, i.e. [$29,$31], [$28,$32], [$27,$33]. These different neighborhood specifications provide consistent results supporting our main hypothesis, i.e. $\delta_{DID} > 0$. In other words, the inverse relationship between the sickness levels of MA beneficiaries and their spending-cost differences was mitigated by the

	DID([$29,$31])	DID([$28,$32])	DID([$27,$33])
β_{RS}	-10.15***	-9.74***	-9.65***
δ_{DID}	2.02***	1.64***	1.51***
β_1	-1.88**	-1.98***	-1.92***
Observations	588,946	779,436	984,339
R^2	0.500	0.508	0.528
Adjusted R^2	0.311	0.309	0.323

Note: Dependent variable is in $1,000 per year unit.
All standard errors are clustered at county level.
*p<0.1; **p<0.05; ***p<0.01

Table 4 Estimation Results of Econometric Model (3)

ACA capitation rate cuts, i.e., $|\beta_{RS}| > |\beta_{RS} + \delta_{DID}|$. As such, we confirm the main hypothesis and empirically identify the strategic cross subsidization practice in MA.

Next, we discuss the effect size of the ACA capitation rate cuts on MA health plans' spending patterns. Specifically, the strategic cross subsidization amount was decreased by around $\delta_{DID} =$ \$1640 (according to the [$28,$32] neighborhood) per patient per year for one unit risk score increase after the ACA capitation rate cuts. In other words, the ACA shock mitigated around 17% of the strategic cross subsidization amount, as measured by $|\frac{\delta_{DID}}{\beta_{RS}}|$. We note that this effect size is economically significant when taking into account the scale of the MA population. For example, there were around 15 million MA enrollees in 2014 (c.f. Jacobson et al. (2016)), where the mean annual change in individual risk score was 0.14 (c.f. Tables 9). Therefore, the mitigated strategic cross subsidization in 2014 was around $0.14 \times \$1640 = \$229.6$ per patient or 15 million $\times\, 0.14 \times$ $\$1640 = \$3,444$ million $\approx \$3$ billion in total.[13]

7. Assessing the Impact of Strategic Cross Subsidization Practice on Risk Selection

In the presence of strategic cross subsidization, Hypothesis 3 posits that the ACA policy change, which imposed more substantial capitation payment reductions for high-risk patients, would mitigate risk selection in MA, as detailed in §3. To empirically assess this hypothesis, we rely on county-level data as described in §5.2 rather than individual-level data from CDM. The reason is that OPTUM does not publish the design weights, i.e., the inverse of sampling probability, for

[12] In the rest of this paper, all individual identifiers i are county-patient specific.

[13] Alternatively, we also make use of quartile statistics of the individual risk score change in 2014 to provide a lower bound for this estimate. Specifically, 25% of our studied population had at least a 0.3 unit increase in risk score during 2014. Therefore, the mitigated strategic cross subsidization in 2014 was at least $0.25 \times 0.3 \times \$1640 = \$123$ per patient or 15 million $\times\, 0.25 \times 0.3 \times \$1640 = \$1,845$ million $\approx \$2$ billion in total.

users to recover population estimates. In the rest of this section, §7.1 develops a generalized triple-difference model to examine this association. §7.2 presents and discusses the estimation results.

7.1. The Econometric Model to Test Hypothesis 3

One way to evaluate Hypothesis 3 is to analyze whether counties phased into the ACA benchmarks earlier had higher county average MA risk scores in comparison to other counties. That is, the ACA capitation payment reductions should alleviate risk selection in a staggered manner, with counties that phased in earlier experiencing a more mitigated risk selection.

However, since the ACA shock affects risk selection through multiple channels, it is difficult to isolate the impact of strategic cross subsidization practice on risk selection in this shock. For example, along with the capitation payment cuts, the ACA also created a Quality Bonus Payments (QBPs) program to incentivize better quality provision in MA through bonus payments.[14] It is shown that the QBPs program enabled high-quality health plans to select more healthier Medicare beneficiaries (c.f. Fioretti et al. (2019)). As such, MA health plans with more quality bonus payments might have more severe risk selection problems. Therefore, the existence of alternative channels through which the ACA shocks can affect risk selection might confound the estimation of the association between strategic cross subsidization and risk selection.

To control for these alternative channels, we use a generalized triple-difference model to isolate the impact of strategic cross subsidization on risk selection, i.e.

$$
\begin{aligned}
RiskScore_{c,t} = &\beta_0 + \delta_{DID} Treatment_{c,t} + \beta_{Star(DID)} Treatment_{c,t} \times Star_{c,12} \\
&+ \beta_{Star(Post)} Star_{c,12} \times I\{PostACA\} + \beta_{Star} Star_{c,t} \\
&+ \beta_{McPop} log(McPop_{c,t}) + \beta_{Rebate} Rebate_{c,t} + \alpha_c + \gamma_t + \epsilon_{c,t}.
\end{aligned} \tag{5}
$$

That is, in addition to assigning counties to the treatment/control groups, we also group counties by the county-average initial star ratings at the beginning of the QBPs program, i.e. $Star_{c,12}$, which determines the county-average additional quality bonus payments from ACA.[15] As such, the treatment effect is moderated by the additional quality bonus payments from ACA. In addition, we also control for the population size of Medicare beneficiaries, i.e., $McPop_{c,t}$, and the county-average funds available for MA health plans to provide supplementary benefits, i.e., $Rebate_{c,t}$. Both factors affect the flexibility for MA health plans to conduct strategic cross subsidization as well as risk selection.

	[\$29,\$31]	[\$28,\$32]	[\$27,\$33]
δ_{DID}	0.185**	0.139**	0.109**
$\beta_{Star(DID)}$	-0.050**	-0.045***	-0.036***
$\beta_{Star(Post)}$	-0.028**	-0.014	-0.020**
β_{Star}	-0.035***	-0.009	-0.008
β_{McPop}	-0.011*	-0.011**	-0.011***
β_{Rebate}	2×10^{-5}	1×10^{-5}	$2 \times 10^{-5**}$
Observations	226	384	580
R^2	0.860	0.870	0.860
Adjusted R^2	0.825	0.841	0.831
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 5 Estimation Results of Econometric Model (5) for HMO Plans

7.2. Results for Hypothesis 3

Table 5 displays the estimation results of the econometric model (5) for HMO plans.[16] The results, estimated using two-way fixed effect estimators, confirm Hypothesis 3. Specifically, the parameter of interest, δ_{DID}, is statistically significant and positive in all three neighborhood specifications, indicating that the treatment group counties experienced a reduction in risk selection compared to the control group counties. Furthermore, the treatment effect is moderated by QBP star ratings, with $\beta_{Star(DID)} < 0$.

The estimation results of the other parameters also have the expected signs. Specifically, we find that MA health plans with higher star ratings tend to have worse risk selection, $\beta_{Star} < 0$. Besides, risk selection tends to be worsening for counties with increasing eligible Medicare populations and thus more flexibility in conducting risk selection, i.e., $\beta_{McPop} < 0$. Lastly, the amount of rebate available for MA health plans to provide supplementary benefits did not significantly impact risk selection during our study period, i.e., $\beta_{Rebate} \approx 0$. One reason is that the rebate amount in MA had stayed relatively the same throughout our study period (c.f. Guterman et al. (2018)).

As for the overall impact of the ACA shock on risk selection, we find that risk selection was mitigated in counties with low initial star ratings, but worsened in counties with high initial star ratings. Specifically, for counties with low initial star ratings (e.g., $Star_{c,12} = 3$), risk scores in the treatment group increase after the ACA shock, i.e., $\delta_{DID} + \beta_{Star(DID)} Star_{c,12} > 0$. However, for

[14] See Appendix A for details.

[15] We note that these initial star ratings in 2012 were mainly determined by MA health plans' performance in the year 2010, and thus are not affected by the phase-in group assignment published in 2011 (c.f. Centers for Medicare & Medicaid Services (2011)).

[16] Results for both HMO and PPO plans can be found in Appendix C.5.

counties with high initial star ratings (e.g., $Star_{c,12} = 4$), risk scores in the treatment group may even decrease after the ACA shock, i.e., $\delta_{DID} + \beta_{Star(DID)} Star_{c,12} < 0$. A possible explanation is that the additional capitation payments from the QBPs program after the ACA shock may help MA health plans conduct more risk selection. Therefore, for these high-quality plans, the negative impact of ACA on risk scores through the channel of QBPs outweighs the positive impact of ACA on risk scores through the channel of reduced strategic cross subsidization. These results highlight the necessity to examine the heterogeneous treatment effect, instead of the population average treatment effect, when studying the impact of the ACA shock on risk selection in MA.

8. Robustness Checks

This section checks the robustness of our findings. Specifically, §8.1 conducts a matching analysis. §8.2 examines the parallel trend assumption of our DID analysis. Additional robustness checks are presented in Appendix C.

8.1. A Matching analysis

To check whether the discontinuity in ACA capitation cuts discussed in §6.2.2 indeed provides such randomization, this section conducts a coarsened exact matching (CEM) on relevant county- and individual-level characteristics and re-estimate our main econometric model (3) using the matched sample.

First, we conduct a CEM using relevant county- and individual-level characteristics introduced in §5 and present the results in Tables 11 and 12 of Appendix B. Specifically, FFS $\text{Rate}_{c,2010}$, $RUCC_{c,2010}$ and $\text{Income}_{c,2010}$ may affect treatment assignments through MA benchmark payments. Here we focus on 2010 because the treatment and control groups were assigned based on the county-level benchmark difference in 2010. $\text{Star}_{c,2012}$ measures the balance between treatment and control groups in health plan quality and QBPs in 2012. We focus on the year 2012 because the star ratings in 2012 were mainly determined based on MA health plans' performance in 2010, the last year before the treatment and control assignments were announced. $\text{Age}_{i,t}$ is a proxy for the balance of the patient population's health in our study period. In addition, we re-estimate our main econometric model using the matched data and present the estimation results (Table 13 of Appendix B). The results based on matched data are consistent with our main results (Table 4).

8.2. Parallel Trend Assumption

A key assumption of our DID empirical strategy is that counties in treatment and control groups were otherwise similar except for their ACA capitation payment cuts. In other words, in a counter-factual world where the ACA capitation payment cuts were not introduced, counties in treatment and control groups would follow the same "parallel trends" so that the treatment effect estimate would not be statistically different between the treatment and control groups. In this section tests, we test this assumption using an event-study design, i.e.,

$$\begin{aligned}
&\text{MA Spending}_{i,t} - \text{Estimated Cost}_{i,t} \\
=&\beta_0 + \sum_{s=2009}^{2010} \left(\delta_{DID(s)} RiskScore_{i,t} \times Treatment_{i,s-2011} + \beta_s Treatment_{i,s-2011}\right) \\
&+ \sum_{s=2012}^{2015} \left(\delta_{DID(s)} RiskScore_{i,t} \times Treatment_{i,s-2011} + \beta_s Treatment_{i,s-2011}\right) \\
&+\beta_{RS} RiskScore_{i,t} + \alpha_i + \gamma_t + \epsilon_{i,t}. \qquad (6)
\end{aligned}$$

Specifically, if the parallel trend assumption holds, the parameter of interest $\delta_{DID(s)}$ should stay close to 0 in the pre-treatment period ($s = 2009, 2010$), which implies that the degree of strategic cross subsidization was similar between treatment and control groups before the ACA cut.[17]

Estimation results of the event-study model (6) are presented in Table 14 of Appendix B. Consistent with the identification assumption, we do not find a statistically significant difference in strategic cross subsidization practice between the treatment and control groups before the implementation of ACA cuts, i.e. $\delta_{DID(2009)} = \delta_{DID(2009)} = 0$. That is, MA health plans in both treatment and control groups have followed parallel trends in terms of their strategic cross subsidization practice. Their trends started diverging when the staggered rollout of ACA cuts began in 2012, i.e. $0 < \delta_{DID(2012)} < \delta_{DID(2013)} < \delta_{DID(2014)}$. When the staggered rollout is completed in 2015 for the local neighborhood studied in this paper, we find that the trends of treatment and control groups started converging again, i.e. $\delta_{DID(2015)} < \delta_{DID(2014)}$. We can also visually verify this implication of the parallel trend assumption through an event study plot of $\delta_{DID(s)}$ as in Figure 5 of Appendix B.

9. Conclusion

Healthcare capitation payment models, particularly the Medicare capitation program, have been extensively studied in recent decades. However, due to the limited access to MA health plan claims data, little is known about how MA health plans allocate capitation payments across various patient risk groups. More importantly, we have a limited understanding of whether and to what extent there may be perverse incentives at play in MA health plans' operations leading to a systematic mismatch between a patient's health status and the healthcare resources she gets from MA. Our

[17] Here we set 2011 as the base year because the phase-in group assignment was announced in February 2011 (c.f. Centers for Medicare & Medicaid Services (2011)).

paper made use of a large commercial insurance database containing medical claims from more than 2 million MA enrollees to open up the black box of MA health plans' operations and to study the allocation problem of MA capitation payments. Our findings shed new light on how MA health plans operate in serving different patient risk groups and how their operations can affect the risk selection problem in MA.

By opening the black box of MA health plans' operations, this paper also emphasizes the critical role data transparency can play in CMS's administration of MA. Notably, the law explicitly prohibits cross subsidization practices in MA,[18] with CMS holding a legal obligation to address these practices if they occur. Particularly, any overspending or underspending due to strategic cross subsidization practice is an unintended use of capitation payments from CMS's perspective. However, without proper transparency in MA claims data, it becomes challenging for CMS to fulfill its legal mandate. The recent legislation introduced in the U.S. House of Representatives, i.e., the *Medicare Advantage Consumer Protection and Transparency Act*, underscores the significance of data transparency in MA. As the legislator stated, enhancing MA data transparency "will empower watchdogs and lawmakers to verify that Medicare Advantage is working as well as it should be" (Committee for a Responsible Federal Budget 2022).

Our study calls for more future research in this direction. First, as population-based payment models are increasingly adopted by healthcare payers, strategic cross subsidization practice is likely to exist outside of the MA capitation program. Further investigations are needed to understand the role of strategic cross subsidization practice in other population-based payment models. Second, due to our lack of linked clinical data, we are unable to examine the impact of strategic cross subsidization on patient outcomes. We hope our initial investigation on strategic cross subsidization will trigger future research to examine its health outcome implications.

References

Adida, E., Mamani, H., and Nassiri, S. (2016). Bundled payment vs. fee-for-service: Impact of payment scheme on performance. *Management Science*, 63(5):1606–1624.

Agarwal, R., Connolly, J., Gupta, S., and Navathe, A. S. (2021). Comparing medicare advantage and traditional medicare: A systematic review: A systematic review compares medicare advantage and traditional medicare on key metrics including preventive care visits, hospital admissions, and emergency room visits. *Health Affairs*, 40(6):937–944.

Aizawa, N. and Kim, Y. S. (2018). Advertising and risk selection in health insurance markets. *American Economic Review*, 108(3):828–67.

[18] "Payments from any rate cell must not cross-subsidize or be cross-subsidized by payments for any other rate cell" (U.S. Congress 2016).

Ankuda, C. K., Ornstein, K. A., Covinsky, K. E., Bollens-Lund, E., Meier, D. E., and Kelley, A. S. (2020). Switching between medicare advantage and traditional medicare before and after the onset of functional disability: Measuring and characterizing enrollees who switch between medicare advantage and traditional medicare in the twelve months before and after onset of a functional disability. *Health Affairs*, 39(5):809–818.

Ata, B., Killaly, B. L., Olsen, T. L., and Parker, R. P. (2013). On hospice operations under medicare reimbursement policies. *Management Science*, 59(5):1027–1044.

Bastani, H., Goh, J., and Bayati, M. (2018). Evidence of upcoding in pay-for-performance programs. *Management Science*.

Biles, B., Pozen, J., and Guterman, S. (2009). The continuing cost of privatization: extra payments to medicare advantage plans jump to $11.4 billion in 2009. Available at https://hsrc.himmelfarb.gwu.edu/cgi/viewcontent.cgi?article=1163&context=sphhs_policy_facpubs (Accessed February 24, 2020).

Brennan, N., Ornstein, C., and Frakt, A. B. (2018). Time to release medicare advantage claims data. *Jama*, 319(10):975–976.

Brown, J., Duggan, M., Kuziemko, I., and Woolston, W. (2014). How does risk selection respond to risk adjustment? new evidence from the medicare advantage program. *American Economic Review*, 104(10):3335–64.

Centers for Medicare & Medicaid Services (2011). Advance Notice of Methodological Changes for Calendar Year (CY) 2012 for Medicare Advantage (MA) Capitation Rates, Part C and Part D Payment Policies and 2012 Call Letter. Available at https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Downloads/Advance2012.pdf (Accessed May 2, 2020).

Centers for Medicare & Medicaid Services (2015). Report to congress: Alternative payment models & medicare advantage. Available at https://www.cms.gov/Medicare/Medicare-Advantage/Plan-Payment/Downloads/Report-to-Congress-APMs-and-Medicare-Advantage.pdf (Accessed January, 9, 2020).

Centers for Medicare & Medicaid Services (2017). Risk Adjustment. Available at https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Risk-Adjustors.html (Accessed June 10, 2018).

Centers for Medicare & Medicaid Services (2019). Medicare Data. Available at https://www.cms.gov/medicare/medicare (Accessed May, 22, 2019).

Chen, C. and Savva, N. (2018). Unintended consequences of hospital regulation: The case of the hospital readmissions reduction program. *Available at SSRN 3236983*.

Committee for a Responsible Federal Budget (2022). Legislation to Bring Transparency to MA Introduced. Available at https://www.crfb.org/blogs/legislation-bring-transparency-ma-introduced (Accessed May 8, 2023).

Cooper, A. L. and Trivedi, A. N. (2012). Fitness memberships and favorable selection in medicare advantage plans. *New England Journal of Medicine*, 366(2):150–157.

Department of Health and Human Services, Office of Inspector General (2022). Some medicare advantage organization denials of prior authorization requests raise concerns about beneficiary access to medically necessary care. Available at https://oig.hhs.gov/oei/reports/OEI-09-18-00260.asp (Accessed November, 28, 2022).

Erhun, F., Mistry, B., Platchek, T., Milstein, A., Narayanan, V., and Kaplan, R. (2015). Time-driven activity-based costing of multivessel coronary artery bypass grafting across national boundaries to identify improvement opportunities: study protocol. *BMJ open*, 5(8):e008765.

Figueroa, J. F., Wadhera, R. K., Frakt, A. B., Fonarow, G. C., Heidenreich, P. A., Xu, H., Lytle, B., DeVore, A. D., Matsouaka, R., Yancy, C. W., et al. (2020). Quality of care and outcomes among medicare advantage vs fee-for-service medicare patients hospitalized with heart failure. *JAMA cardiology*, 5(12):1349–1357.

Fioretti, M., Wang, H., et al. (2019). Subsidizing inequality: Performance pay and risk selection in medicare. Technical report, Sciences Po Departement of Economics.

Flach, L. (2016). Quality upgrading and price heterogeneity: Evidence from brazilian exporters. *Journal of International Economics*, 102:282–290.

Freed, M., Biniek, J. F., Damico, A., and Neuman, T. (2022). Medicare advantage in 2022: Enrollment update and key trends. Available at https://www.kff.org/medicare/issue-brief/medicare-advantage-in-2022-enrollment-update-and-key-trends/ (Accessed September, 5, 2022).

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

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

Ginsburg, P. B. and Lieberman, S. M. (2022). The debate on overpayment in medicare advantage: Pulling it together. *Health Affairs Forefront* Available at https://www.healthaffairs.org/do/10.1377/forefront.20220223.736815/full/ (Accessed August, 4, 2022).

Glazer, J. and McGuire, T. G. (2000). Optimal risk adjustment in markets with adverse selection: an application to managed care. *The American Economic Review*, 90(4):1055–1071.

Guterman, S., Skopec, L., and Zuckerman, S. (2018). Do medicare advantage plans respond to payment changes? a look at the data from 2009 to 2014. *Issue Brief, March. Washington, DC: Commonwealth Fund. Accessed January*, 22:2020.

Halpern, M. T., Urato, M. P., and Kent, E. E. (2017). The health care experience of patients with cancer during the last year of life: analysis of the seer-cahps data set. *Cancer*, 123(2):336–344.

Han, T. and Lavetti, K. (2017). Does part d abet advantageous selection in medicare advantage? *Journal of health economics*, 56:368–382.

Huesch, M. D. (2010). Managing care? medicare managed care and patient use of cardiologists. *Health services research*, 45(2):329–354.

Hung, A., Stuart, B., and Harris, I. (2016). The effect of medicare advantage enrollment on mammographic screening. *The American journal of managed care*, 22(2):e53–9.

Ibanez, M. R., Clark, J. R., Huckman, R. S., and Staats, B. R. (2018). Discretionary task ordering: Queue management in radiological services. *Management Science*, 64(9):4389–4407.

Irarrazabal, A., Moxnes, A., and Opromolla, L. D. (2015). The tip of the iceberg: a quantitative framework for estimating trade costs. *Review of Economics and Statistics*, 97(4):777–792.

Jacobson, G., Casillas, G., Damico, A., Neuman, T., and Gold, M. (2016). Medicare advantage 2016 spotlight: Enrollment market update. Available at http://kff.org/medicare/issue-brief/medicare-advantage-2016-spotlight-enrollment-market-update/ (Accessed March, 4, 2017).

Jacobson, G., Freed, M., Damico, A., and Neuman, T. (2019a). A dozen facts about medicare advantage in 2019. Available at https://www.kff.org/medicare/issue-brief/a-dozen-facts-about-medicare-advantage-in-2019/ (Accessed January, 9, 2020).

Jacobson, G., Neuman, T., and Damico, A. (2019b). Do people who sign up for medicare advantage plans have lower medicare spending? Available at http://files.kff.org/attachment/Issue-Brief-Do-People-Who-Sign-Up-for-Medicare-Advantage-Plans-Have-Lower-Medicare-Spending (Accessed December, 15, 2019).

Jacobson, G. A., Neuman, P., and Damico, A. (2015). At least half of new medicare advantage enrollees had switched from traditional medicare during 2006–11. *Health Affairs*, 34(1):48–55.

James, B. C. and Poulsen, G. P. (2016). The case for capitation. *Harvard business review*, 94(7):19.

KC, D. S., Staats, B. R., Kouchaki, M., and Gino, F. (2020). Task selection and workload: A focus on completing easy tasks hurts performance. *Management Science*.

Kc, D. S. and Terwiesch, C. (2011). The effects of focus on performance: Evidence from california hospitals. *Management Science*, 57(11):1897–1912.

Keenan, P. S., Elliott, M. N., Cleary, P. D., Zaslavsky, A. M., and Landon, B. E. (2009). Quality assessments by sick and healthy beneficiaries in traditional medicare and medicare managed care. *Medical Care*, pages 882–888.

Lai, L. (2022). Budget debate: Healthcare clusters to get fixed sum for patients under care, instead of getting paid per visit. *The Straits Times* Available at https://www.straitstimes.com/singapore/politics/budget-debate-healthcare-clusters-to-get-fixed-sum (Accessed August 2, 2022).

Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of economic literature*, 48(2):281–355.

Li, Q., Trivedi, A. N., Galarraga, O., Chernew, M. E., Weiner, D. E., and Mor, V. (2018). Medicare advantage ratings and voluntary disenrollment among patients with end-stage renal disease. *Health Affairs*, 37(1):70–77.

Look, A. T.-Y. (2019). Did the affordable care act contain costs? *Health Affairs*, 2019.

Lu, S. F. and Lu, L. X. (2017). Do mandatory overtime laws improve quality? staffing decisions and operational flexibility of nursing homes. *Management Science*, 63(11):3566–3585.

Martino, S. C., Elliott, M. N., Haviland, A. M., Saliba, D., Burkhart, Q., and Kanouse, D. E. (2016). Comparing the health care experiences of medicare beneficiaries with and without depressive symptoms in medicare managed care versus fee-for-service. *Health services research*, 51(3):1002–1020.

Medicare Payment Advisory Commission (2019). Report to congress: Medicare payment policy. Available at https://www.medpac.gov/wp-content/uploads/import_data/scrape_files/docs/default-source/reports/mar19_medpac_entirereport_sec_rev.pdf (Accessed August, 5, 2022).

Medicare Payment Advisory Commission (2022). Report to congress: Medicare payment policy. Available at https://www.medpac.gov/wp-content/uploads/2022/03/Mar22_MedPAC_ReportToCongress_SEC.pdf (Accessed August, 5, 2022).

MedPage Today (2020). Medicare advantage enrollees discover dirty little secret — getting out is a lot harder than getting in. Available at https://www.medpagetoday.com/special-reports/exclusives/83661 (Accessed May 31, 2023).

Meyers, D. J., Belanger, E., Joyce, N., McHugh, J., Rahman, M., and Mor, V. (2019). Analysis of drivers of disenrollment and plan switching among medicare advantage beneficiaries. *JAMA internal medicine*, 179(4):524–532.

Meyers, D. J., Mor, V., and Rahman, M. (2018). Medicare advantage enrollees more likely to enter lower-quality nursing homes compared to fee-for-service enrollees. *Health Affairs*, 37(1):78–85.

Morrisey, M. A., Kilgore, M. L., Becker, D. J., Smith, W., and Delzell, E. (2013). Favorable selection, risk adjustment, and the medicare advantage program. *Health Services Research*, 48(3):1039–1056.

Newhouse, J. P., McWilliams, J. M., Price, M., Huang, J., Fireman, B., and Hsu, J. (2013). Do medicare advantage plans select enrollees in higher margin clinical categories? *Journal of health economics*, 32(6):1278–1288.

Patterson, B. W., Batt, R. J., Wilbanks, M. D., Otles, E., Westergaard, M. C., and Shah, M. N. (2016). Cherry picking patients: examining the interval between patient rooming and resident self-assignment. *Academic Emergency Medicine*, 23(6):679–684.

Rahman, M., Keohane, L., Trivedi, A. N., and Mor, V. (2015). High-cost patients had substantial rates of leaving medicare advantage and joining traditional medicare. *Health Affairs*, 34(10):1675–1681.

Rajkumar, R., Conway, P. H., and Tavenner, M. (2014). Cms engaging multiple payers in payment reform. *Jama*, 311(19):1967–1968.

Rice, N., Smith, P., et al. (1999). *Approaches to capitation and risk adjustment in health care: an international survey*. University of York The Centre for Health Economics.

Savva, N., Tezcan, T., and Yildiz, O. (2018). Can yardstick competition reduce waiting times? *Management Science*.

Schwartz, M. L., Kosar, C. M., Mroz, T. M., Kumar, A., and Rahman, M. (2019). Quality of home health agencies serving traditional medicare vs medicare advantage beneficiaries. *JAMA network open*, 2(9):e1910622–e1910622.

She, Z., Ayer, T., and Montanera, D. (2022). Can big data cure risk selection in healthcare capitation program? a game theoretical analysis. *Manufacturing & Service Operations Management*.

Skopec, L., Aarons, J., and Zuckerman, S. (2019a). Did medicare advantage payment cuts affect beneficiary access and affordability. *The American Journal of Managed Care*, 25(9):e261–e266.

Skopec, L., Ramos, C., and Aarons, J. (2019b). Are medicare advantage plans using new supplemental benefit flexibility to address enrollees' health-related social needs? Available at https://www.urban.org/sites/default/files/publication/101067/sdh_medicare_advantage_1.pdf (Accessed May 31, 2023).

Song, Z., Landrum, M. B., and Chernew, M. E. (2013). Competitive bidding in medicare advantage: Effect of benchmark changes on plan bids. *Journal of health economics*, 32(6):1301–1312.

United States Department of Agriculture (2013). 2013 Rural-Urban Continuum Codes. Available at https://www.ers.usda.gov/data-products/rural-urban-continuum-codes/ (Accessed June 10, 2018).

U.S. Congress (2008). 42 CFR § 422.306 - Annual MA capitation rates. Available at https://www.law.cornell.edu/cfr/text/42/422.306 (Accessed May, 23, 2021).

U.S. Congress (2016). 42 CFR § 438.4 - Actuarial soundness. Available at https://www.law.cornell.edu/cfr/text/42/438.4 (Accessed September, 8, 2022).

Verhoogen, E. A. (2008). Trade, quality upgrading, and wage inequality in the mexican manufacturing sector. *The Quarterly Journal of Economics*, 123(2):489–530.

Wallace, P. J., Shah, N. D., Dennen, T., Bleicher, P. A., and Crown, W. H. (2014). Optum labs: building a novel node in the learning health care system. *Health Affairs*, 33(7):1187–1194.

Appendix

A. ACA Reform of MA Capitation Payments

In §4.1, we introduced the capitation payment formula (1). We note that this capitation payment only includes those amounts paid by CMS for MA health plans to provide equivalent healthcare services as those in Traditional Medicare. [19] As discussed in §4.2, MA capitation payment began to phase into the ACA rates starting at 2012 and finished the phase-in process by 2017. Specifically, the Phase-in Group of each county was assigned based on the projected difference between ACA rates and the pre-ACA rates (Centers for Medicare & Medicaid Services 2011b). The first part of this section (A.1) explains how MA capitation payments were determined when the ACA rates were fully phased in. The second part of this section (A.2) shows how the projected difference between ACA and the pre-ACA rates was calculated, and discusses how MA capitation payments during the transition period (2012-2017) were determined. To better explain how changes in MA capitation payment would affect decisions of MA health plans and beneficiaries on a yearly basis, we have provided a reference for the calendar of key dates for MA contract applications at the end of this section (A.3).

A.1. MA Capitation Payments with Fully Phased-in ACA Benchmark (After 2017)

The ACA reform made two major changes to the formula of MA capitation rates (1). First, the benchmark payment would gradually decrease from 2012-2017. Second, MA health plans with higher quality would receive bonus capitation payments assigned by the QBPs program. This section provides a detailed explanation of how these ACA benchmarks and bonus capitation payments were determined.

The benchmark payments ($Benchmark_{c,t}$) in (1) were significantly reduced after 2012 when the ACA was taken into effect. Specifically, the pre-ACA benchmark payments were set equal to $\text{Max}\{Update_{c,t}, \text{FFS Rate}_{c,t}\}$, where $Update_{c,t}$ reflects the annual inflation in national per capita costs of Traditional Medicare, while $\text{FFS Rate}_{c,t}$ measures the local area per capita costs of Traditional Medicare (U.S. Congress 2008). In other words, the pre-ACA benchmark payments were often higher than the local area per capita costs of Traditional Medicare ($\text{FFS Rate}_{c,t}$). In contrast, ACA benchmark were based only on the local area per capita costs of Traditional Medicare ($\text{FFS Rate}_{c,t}$). Besides, benchmark payments after 2012 were further adjusted based on the quartile of local area per capita costs of Traditional Medicare in the national ranking. As shown in table 6,

[19] Particularly, it does not include premiums and rebates received by MA health plans. Premiums are paid by MA beneficiaries in order to join certain MA health plans. However, only less than 10% of the MA health plans charged premiums (Song et al. 2013). Rebates are paid by CMS for MA health plans to provide supplementary benefits, i.e. benefits not covered in Traditional Medicare, to their enrollees.

counties ranked in the lower quartile were assigned higher quartile adjustment ($quartile_c$).Hence, after adjusting for the quartile payments, the ACA benchmark rates were FFS $\text{Rate}_{c,t} \times quartile_c$.

FFS Quartile	Quartile Adjustment
4^{th} (highest)	95 Percent
3^{rd}	100 Percent
2^{nd}	107.5 Percent
1^{st} (lowest)	115 Percent

Table 6 **Medicare Advantage County Quartile Payment** ($quartile_c$)

The Quality Bonus Payments (QBPs) program was created to let MA health plans earn back parts of the reduction in ACA benchmark payments when these health plans achieved the targeted service quality. Specifically, CMS assigned each MA health plan a star rating based on their quality in a list of monitored services[20], and gave bonus capitation payments to these MA health plans according to these star ratings. In particular, MA health plans with higher star ratings would have higher benchmark payments (FFS $\text{Rate}_{c,t} \times bonus^h_{c,t}$) in (1). The exact bonus amount can be calculated based on Table 7. Furthermore, for a selected set of counties, they can receive a double bonus on their benchmark payments. For example, an MA health plan with star rating 3 in 2012 would have an additional bonus $bonus^h_{c,t} = 3\%$ ($bonus^h_{c,t} = 6\%$ if county c is a double-bonus county).

Star Rating	2012	2013	2014	Post-2014
5 Stars	5%	5%	5%	5%
4 or 4.5 Stars	4%	4%	5%	5%
3.5 Stars	3.5%	3.5%	3.5%	0
3 Stars	3%	3%	3%	0
Below 3 Stars	0	0	0	0

Table 7 **QBP Bonus Benchmark Payment** ($bonus^h_{c,t}$)

In summary, if the ACA capitation rates phased in immediately after 2012, the ACA benchmark payments for each MA health plan at county c would be

$$ACA\_Benchmark_{c,t} = \text{FFS Rate}_{c,t} \times \left(quartile_c + bonus^h_{c,t}\right). \tag{7}$$

Before taking bonus payments into account, FFS $\text{Rate}_{c,t} \times quartile_c$ were significantly lower than the pre-ACA benchmark payments ($\text{Max}\{Update_{c,t}, \text{FFS Rate}_{c,t}\}$). In 2015, the nationwide benchmark payment reduction was estimated to be 9.3% of the pre-ACA benchmark before adjusting for bonus payments. Through the QBPs program, MA health plans on average reduced these benchmark payment losses to 6.8% of the pre-ACA benchmark (Piper and Friedman 2016). As such, the

[20] The list of monitored services and their targeted quality can be found at Centers for Medicare & Medicaid Services (2020).

QBPs program indeed offered incentives for MA health plans to reduce their losses in benchmark payments through better service quality.

A.2. MA Capitation Payments during the Transition Period (2012-2017)

As discussed in §4, the ACA benchmark payment payments gradually phased into counties across U.S. during the 2012-2017 period. Specifically, the actual MA benchmark payments for each MA health plan at county c during the 2012-2017 period also depended on the pre-ACA benchmark ($\text{Max}\{Update_{c,t}, \text{FFS Rate}_{c,t}\}$) as well as the phase-in factor ($PhaseInFactor_{c,t}$) listed in Table 1, and were calculated as

$$\begin{aligned} Benchmark_{c,t} = &\text{Max}\{Update_{c,t}, \text{FFS Rate}_{c,t}\} \times (1 + bonus^{h}_{c,t}) \times (1 - PhaseInFactor_{c,t}) \\ &+ \text{FFS Rate}_{c,t} \times \left(quartile_c + bonus^{h}_{c,t}\right) \times PhaseInFactor_{c,t}, \end{aligned} \tag{8}$$

where $PhaseInFactor_{c,t} \in [0,1]$. Clearly, the actual MA benchmark payments during the 2012-2017 period was a weighted average of $\text{Max}\{Update_{c,t}, \text{FFS Rate}_{c,t}\} \times (1 + bonus^{h}_{c,t})$ and $ACA\_Benchmark_{c,t}$, where counties c with higher $PhaseInFactor_{c,t}$ would phase into the ACA benchmark earlier.

Furthermore, the ACA phase-in factor ($PhaseInFactor_{c,t}$) was determined based on the projected difference between pre-ACA and ACA benchmark rates. Specifically, this projected difference was calculated as

$$\begin{aligned} &0.5 \times [\text{2010 Rate(a) After Budget Neutrality Adjustment}] \\ &- 0.5 \times ([\text{Estimated 2010 FFS Rate}] - [\text{2010 IME Phase-out Dollar Amount}]) \\ &\times \left([\text{Quartile Percent}] + (\mathbb{1}_{[\text{Qualifying County}]=1} + 1) \times 0.015\right) \end{aligned}$$

(Centers for Medicare & Medicaid Services 2011b). All variables in this formula can be found in the risk2012.csv of Rate calculation data (ZIP) of Centers for Medicare & Medicaid Services (2011a). Counties with projected difference less than 30 were assigned to the Two Year Phase-in Group; counties with projected difference at least 30 and less than 50 were assigned to the Four Year Phase-in Group; counties with projected difference at least 50 were assigned to the Six Year Phase-in Group. The ACA phase-in factor ($PhaseInFactor_{c,t}$) of each county was then determined by its Phase-in Group as in Table 1.

A.3. A Calendar of Key Dates for MA Contract Applications

1. In the first week of April in year t, CMS announces the annual MA capitation rates and risk adjustment methods for contract year $t+1$ (Centers for Medicare & Medicaid Services 2021b).

2. On June 1 in year t, health plans submit their applications for MA contracts in year $t+1$. Each of these applications must include a Plan Benefit Package (PBP) that describes the coverage details and other supplementary materials (e.g. provider/pharmacy directories)
3. By August in year t, CMS completes the review for health plans' applications for MA contracts in year $t+1$.
4. By mid-September in year t, CMS signs MA contracts in year $t+1$ with health plans.
5. By October 15 in year t, MA health plans send out their plan benefit designs for year $t+1$ including provider/pharmacy directories to Medicare beneficiaries.
6. From October 15 to December 7 in year t, Medicare beneficiaries choose whether to enroll in an MA health plan or stay in Traditional Medicare in year $t+1$
7. Starting in January 1 of year $t+1$, CMS made advanced monthly capitation payments to MA health plans based on the monthly enrollment numbers of these plans.

This calendar of key dates, provided by Integrated Care Resource Center (2020), demonstrates the order of events in MA contract application. Specifically, in order to apply for MA contracts in year $t+1$, each MA health plan has to prepare a plan benefit package (PBP), outlining the coverage details and provider/pharmacy directories before June 1 in year t. Based on the information in PBP, MA beneficiaries make their enrollment decisions for year $t+1$ in the open enrollment period (typically between October and December in year t). Starting in January 1 of year $t+1$, CMS made predetermined monthly capitation payments to MA health plans, regardless of the number of patient visits or claims generated in these months.

Additionally, this calendar also demonstrates that MA capitation payment rates (1) are adjusted at a yearly level. Specifically, the risk scores ($RiskScore_{i,t}$ in (1)) for patient i on year t are calculated annually based on this patients' medical diagnoses documented in the previous year $t-1$, and do not change during year t.[21] In addition, the county-level benchmarks ($Benchmark_{c,t}$ in (1)) are also adjusted yearly. Particularly, the county-level benchmarks for the year t are announced by CMS "not later than the first Monday in April" of year $t-1$ (Centers for Medicare & Medicaid Services 2018).

[21] "The risk score is computed for each beneficiary for a given year and applied prospectively. The risk score generally follows the beneficiary for one calendar year" (Centers for Medicare & Medicaid Services 2021b).
See also "The CMS-HCC risk adjustment model is prospective—it uses a profile of major medical conditions in the base year, along with demographic information (age, sex, Medicaid dual eligibility, disability status), to predict Medicare expenditures in the next year"(Centers for Medicare & Medicaid Services 2018).

B. Supplementary Figures and Tables

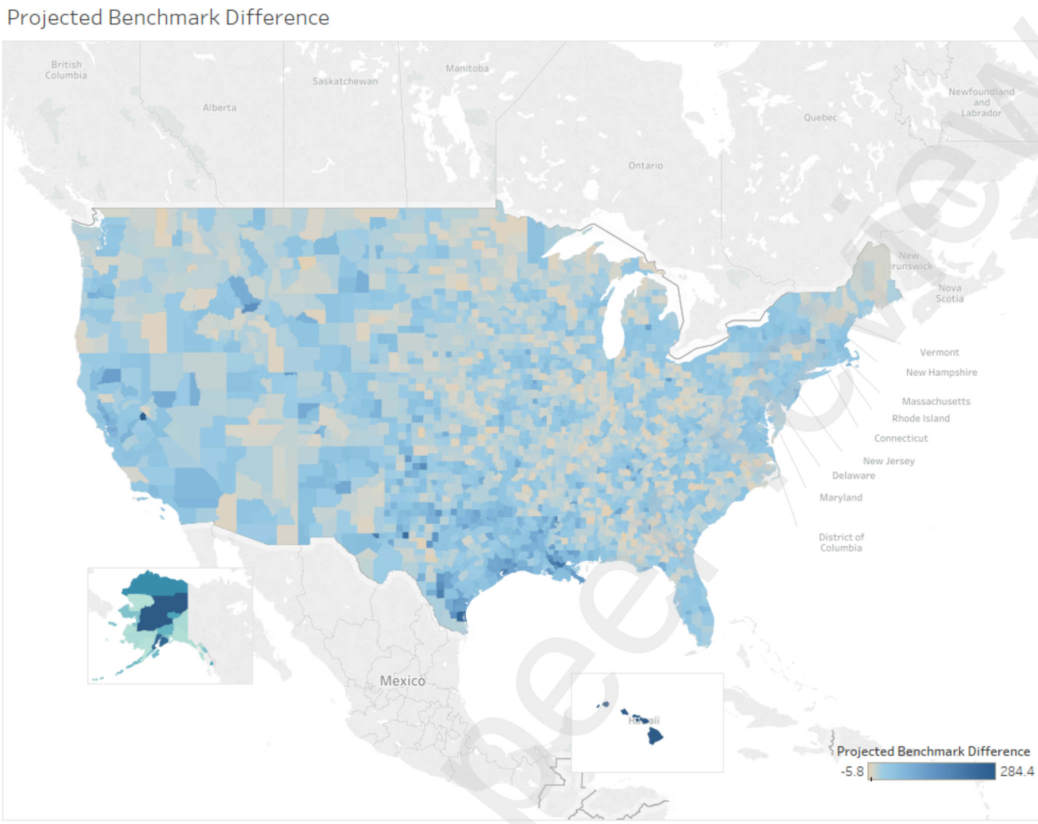

Figure 2 **The Projected Difference between the ACA and Pre-ACA Benchmark Payments Per Beneficiary Per Month (PBPM)**

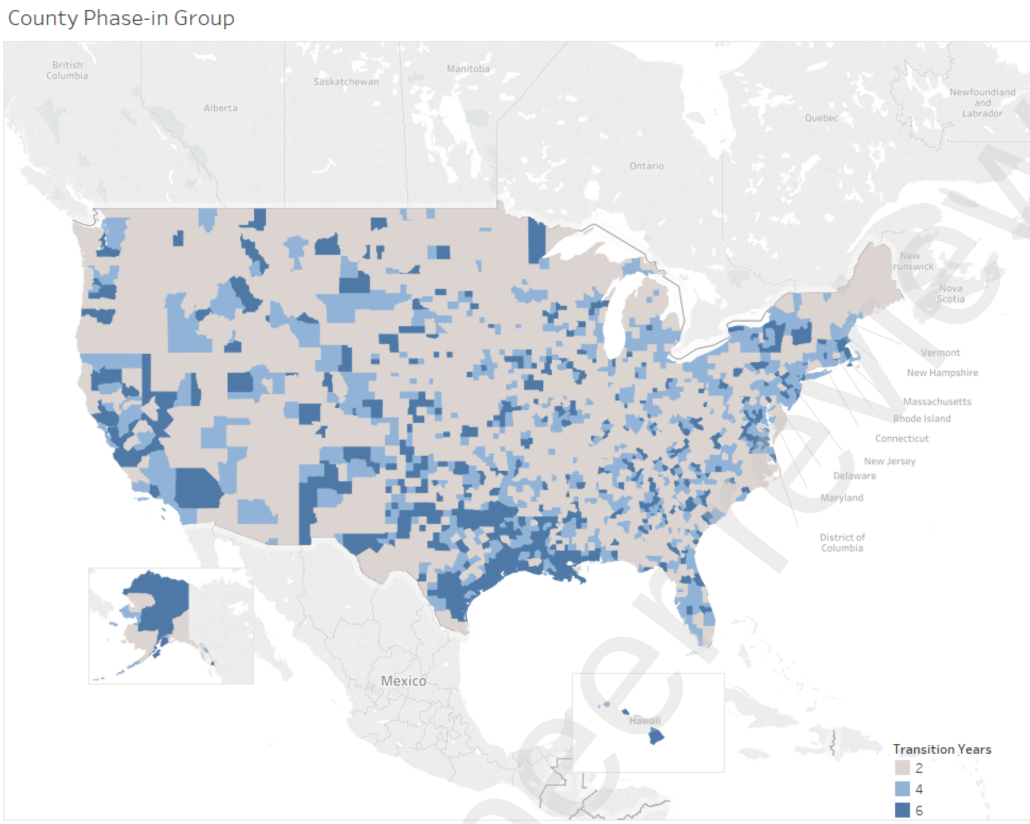

Figure 3 **Phase-in Group of Counties across the U.S.:** Counties in brown were in the Two Year Phase-in Group, who phased in the ACA benchmark by 2013. Counties in light blue were in the Four Year Phase-in Group, who phased in the ACA benchmark by 2015. Counties in dark blue were in the Six Year Phase-in Group, who phased in the ACA benchmark by 2017.

Group	Time	Outcomes	First Difference	Second Difference	Third Difference
	Before	$\beta_0 + \alpha_i$			
Control			$\beta_{RS} + \gamma_t$		
	After	$\beta_0 + \alpha_i + \beta_{RS} + \gamma_t$			
				$\beta_1 + \delta_{DID}$	δ_{DID}
	Before	$\beta_0 + \alpha_i$			
Treatment			$\beta_{RS} + \gamma_t + \beta_1 + \delta_{DID}$		
	After	$\beta_0 + \alpha_i + \beta_{RS} + \gamma_t + \beta_1 + \delta_{DID}$			

Table 8 **The Identification Strategy of Econometric Model (3)**

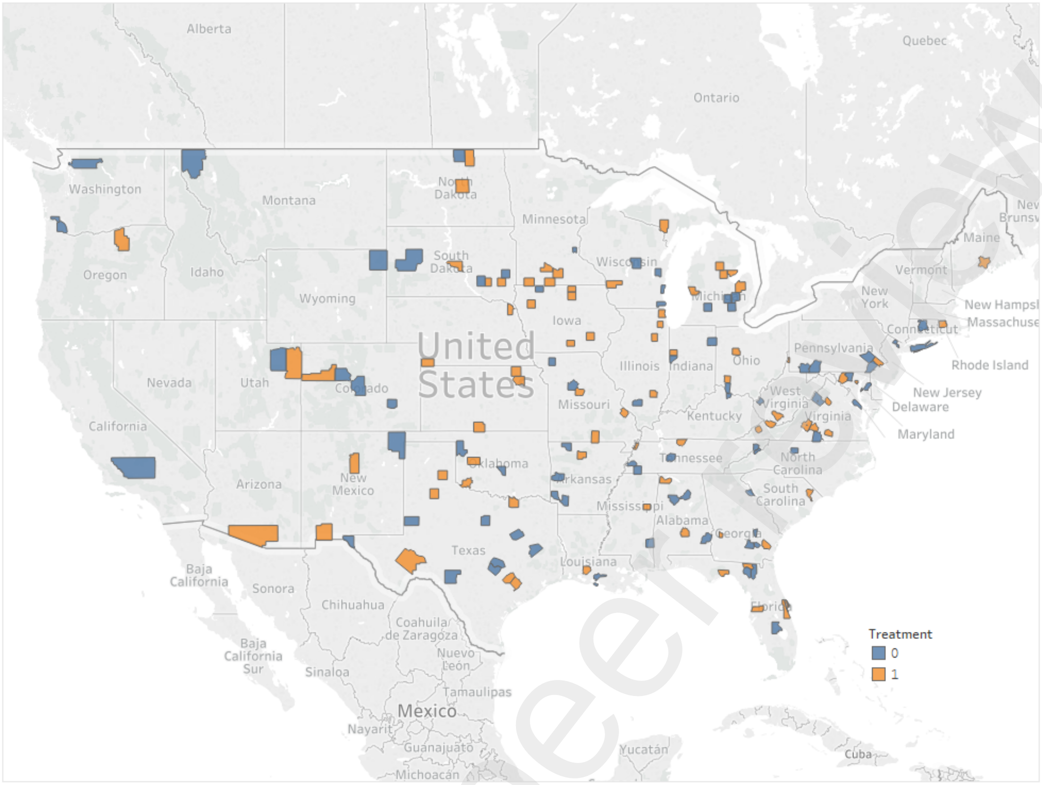

Figure 4 **Counties in the DID Analysis with Projected Benchmark Difference in** $[\$27, \$33]$**:** There were 155 of these counties in our data. All counties are in the control group before 2012. In 2012, 77 counties with projected benchmark difference in $[\$27, \$30)$ are assigned to the treatment group (Treatment=1), while the remaining 78 counties, with projected benchmark difference in $[\$30, \$33]$, stay in the control group (Treatment=0).

	Mean	2009	2010	2011	2012
County	HMO Enrollment	863	887	1,045	1,389
	HMO Star Rating	3.16	3.41	3.45	3.59
	PPO Enrollment	122	190	224	426
	PPO Star Rating	3.08	3.25	3.38	3.33
Individual	Risk Score	1.04	1.08	1.09	1.10
	Annual Change in Risk Score	NA	0.21	0.16	0.15
	MA Spending	$12,255	$11,439	$11,187	$11,302
	Annual Change in MA Spending	NA	$1,378	$1,249	$1,203

Table 9 **Summary Statistics of Year 2009 - 2012**

	Mean	2013	2014	2015
County	HMO Enrollment	1,450	1,376	1,510
	HMO Star Rating	3.71	3.74	3.78
	PPO Enrollment	550	683	733
	PPO Star Rating	3.42	3.71	4.08
Individual	Risk Score	1.11	1.12	1.14
	Annual Change in Risk Score	0.14	0.14	0.15
	MA Spending	$12,550	$12,121	$12,102
	Annual Change in MA Spending	$2,458	$1,172	$707

Table 10 **Summary Statistics of Year 2013 - 2015**

		FFS Rate$_{c,2010}$		RUCC$_{c,2010}$		Income$_{c,2010}$	
		Treatment	Control	Treatment	Control	Treatment	Control
Mean	All Data	$707.30	$724.24	1.75	1.73	$48,007	$50,936
	Matched Data	$684.85	$680.34	2.12	2.12	$43,964	$44,161
Percent Balance Improvement		73.4		100.0		93.3	

Table 11 **County- and Individual-level Characteristics of Treatment and Control Groups**

		Star$_{c,2012}$		Age$_{i,t}$	
		Treatment	Control	Treatment	Control
Mean	All Data	3.62	3.49	74.70	74.19
	Matched Data	3.62	3.65	74.61	74.59
Percent Balance Improvement		81.8		97.8	

Table 12 **County- and Individual-level Characteristics of Treatment and Control Groups (continued)**

	[$29,$31]	[$28,$32]	[$27,$33]
β_{RS}	-10.18***	-9.68***	-9.68***
δ_{DID}	1.53***	1.56***	1.56***
β_1	2.02***	-2.62***	-2.62***
Observations	279,358	341,983	355,553
R^2	0.465	0.468	0.469
Adjusted R^2	0.301	0.296	0.295

Note: Dependent variable is in $1,000 per year unit.
All standard errors are clustered at county level.
*p<0.1; **p<0.05; ***p<0.01

Table 13 Estimation Results of Econometric Model (3) using Matched Data

	[$29,$31]	[$28,$32]	[$27,$33]
β_{RS}	-10.27***	-9.76***	-9.66***
$\delta_{DID(2009)}$	0.01	-0.17	-0.19
$\delta_{DID(2010)}$	0.38	-0.11	-0.24
$\delta_{DID(2012)}$	1.16**	0.70*	0.60
$\delta_{DID(2013)}$	2.38***	1.91***	1.74***
$\delta_{DID(2014)}$	2.81***	2.28***	2.19***
$\delta_{DID(2015)}$	2.47***	1.93***	1.74***
β_{2009}	1.35	1.05*	1.03**
β_{2010}	0.23	0.60	0.89
β_{2012}	-1.38**	-1.36***	-1.13**
β_{2013}	-0.62	-0.95	-0.88
β_{2014}	-1.84*	-1.77***	-1.83***
β_{2015}	-2.75**	-2.92***	-2.93***
Observations	588,946	779,436	984,339
R^2	0.501	0.508	0.528
Adjusted R^2	0.311	0.310	0.323

Note: Dependent variable is in $1,000 per year unit.
All standard errors are clustered at county level.
*p<0.1; **p<0.05; ***p<0.01

Table 14 Estimation Results of Econometric Model (6)

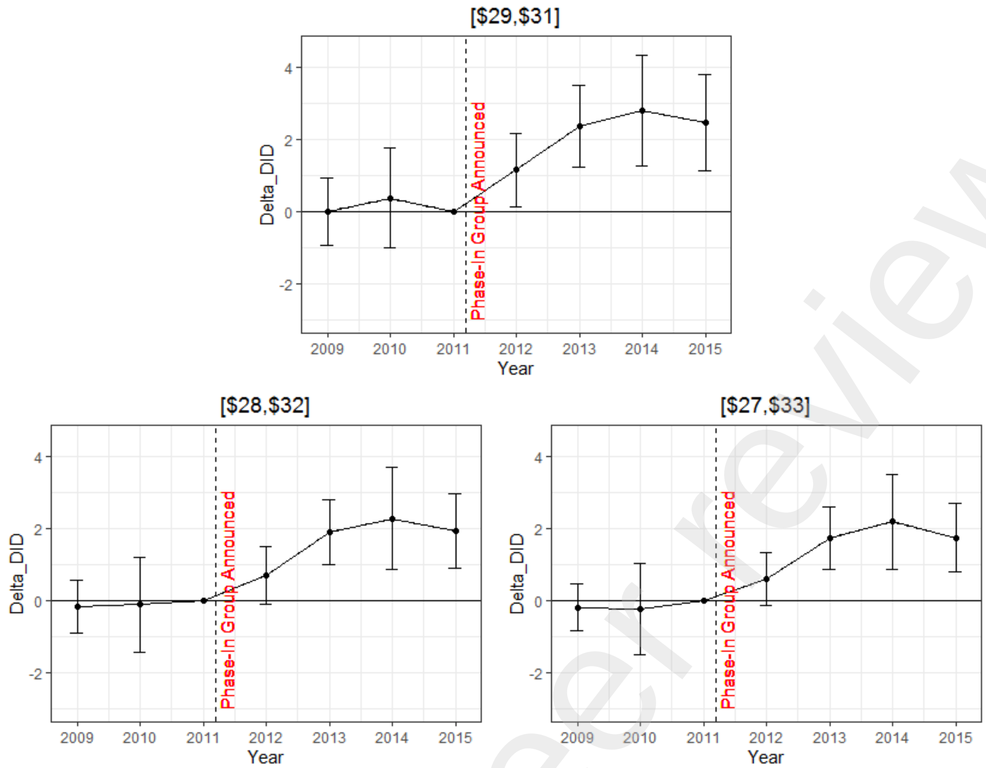

Figure 5 **Event Study Plots of $\delta_{DID(s)}$ in Econometric Model (6)**

	HIV/AIDS		Cancer		Diabetes		Cardiovascular	
	Before	After	Before	After	Before	After	Before	After
Spending-Cost Difference	$11,105.67	-$2,279.40	$7,731.53	$2,703.56	$5,187.45	$725.40	$8,357.58	$1,436.44
Risk Score	1.36	2.63	1.09	1.76	0.97	1.53	1.02	1.91

Table 15 **Case Study:** Changes in mean spending-cost differences and risk scores when MA enrollees developed HIV/AIDS (HCC1), Cancer (HCC7-HCC10)[22], Diabetes (HCC15 - HCC19)[23] or Cardiovascular diseases (HCC79-HCC83, HCC92)[24]

	IBD		Renal		Cerebrovascular		Hepatitis	
	Before	After	Before	After	Before	After	Before	After
Spending-Cost Difference	$9,393.79	$2,845.36	$5,843.59	-$686.18	$10,604.92	$2,482.19	$8,396.99	$2,033.09
Risk Score	1.27	1.84	1.18	2.05	1.37	2.19	1.28	1.91

Table 16 **Case Study (continued):** Changes in mean spending-cost differences and risk scores when MA enrollees developed Inflammatory Bowel Disease (IBD) (HCC33), Renal diseases (HCC130, HCC131)[25], Cerebrovascular diseases (HCC95, HCC96, HCC100, HCC101)[26] or Chronic Hepatitis (HCC27).

[22] Cancer conditions include Metastatic Cancer and Acute Leukemia (HCC7), Lung, Upper Digestive Tract, and Other Severe Cancers (HCC8), Lymphatic, Head and Neck, Brain, and Other Major Cancers (HCC9) and Breast, Prostate, Colorectal and Other Cancers and Tumors (HCC10).

[23] Diabetes conditions include Diabetes with Renal or Peripheral Circulatory Manifestation (HCC15), Diabetes with Neurologic or Other Specified Manifestation (HCC16), Diabetes with Acute Complications (HCC17), Diabetes with Ophthalmologic or Unspecified Manifestation (HCC18) and Diabetes without Complication (HCC19).

[24] Cardiovascular diseases include Cardio-Respiratory Failure and Shock (HCC79), Congestive Heart Failure (HCC80), Acute Myocardial Infarction (HCC81), Unstable Angina and Other Acute Ischemic Heart Disease (HCC82), Angina Pectoris/Old Myocardial Infraction (HCC83) and Specified Heart Arrhythmias (HCC92).

[25] Renal diseases include Dialysis Status (HCC130) and Renal Failure (HCC131).

[26] Cerebrovascular diseases include Cerebral Hemorrhage (HCC95), Ischemic or Unspecified Stroke (HCC96), Hemiplegia/Hemiparesis (HCC100) and Cerebral Palsy and Other Paralytic Syndromes (HCC101).

C. Additional Robustness Checks

This section conducts additional robustness checks to examine our findings. Specifically, we check whether results for the main hypothesis are robust against the panel attrition problem in §C.1 and measurement errors in the dependent variable in §C.2. In addition, we present several placebo tests to further examine our strategic cross subsidization hypothesis in §C.3. Lastly, we examine the potential time-series correlation problem in our DID estimation in §C.4.

C.1. Panel Attrition

A potential threat to our identification strategy is the panel attrition problem. Specifically, when testing our main hypothesis, we rely on the assumption that a representative sample of Medicare beneficiaries would continuously enroll in their MA health plans throughout the study period. Otherwise, there could be an alternative explanation for the estimation results in §6 as follows: if a significant number of sicker MA enrollees within each risk group dropped out from the sample during the study period, we would get the same estimation results as in §6 while the strategic cross subsidization problem in MA was getting worse instead of improving. In particular, since the CDM data do not allow users to link the geographic location and mortality information of patients, sicker MA enrollees within each risk group can drop out from the sample because they were deceased. In other words, if the main driver of the estimation results in §6 is panel attrition, the narrowing spending-cost difference we observed after the ACA shock does not necessarily imply mitigated strategic cross subsidization in MA.

To address this concern, we conduct a subsample analysis focusing on the continuously enrolled MA patients throughout the 2009-2015 study period. By focusing on the balanced panel data of this subsample, we can eliminate the potential estimation bias due to churning behavior or unrecorded death. The estimation results for this subsample are presented in Table 17. We note that these results are consistent with our main results in Table 4, i.e. $\beta RS < 0$ and $\delta DID > 0$. As such, this subsample analysis rules out the aforementioned alternative explanation of the observed mitigated strategic cross subsidization in MA.

C.2. Fat-Tailed Distributions in the Spending-Cost Difference

Another potential threat to our analysis arises from the fat-tailed distributions in spending-cost differences. It is well known that insurance claims commonly follow fat-tailed distributions, where the spending patterns can be significantly skewed by some large medical claims (Cooke et al. 2014). As such, one may be concerned that the empirical evidence of strategic cross subsidization in §6 could be mainly driven by several extreme medical claims, and thus was not representative of the MA market.

	[\$29,\$31]	[\$28,\$32]	[\$27,\$33]
β_{RS}	-8.52***	-8.25***	-8.00***
δ_{DID}	1.99***	1.77***	1.54***
β_1	-1.06	-1.39***	-1.10***
Observations	225,069	271,462	296,043
R^2	0.318	0.313	0.319
Adjusted R^2	0.220	0.213	0.218

Note: Dependent variable is in \$1,000 per year unit.
All standard errors are clustered at county level.
*p<0.1; **p<0.05; ***p<0.01

Table 17 Estimation Results of Econometric Model (3) for Continuously Enrolled MA Patients

To de-emphasize these potentially confounding fat tails in spending-cost differences, we re-estimate (3) our main hypothesis with log-transformed dependent variables. Specifically, we estimate the following counterpart of (3):

$$\begin{aligned} & Log[(\text{MA Spending}_{i,t} - \text{Estimated Cost}_{i,t})^*] \\ =& \beta_0 + \delta_{DID} RiskScore_{i,t} \times Treatment_{i,t} + \beta_1 Treatment_{i,t} \\ +& \beta_{RS} RiskScore_{i,t} + \alpha_i + \gamma_t + \epsilon_{i,t}, \end{aligned} \tag{9}$$

where the dependent variable is defined as

$$\begin{aligned} Log[(\text{MA Spending}_{i,t} - \text{Estimated Cost}_{i,t})^*] := Log[&\text{MA Spending}_{i,t} - \text{Estimated Cost}_{i,t} \\ &- \min_{i\in I, t\in T}\{\text{MA Spending}_{i,t} - \text{Estimated Cost}_{i,t}\} + 1]. \end{aligned}$$

This transformation normalizes the lowest spending-cost difference to 0 and applies log transformation to the normalized spending-cost difference, a common technique to deal with skewed data (c.f. Feng et al. (2013)). The estimation results for (9) is shown in Table 18. We can see that the parameters of interest, i.e. δ_{DID}, have the expected sign and remain statistically significant as those in Table 4. As such, these findings indicate that strategic cross subsidization is not confounded by the fat-tailed distributions of spending-cost differences.

C.3. Two Placebo Tests of the Strategic Cross Subsidization Mechanism

To test whether the observed effect, as in Table 4, is caused by the strategic cross subsidization mechanism, this section conducts two placebo tests. First, we restrict the DID regression (3) to a subsample of MA health plans that operated in both treatment and control counties. Second, we conduct the DID regression (3) in a subsample consisting of only high- or low-risk patients. In both tests, we should not observe significant results in placebo groups if the observed effect in Table 4 was driven by the strategic cross subsidization mechanism.

	[$29,$31]	[$28,$32]	[$27,$33]
β_{RS}	-0.018***	-0.018***	-0.017***
δ_{DID}	0.003***	0.002***	0.002***
β_1	-0.003***	-0.003***	-0.003***
Observations	588,946	779,436	984,339
R^2	0.491	0.498	0.511
Adjusted R^2	0.297	0.295	0.299

Note: All standard errors are clustered at county level.
*p<0.1; **p<0.05; ***p<0.01

Table 18 Estimation Results of Econometric Model (9)

As the first placebo test, we examine whether our DID estimation (3) is significant among MA health plans that were less likely to respond differentially to the ACA capitation payment cuts. We contend that MA health plans that operated in both treatment and control counties are ideal candidates for this placebo test. The reason is that it would be operationally difficult, if not impossible, for these health plans to restrict their strategic cross subsidization practice only in treatment counties while continuing such practice without restriction in control counties. Specifically, within an MA health plan, it is required by law that the benefit design cannot differentiate based on the county in which an MA enrollee resides.[27] Therefore, if an MA health plan operates in both treatment and control counties, it cannot reduce benefits for low-risk patients in treatment counties while keeping benefits for low-risk patients in control counties unchanged. As such, these MA health plans were less likely to respond differentially to the ACA capitation payment cuts based on treatment status and thus are assigned to the placebo group.

	[$29,$31]	[$28,$32]	[$27,$33]
β_{RS}	-10.70***	-10.12***	-12.97***
δ_{DID}	1.15	0.68	0.56
β_1	-1.96	-1.62**	-1.45**
Observations	418,912	571,724	736,340
R^2	0.536	0.534	0.545
Adjusted R^2	0.333	0.326	0.329

Note: Dependent variable is in $1,000 per year unit.
All standard errors are clustered at county and health plan levels.
*p<0.1; **p<0.05; ***p<0.01

Table 19 Estimation Results of Econometric Model (3) for MA Health Plans in the Placebo Group

[27] "All plan benefits must be offered uniformly to all enrollees residing in the service area of the plan" (Centers for Medicare & Medicaid Services 2021a).

	[\$29,\$31]	[\$28,\$32]	[\$27,\$33]
β_{RS}	-10.42***	-10.19***	-12.74***
δ_{DID}	2.42***	2.21***	2.14***
β_1	-2.33**	-2.76***	-3.25***
Observations	170,034	207,712	247,999
R^2	0.542	0.563	0.600
Adjusted R^2	0.330	0.340	0.377

Note: Dependent variable is in \$1,000 per year unit.
All standard errors are clustered at county and health plan levels.
*p<0.1; **p<0.05; ***p<0.01

Table 20 Estimation Results of Econometric Model (3) for MA Health Plans not in the Placebo Group

To operationalize the first placebo test, we re-estimate the econometric model (3) for MA health plans enrolling patients in both treatment and control counties separately. Particularly, we should not expect to find any statistically significant effect based on treatment status if strategic cross subsidization was the underlying mechanism. As shown in Table 19, we do not find any statistically significant treatment effect (δ_{DID}) among patients enrolled in these placebo MA health plans. In addition, among patients enrolled in the non-placebo MA health plans, Table 20 shows a statistically significant treatment effect (δ_{DID}), which is consistent with our main results in Table 4.

As the second placebo test, we examine whether our DID estimation (3) is significant among subpopulations consisting of only high- or low-risk patients. Again, if strategic cross subsidization was the underlying mechanism, we should not observe a statistically significant treatment effect (δ_{DID}) within these subpopulations. The reason is that the strategic cross subsidization hypothesis predicts that the spending-cost differences of the low-risk population would have a more significant decrease compared to that of the high-risk population. As such, without the presence of both populations in the sample, it is unlikely to detect this effect.

To implement the second placebo test, we re-estimate the econometric model (3) for low-risk MA patients, i.e., MA patients with risk scores below the first quartile (0.58), and high-risk MA patients, i.e., MA patients with risk scores above the third quartile (1.27), separately. As shown in Tables 21 and 22, we do not find any statistically significant change in spending-cost differences caused by RCS practice, i.e. $\delta_{DID} = 0$, which is consistent with our strategic cross subsidization hypothesis.

	[\$29,\$31]	[\$28,\$32]	[\$27,\$33]
β_{RS}	-16.07***	-14.65***	-14.94***
δ_{DID}	1.87	1.43	0.99
β_1	-0.28	-0.46	-0.41
Observations	156,347	213,415	273,771
R^2	0.700	0.727	0.755
Adjusted R^2	0.440	0.472	0.510
Note:	Dependent variable is in \$1,000 per year unit. All standard errors are clustered at county level. *p<0.1; **p<0.05; ***p<0.01		

Table 21 Estimation Results of the Main Econometric Model for Low-Risk MA Patients

	[\$29,\$31]	[\$28,\$32]	[\$27,\$33]
β_{RS}	-10.91***	-10.84***	-11.05***
δ_{DID}	0.11	-0.03	0.08
β_1	-2.98	-2.93**	-2.51**
Observations	149,301	193,411	241,970
R^2	0.651	0.650	0.665
Adjusted R^2	0.381	0.367	0.379
Note:	Dependent variable is in \$1,000 per year unit. All standard errors are clustered at county level. *p<0.1; **p<0.05; ***p<0.01		

Table 22 Estimation Results of the Main Econometric Model for High-Risk MA Patients

C.4. Time-Series Correlation in DID Estimation

One potential pitfall in DID estimation with many years of data, as noted by Bertrand et al. (2004), is the time-series correlation problem. That is, changes in the outcome variable can be just serially correlated, instead of being driven by the treatment variable.

To address this concern, we re-estimate the econometric model (3) with just 3 years of data, i.e. 2009, 2011, and 2013. This subsample consists of data from the base year, i.e. 2011, and one year of pretreatment and posttreatment, respectively, i.e. 2009 and 2013. That is, we estimate

$$\begin{aligned}
&\text{MA Spending}_{i,t} - \text{Estimated Cost}_{i,t}\\
=&\beta_0 + \delta_{DID(2009)} RiskScore_{i,2009} \times Treatment_{i,-2} + \beta_{2009} Treatment_{i,-2}\\
&+\delta_{DID(2013)} RiskScore_{i,2013} \times Treatment_{i,2} + \beta_{2013} Treatment_{i,2}\\
&+\beta_{RS} RiskScore_{i,t} + \alpha_i + \gamma_t + \epsilon_{i,t}.
\end{aligned} \tag{10}$$

This "ignoring-time-series-information" approach (c.f. Bertrand et al. (2004)) avoids time-series correlation in the pretreatment and posttreatment periods, and is commonly used by empirical researchers (e.g. Verhoogen (2008)). Estimation results, as presented in Table 23, are consistent with our main hypothesis, i.e. $\beta_{RS} < 0$ and $\delta_{DID(2013)} > 0$.

	[\$29,\$31]	[\$28,\$32]	[\$27,\$33]
β_{RS}	-7.324***	-6.975***	-6.619***
$\delta_{DID(2009)}$	0.952	0.918	0.694
$\delta_{DID(2013)}$	2.110***	1.907***	1.558***
β_{2009}	-1.338	-0.761	-0.372
β_{2013}	-0.885	-1.037	-0.457
Observations	220,149	287,283	360,087
R^2	0.686	0.689	0.704
Adjusted R^2	0.338	0.318	0.323
Note:	Dependent variable is in \$1,000 per year unit. All standard errors are clustered at county level. *p<0.1; **p<0.05; ***p<0.01		

Table 23 Estimation Results of Econometric Model (10)

C.5. An Analysis of both HMO and PPO plans for Hypothesis 3

	[\$29,\$31]	[\$28,\$32]	[\$27,\$33]
δ_{DID}	0.260***	0.133***	0.093**
$\beta_{Star(DID)}$	-0.068***	-0.039***	-0.027**
$\beta_{Star(Post)}$	-0.013	-0.006	-0.012*
β_{Star}	-0.012	-0.005	-0.004
β_{McPop}	0.003	0.002	0.003
β_{Rebate}	2×10^{-5}	-0×10^{-5}	1×10^{-5}
Observations	476	825	1,275
R^2	0.842	0.851	0.843
Adjusted R^2	0.805	0.817	0.809
Note:	*p<0.1; **p<0.05; ***p<0.01		

Table 24 Estimation Results of Econometric Model (5) for both HMO and PPO plans

To complement the results presented in Table 5, this section presents an analysis of both HMO and PPO plans for Hypothesis 3. Specifically, since CMS publishes its data for each type of health plan separately, we first pool data from HMO and PPO together and then estimate the econometric model (5) using the weighted least square estimator. Here, each observation is weighted by its county-level enrollment fraction, defined as the HMO (or PPO) MA enrollment number in a county divided by the total HMO and PPO MA enrollment number of this county. The estimation results, as presented in Table 24, are consistent with Hypothesis 3, i.e. $\delta_{DID} > 0$ and $\beta_{Star(DID)} < 0$.

Appendix Reference

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly journal of economics*, 119(1):249–275.

Centers for Medicare & Medicaid Services (2011a). 2012 Medicare Advantage ratebook and Prescription Drug rate information. Available at https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Ratebooks-and-Supporting-Data.html (Accessed May 5, 2020).

Centers for Medicare & Medicaid Services (2011b). Advance Notice of Methodological Changes for Calendar Year (CY) 2012 for Medicare Advantage (MA) Capitation Rates, Part C and Part D Payment Policies and 2012 Call Letter. Available at https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Downloads/Advance2012.pdf (Accessed May 2, 2020).

Centers for Medicare & Medicaid Services (2018). Report to congress: Risk adjustment in medicare advantage. Available at https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Downloads/RTC-Dec2018.pdf (Accessed September, 19, 2021).

Centers for Medicare & Medicaid Services (2020). Part C and D Performance Data. Available at https://www.cms.gov/Medicare/Prescription-Drug-Coverage/PrescriptionDrugCovGenIn/PerformanceData (Accessed May 22, 2108).

Centers for Medicare & Medicaid Services (2021a). Medicare managed care manual: Chapter 4 - benefits and beneficiary protections organizations. Available at https://www.cms.gov/Regulations-and-Guidance/Guidance/Manuals/Downloads/mc86c04.pdf (Accessed September, 19, 2021).

Centers for Medicare & Medicaid Services (2021b). Medicare managed care manual: Chapter 8 - payments to medicare advantage organizations. Available at https://www.cms.gov/Regulations-and-Guidance/Guidance/Manuals/downloads/mc86c08.pdf (Accessed September, 19, 2021).

Cooke, R. M., Nieboer, D., and Misiewicz, J. (2014). *Fat-tailed distributions: Data, diagnostics and dependence*, volume 1. John Wiley & Sons.

Feng, C., Wang, H., Lu, N., and Tu, X. M. (2013). Log transformation: application and interpretation in biomedical research. *Statistics in medicine*, 32(2):230–239.

Integrated Care Resource Center (2020). Key 2020 Medicare Dates. Available at https://www.integratedcareresourcecenter.com/sites/default/files/ICRC_Key_2020_Medicare_Dates.pdf (Accessed September, 27, 2021).

Piper, B. J. and Friedman, J. M. (2016). Bonus: The ACA reduced Medicare Advantage benchmark payment rates ... how much have medicare advantage organizations earned back through quality bonus payments? Available at http://www.theactuarymagazine.org/bonus/ (Accessed May, 21, 2018).

Song, Z., Landrum, M. B., and Chernew, M. E. (2013). Competitive bidding in medicare advantage: Effect of benchmark changes on plan bids. *Journal of health economics*, 32(6):1301–1312.

U.S. Congress (2008). 42 CFR § 422.306 - Annual MA capitation rates. Available at https://www.law.cornell.edu/cfr/text/42/422.306 (Accessed May, 23, 2021).

Verhoogen, E. A. (2008). Trade, quality upgrading, and wage inequality in the mexican manufacturing sector. *The Quarterly Journal of Economics*, 123(2):489–530.