

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

Zhaowei She

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

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

Capitation payment models have been increasingly adopted by the payers in the U.S. healthcare market during the past decade. However, healthcare services provided in Medicare Advantage (MA), the largest capitation program in the U.S., have been suggested to be more appealing to healthier patients and less appealing to sicker patients. The mismatch between a patient's health status and the benefits she gets from MA suggests that there may be a misallocation problem in MA. Despite extensive research on Medicare capitation program, little is known about how MA health plans actually allocate these capitation payments to different patients due to limited access to MA health plans' claims data. This paper utilizes a large commercial insurance database containing claims from more than 2 million MA enrollees to study the allocation problem of MA capitation payments. We empirically demonstrate that MA inadvertently incentivizes MA health plans to reallocate parts of the capitation payments from the sick to cross subsidize the healthy. By exploiting an exogenous policy shock on MA capitation payments through a Difference-in-Difference (DID) design, we identify, the first time in the literature, this reverse cross subsidization practice. Furthermore, we show that the reverse 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 the high-risk patients.

1. Introduction

Capitation has been highlighted as one of the most promising healthcare payment models to contain costs and increase efficiency in the U.S. health care system ([James and Poulsen 2016](#)). 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 Medicare, the largest healthcare insurance program in the U.S. 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 10.5 million enrollees in

2009 to 22 million in 2019, corresponding to more than one third of the total Medicare population (Jacobson et al. 2019a). Furthermore, the Congressional Budget Office (CBO) projects that by 2029 nearly half of the total Medicare population will enroll in MA (Jacobson et al. 2019a).

While CMS is the payer in both Traditional Medicare (TM) and Medicare Advantage (MA), it plays very different roles in these two programs from an operations perspective. Specifically, CMS is both a payer and a manager of service operations in TM (Medicare Payment Advisory Commission 2019). In contrast, CMS is a pure payer in MA where it completely outsources the provision and management of healthcare services for its Medicare Advantage enrolled patients to private MA health plans. In MA, 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 overall health of their enrollees.¹

The allocation of capitation payments and healthcare resources in MA are 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 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 services that are appealing to healthier Medicare beneficiaries, such as gym memberships (Cooper and Trivedi 2012). 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,

¹ 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.

is there a systematic mismatch between a patient's health status and the healthcare benefits 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 the healthcare resource allocation problem in MA. 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. We hypothesize that there is a systematic difference between MA health plans' actual healthcare spending and the estimated healthcare cost for patients in different risk groups.

In addition, we empirically examine the cause of the spending-cost differences in MA. Specifically, we hypothesize that MA health plans strategically overspend on low-risk patients at the cost of high-risk patients. In fact, even though MA capitation payments are adjusted based on enrollees' health status², it has been theoretically shown that low-risk patients might be more profitable for MA health plans to enroll compared to high-risk patients (Glazer and McGuire 2000, She et al. 2020). Subsequently, MA health plans might be better off overspending to attract low-risk patients and underspending to deter high-risk patients, which we empirically study in this paper.

The main empirical challenge of this study is to correctly identify the overspending and underspending behaviors of MA health plans from the claims data. Specifically, for an observed difference between MA health plans' actual spending on a patient and the estimated healthcare cost of this patient, there may be alternative explanations besides the overspending and underspending behaviors of MA health plans. For example, such difference may be due to overestimation or underestimation of patients' healthcare cost (Brown et al. 2014).³ To address such alternative explanations, we exploit a policy shock from the Affordable Care Act (ACA) which reduced MA capitation payments over a period of time in a staggered manner. This quasi-natural experiment enables us to empirically identify the overspending and underspending behaviors of MA health

² See Social Security Act §1853(a)(1)(B)(ii): "The Secretary shall adjust the payment amount (of fixed monthly payments to Medicare Advantage insurers) for such risk factors as age, disability status, gender, institutional status, and such other factors as the Secretary determines to be appropriate, including adjustment for health status . . . , so as to ensure actuarial equivalence." (U.S. Congress 1999)

³ In fact, since the explanatory power of the current CMS-HCC risk adjustment formula is weak ($R^2 = 11\%$), it has been empirically shown that this risk adjustment formula often over-estimate or under-estimate the healthcare costs of patients in MA (Brown et al. 2014).

plans through a Difference-in-Difference (DID) design, which provides the first empirical evidence that MA health plans reallocate parts of the capitation payments from sicker patients to cross subsidize the healthier patients in MA, a practice which we refer to as *reverse cross subsidization*.

To explore the implications of reverse cross subsidization practice on the MA capitation program, we further investigate a possible connection between reverse cross subsidization and the risk selection problem in MA. Specifically, while explicit patient selection is not legal in MA, MA health plans can still over-provide certain services to attract healthier patients and under-provide other service 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, 2019b). This patient self-selection problem in MA, which is commonly referred to as risk selection, has been regarded as “the central problem that inhibits the smooth, efficient functioning” of this market (Geruso and Layton 2017). Existing literature mostly attribute the risk selection in MA to the *miscalculation problem* of MA capitation payments (Brown et al. 2014, Glazer and McGuire 2000). That is, CMS would overpay (or underpay) MA health plans to enroll certain patients when the CMS-HCC risk adjustment formula over-estimates (or under-estimates) the healthcare costs of these patients. However, this body of literature has not recognized the possibility that risk selection in MA can also be connected to the *misallocation problem* of MA capitation payments. In other words, even if CMS estimates the healthcare cost of each MA patient correctly and does not miscalculate their capitation payments, MA health plans might still strategically overspend on the low-risk patients and underspend on the high-risk patients, which would also result in risk selection in MA. In this paper, we empirically test this possibility, and show that the reverse cross subsidization practice is associated with risk selection in MA.

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 reverse 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 or hospitals (Ibanez et al. 2018, Kc and Terwiesch 2011). Our work joins a small number of studies employing nationwide data (Lu and Lu 2017) and investigates a broad but unrecognized resource misallocation problem in health plans. In addition, except for a recent theoretical paper (c.f. She et al. (2020)), reverse cross subsidization practice, which is a key focus of this paper, has not been examined in the extant literature. Our work is a first attempt to distinguish its presence with patient level data. Second, we empirically document reverse cross

subsidization as an important driver of risk selection in Medicare Advantage. Currently, CMS has not taken any measures to limit reverse cross subsidization practice in MA, which leaves sick patients vulnerable to risk selection induced by reverse cross subsidization. Our work highlights the need for CMS to consider restricting reverse cross subsidization in MA, rather than the current practice of improving risk adjustment, when addressing the risk selection problem.

The remainder of this paper proceeds as follows. In §2, we review the relevant literature and develop our research hypotheses. In §3, 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 §4, we introduce our data sources and construct the variables used in our econometric models. In §5, we develop the econometric model to test our main hypothesis. In §6, we empirically assess the impact of MA health plans' reverse cross subsidization practice on the risk selection problem in MA. In §7, we conduct various robustness checks to examine our main findings. We summarize our findings and conclude in §8.

2. Literature Review and Hypothesis

In this section, we first review the related healthcare management literature on the MA capitation program and risk selection. Subsequently, we develop our research hypotheses to examine the potential reverse cross subsidization practice in MA.

2.1. Literature Review

Capitation payment models have long been one of the most debated issues in healthcare payment reforms. On one hand, healthcare payers, for example the 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 discrepancy between the expectations and the actual performance of capitation programs in practice. Indeed several studies noted that healthcare services provided in MA tend to be more appealing to healthier Medicare beneficiaries (e.g. gym membership) and less appealing to sicker Medicare beneficiaries (e.g. low-quality nursing home care) (Li et al. 2018, Cooper and Trivedi 2012, Meyers et al. 2018). 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 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.

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 practices. 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”. 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. \(2020\)](#) 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. While [She et al. \(2020\)](#) offers 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 reverse cross subsidization in MA, and explains how it is associated with inequality of healthcare access and risk selection in MA.

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 subsidisation 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 on healthcare payment models based on past cost estimates (c.f. [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 to 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 observation beds were strategically utilized by hospitals 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 identify the reverse cross subsidization behaviors of MA health plans as a strategic response to the MA capitation payment model.

2.2. 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 patient. 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 association between an MA enrollee’s health status and the healthcare benefits she gets from MA. Specifically, we want to understand whether there is a systematic mismatch between MA health plans’ actual healthcare spending and the estimated healthcare costs across patient risk groups. We hypothesize that as an MA patient became sicker, her MA health plan would spend less on her healthcare services compared to her estimated healthcare cost. This hypothesis is motivated by the findings in the medical literature that MA health plans tend to provide better preventive care (c.f. [Cooper and Trivedi \(2012\)](#), [Hung et al. \(2016\)](#)) but worse nursing home care or postacute care (c.f. [Meyers et al. \(2018\)](#), [Rahman et al. \(2015\)](#)). As an MA patient became sicker, she would utilize less preventive care but more nursing home care or postacute care. Therefore, we hypothesize that

Hypothesis 1: *There is a negative association 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 to which we refer as reverse cross subsidization. To this end, we examine whether the negative association studied in Hypothesis 1 is caused by the reverse cross subsidization practice of MA health plans. Specifically, if this causal relationship is established, it would constitute evidence of a potential misallocation problem in MA, i.e. MA health plans strategically overspend on low-risk patients and underspend on high-risk patients in MA. Hence, our main hypothesis is as follows:

Main Hypothesis (Hypothesis 2): *There exists reverse cross subsidization practice in MA, which causes the negative association characterized in Hypothesis 1.*

Notice that, the negative association put forward in Hypothesis 1 can also be generated by other alternative mechanisms. For example, [Brown et al. \(2014\)](#) documented that the healthcare costs of high-risk MA patients are often systematically over-estimated by the CMS-HCC risk adjustment formula, compared to those of low-risk MA patients. Consequently, an alternative hypothesis would be that the negative association put forward in Hypothesis 1 is caused by the systematic errors in healthcare cost estimations. The main goal of our paper is therefore to disentangle our main hypothesis, i.e. reverse cross subsidization, from these alternative hypotheses.

Lastly, a natural implication of the reverse cross subsidization practice of MA health plans would be risk selection. On one hand, through reverse cross subsidization, MA health plans would be able to provide greater benefits to satisfy the needs of the healthier Medicare beneficiaries, e.g. free fitness membership (c.f. [Cooper and Trivedi \(2012\)](#)). On the other hand, as MA health plans reallocate parts of the capitation payments from the sicker Medicare beneficiaries to spend on healthier Medicare beneficiaries, sicker patients would get less healthcare benefits in MA, e.g. lower quality long-term and postacute care (c.f. [Meyers et al. \(2018\)](#), [Rahman et al. \(2015\)](#)). As a result, low-risk patients would be more likely to enroll in MA compared to the high-risk patients. This potential impact of the reverse cross subsidization practice on MA is summarized in the following hypothesis :

Hypothesis 3: *The reverse cross subsidization practice of MA health plans is associated with the risk selection problem in MA.*

3. Institutional Background of Medicare Advantage (MA)

The first part of this section provides a general introduction of the Medicare Advantage program, with emphasis on its capitation payment model (§3.1). The second part discusses the Affordable Care Act (ACA) as a shock to this capitation payment model in MA (§3.2).

3.1. Medicare Advantage and Its Capitation Payment Model

Medicare Advantage (Medicare Part C) is an insurance program which 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 of Medicare beneficiaries have enrolled in Medicare Advantage, which makes MA the largest capitation program in the U.S. ([Jacobson et al. 2019a](#)).

A main responsibility of CMS, as a payer in MA, is to use capitation payments to incentivize MA health plans to deliver similar healthcare services to MA enrollees as those in Traditional Medicare.⁴ Specifically, the Social Security Act requires CMS to set the MA capitation rates so that the medical benefits delivered in MA are actuarially equivalent to those in Traditional Medicare (Medicare Part A and Part B).⁵ 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

$$CapitationRate_{i,t} = RiskScore_{i,t} \times Benchmark_{c,t}. \quad (1)$$

Other details of MA capitation payments are provided in [Appendix A](#).

3.2. 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

⁴ [UnitedHealthcare Ins. Co. v. Azar](#), Civil Action No. 16-157 (RMC) (D.D.C. Jan. 27, 2020): “A Medicare Advantage insurer must provide, at a minimum, the same level of benefits provided by traditional Medicare itself. See 42 U.S.C. § 1395w-22(a).” ([U.S. District Court 2020](#))

⁵ See Social Security Act §1853(a)(1)(B)(ii): “The Secretary shall adjust the payment amount (of fixed monthly payments to Medicare Advantage insurers) for such risk factors as age, disability status, gender, institutional status, and such other factors as the Secretary determines to be appropriate, including adjustment for health status . . . , so as to ensure actuarial equivalence.” ([U.S. Congress 1999](#))

See also [UnitedHealthcare Ins. Co. v. Azar](#), 330 F. Supp. 3d 173 (D.D.C. 2018): “The Social Security Act’s “actuarial equivalence” requirement mandates the ‘application of actuarial principles’ to ensure that MAOs are paid ‘a sum equal to the cost that CMS would expect to bear in providing traditional Medicare.’” ([U.S. District Court 2018](#))

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 per beneficiary per month (PBPM) in some counties to around \$300 PBPM in other counties. In other words, some counties have disproportionately large 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 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 $t \in \{2012, \dots, 2017\}$ 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 was 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

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”.

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 (Centers for Medicare & Medicaid Services 2011c, 2018d).

Year	Two Year Phase-in Group		Four Year Phase-in Group		Six Year Phase-in Group	
	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 group. 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 2011b](#)).

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 cuts less than \$30 per beneficiary per month (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 period⁸, and thus smooths the transition process. In §5.2.2, we describe in detail how we exploit this Phase-in Group assignment rule to implement our DID design.

4. 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 2009-2012. This study period contains years before (2009-2011) and after (2012) the ACA capitation payment cuts in which the same CMS-HCC risk adjustment formula was used in MA.⁹ 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. §4.1 and §4.2 introduce these two data sources and explain construction of the relevant variables from these datasets.

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 The CMS-HCC risk adjustment formula was updated in 2009, 2013, 2014 and 2017 ([Centers for Medicare & Medicaid Services 2018](#)).

4.1. Individual-level Data

The individual-level data used in this study comes from CDM, a commercial insurance claims database which 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 2012, 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.¹⁰ As the healthcare cost estimate of an MA enrollee depends on the county in which this enrollee lived, i.e. Estimated Cost $_{i,t} := RiskScore_{i,t} \times FFS\ Rate_{c,t}$, it is not possible to make 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.

The final sample contains MA claims data from 2,601,298 distinct Medicare beneficiaries living in 2,789 counties or county equivalent in 50 states and DC in the U.S. across 4 years (2009-2012). The individual-level dataset is used to construct variables to test Hypotheses 1 and 2. These variables include:

- $RiskScore_{i,t}$: 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, 2011d). The CDM provides these individual characteristics of Medicare beneficiary i to calculate his/her risk score.
- MA Spending $_{i,t}$: the total amount an MA health plan spent on a 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 .¹¹
- FFS Rate $_{c,t}$: 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 2009-2012 county-level Traditional Medicare data publicly available on the CMS's website (c.f. Centers for Medicare & Medicaid Services (2018a))

¹⁰ These "movers" accounted for about 8% of the total Medicare beneficiaries in the study sample.

¹¹ 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.

- Estimated Cost $_{i,t}$: 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 Cost $_{i,t} = RiskScore_{i,t} \times FFS Rate_{c,t}$ based on the capitation payment formula (1), with the MA benchmark payment $Benchmark_{c,t}$ replaced by the TM benchmark spending FFS Rate $_{c,t}$.
- Treatment $_{i,t}$: 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 §5.2.2.

4.2. 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 \(2018c,b,d,c,a, 2020\)](#). We focus on the analysis of non-special needs Health Maintenance Organizations (HMOs) in MA, as it was the dominant health plan type in this market ([Song et al. 2013](#)). To reduce sample variance, we exclude those counties with less than 100 HMO enrollees from the study sample. The final sample in this dataset contains observations from 1,801 counties or county equivalent in 50 states and DC in the U.S. across 4 years (2009-2012).

The county-level dataset is used to construct variables to test Hypothesis 3. These variables include:

- RiskScore $_{c,t}$: the average risk scores of MA beneficiaries in county c of year t ;
- Treatment $_{c,t}$: 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 §5.2.2.
- CapitationRate $_{c,t}$: the average MA capitation payment for a standard Medicare beneficiary in county c at year t .
- Star $_{c,t}$: the average star rating of health plans in county c at year t .

5. Empirical Evidence of Reverse Cross Subsidization

This section provides empirical evidence of reverse cross subsidization in Medicare Advantage. Specifically, §5.1 develops the econometric model to test Hypothesis 1, and empirically shows that there is a negative association between the sickness levels of patients and the spending-cost differences of such patients. §5.2 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 negative association established in Hypothesis 1 is caused by the reverse cross subsidization practice of MA health plans.

5.1. Hypothesis 1

Hypothesis 1 states that there is a negative association 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 §5.1.1, and discuss the test results in §5.1.2.

5.1.1. 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 (MA Spending $_{i,t}$ – 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 Spending $_{i,t}$ – Estimated Cost $_{i,t}$.

The corresponding econometric model is thus specified as:

$$\text{MA Spending}_{i,t} - \text{Estimated Cost}_{i,t} = \beta_0 + \beta_{RS} \text{RiskScore}_{i,t} + \alpha_i + \gamma_t + \epsilon_{i,t}, \quad (2)$$

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}	-12.75***
Observations	5,712,882
R ²	0.635
Adjusted R ²	0.359

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)

5.1.2. 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 a 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 \$12,000 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 a negative association between the sickness of MA beneficiaries and their spending-cost differences.

To better understand this negative association, 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 would on average have a 1.09 points increase in risk scores. However, her spending-cost difference would on average decrease from \$8,862.85 to -\$1,890.74. 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 initiating dialysis. Table 3 reports similar results also for MA enrollees who have developed cancer. Furthermore, Table 14 and 15 in Appendix B find similar results in conditions including Diabetes¹², Cardiovascular diseases¹³, Inflammatory Bowel Disease (IBD), Cerebrovascular diseases¹⁴, Chronic Hepatitis and HIV/AIDS. These findings provide further support to the main estimation results of econometric model (2).

Finally, we reiterate that (2) only captures the association between the sickness levels of patients and the spending-cost differences of such patients. However, it cannot identify the causal impact of risk scores on the spending-cost differences. For example, a measurement error in $RiskScore_{i,t}$ may confound the impact of risk scores on the spending-cost differences. In particular, if the CMS-HCC risk adjustment formula systematically over-estimated the healthcare costs in high-risk patients in comparison to the low-risk patients, we would observe a negative association ($\beta_{RS} < 0$) in (2) simply due to this measurement error. In order to support our argument that this negative association can be explained by potential reverse cross subsidization practice of MA health plans,

¹² 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).

¹³ 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 Infarction (HCC83) and Specified Heart Arrhythmias (HCC92).

¹⁴ Cerebrovascular diseases include Cerebral Hemorrhage (HCC95), Ischemic or Unspecified Stroke (HCC96), Hemiplegia/Hemiparesis (HCC100) and Cerebral Palsy and Other Paralytic Syndromes (HCC101).

	Renal Diseases		Cancer	
	Before	After	Before	After
Spending-Cost Difference	\$8,862.85	-\$1,890.74	\$10,171.48	\$1,416.14
Risk Score	1.23	2.32	1.11	1.93

Table 3 Case Study: spending-cost differences and risk scores of MA enrollees before and after they develop renal diseases (HCC130, HCC131)¹⁵ and Cancer (HCC7-HCC10)¹⁶

we need a suitable empirical strategy which would separate this misallocation problem from other explanations such as possible measurement error, which we discuss next.

5.2. 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. §5.2.1 discusses that the ACA capitation rate cuts provided such an exogenous shock and thus can serve as our source of identification. §5.2.2 discusses the assignment of treatment and control groups in this quasi-natural experiment. Finally, §5.2.3 develops the econometric model based on this empirical strategy.

5.2.1. Identifying Reverse Cross Subsidization through ACA Capitation Rate Cuts

As discussed in §3.2, 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 reverse cross subsidization practice of MA health plans through examining how the observed negative association between risk scores and spending-cost differences would change after the ACA shock. As illustrated in Figure 1a, if there was

¹⁵ Renal diseases include Dialysis Status (HCC130) and Renal Failure (HCC131).

¹⁶ 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).

no reverse cross subsidization, the ACA shock, with its deeper payment cut for high risk patients, would simply amplify the negative relationship between risk scores and spending-cost differences. In contrast, if the negative association between risk scores and spending-cost differences was caused by the reverse cross subsidization practice of MA health plans, the ACA shock would mitigate this negative association. The reason is as follows. With reverse 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 cut 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, reverse cross subsidization would subside as in Figure 1b.

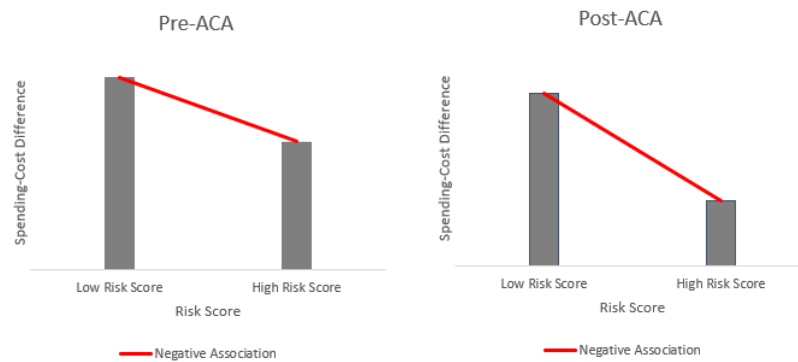


Figure 1a Identification Strategy Based on ACA Cut (No Reverse Cross Subsidization): If the negative association between risk scores and spending-cost differences was not caused by reverse cross subsidization, the ACA shock would not mitigate this negative association. Here, the negative associations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

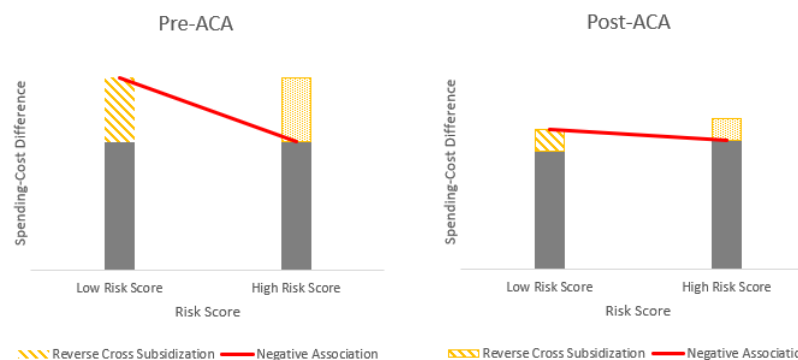


Figure 1b Identification Strategy Based on ACA Cut (Reverse Cross Subsidization): If the negative association between risk scores and spending-cost differences was caused by the reverse cross subsidization practice of MA health plans, the ACA shock would mitigate this negative association. Here, the negative associations 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 negative association between risk scores and spending-cost differences had been mitigated after the ACA shock. In addition, since the ACA reform does not affect how $RiskScore_{i,t}$ was estimated, a Difference-in-Difference study of this ACA shock would not be confounded by measurement errors in risk scores. Therefore, the reverse cross subsidization practice in MA is identified if the ACA shock mitigated the negative association between risk scores and spending-cost differences.

5.2.2. Treatment Assignment

To effectively isolate the impact of ACA capitation rate cuts on the negative association 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 §3.2, 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 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 were 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 difference below \$30 PBPM were assigned to the Two Year Phase-in Group, while counties with projected benchmark difference 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).¹⁷ That is, by focusing on the counties with benchmark

¹⁷ 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. This 25% capitation payment cut could not be offset by other ACA programs such as the Quality Bonus Payments (QBPs), which accounted for at most 5% of the capitation payment changes (See Appendix A).

difference 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 difference slightly below the \$30 cut-off value are assigned to the treatment group, while counties with projected benchmark difference 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.

5.2.3. The Econometric Model to Test the Main Hypothesis

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

$$\begin{aligned} & \text{MA Spending}_{i,t} - \text{Estimated Cost}_{i,t} \\ &= \beta_0 + \delta_{DID} \text{RiskScore}_{i,t} \times \text{Treatment}_{i,t} + \beta_1 \text{Treatment}_{i,t} \\ &+ \beta_{RS} \text{RiskScore}_{i,t} + \alpha_i + \gamma_t + \epsilon_{i,t}, \end{aligned} \quad (3)$$

where the treatment $\text{Treatment}_{i,t}$ is assigned as described in §5.2.2. In particular, since treatment and control groups are assigned at the county-level, the individual identifiers i become county-patient specific.¹⁸ In this specification, if our main hypothesis was true, then $\delta_{DID} > 0$. This is because the negative association between risk scores and spending-cost differences (β_{RS}) estimated in (2) would be mitigated by the ACA capitation rate cuts, i.e. $\beta_{RS} < \beta_{RS} + \delta_{DID}$.

The econometric model (3) empirically separates the potential measurement error problem in $\text{RiskScore}_{i,t}$ from the misallocation problem, and controls for other temporal changes happening simultaneously with the ACA capitation rate cuts. The identification strategy of this econometric model is summarized in Table 4. First, we note that the association between $\text{RiskScore}_{i,t}$ and $\text{MA Spending}_{i,t} - \text{Estimated Cost}_{i,t}$, i.e. $\beta_{RS} + \delta_{DID}$, is decomposed into two components. The first component β_{RS} measures the association independent of the ACA shock, while the second component δ_{DID} captures the change of this association caused by the ACA shock. As illustrated in §5.2.1, since the ACA shock does not change how risk scores are estimated, its treatment effect δ_{DID} can only be explained by the misallocation problem, instead of potential measurement error in risk scores. Second, as illustrated in Table 4, temporal changes γ_t are controlled by the second difference in (3).

¹⁸ In the rest of this paper, all individual identifiers i are county-patient specific.

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

Table 4 The Identification Strategy of Econometric model (3)

5.2.4. Results for the Main Hypothesis

The econometric model (3) is estimated using two-way fixed effect estimators. Table 5 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 negative association between the sickness levels of MA beneficiaries and their spending-cost differences was mitigated by the ACA capitation rate cuts, i.e. $|\beta_{RS}| > |\beta_{RS} + \delta_{DID}|$. As such, we confirm the main hypothesis and empirically identify the reverse cross subsidization practice in MA.

Next, we discuss the effect size of the ACA capitation rate cuts on MA health plans' spending patterns. First, the reverse cross subsidization amount was decreased by around $\beta_{RS} = \$900$ (according to the [\$28,\$32] neighborhood) per patient per unit risk score after the ACA capitation rate cuts. In other words, the ACA shock mitigated around 7% of the reverse cross subsidization amount, as measured by $|\frac{\delta_{DID}}{\beta_{RS}}|$. Furthermore, we also find that the average annual spending-cost difference in MA was reduced by around $\beta_1 = \$2,260$ (according to the [\$28,\$32] neighborhood) per patient after the ACA capitation rate cuts. Therefore, while the ACA capitation rate cuts constrained MA health plans from conducting further reverse cross subsidization, these cuts also made MA health plans spend less on their enrollees in general.

	[\$29,\$31]	[\$28,\$32]	[\$27,\$33]
β_{RS}	-12.99***	-12.89***	-12.62***
δ_{DID}	1.00***	0.90***	0.72**
β_1	-2.70***	-2.26***	-2.16***
Observations	285,879	368,844	453,612
R ²	0.604	0.608	0.635
Adjusted R ²	0.363	0.358	0.384

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 5 Estimation Results of Econometric Model (3)

6. Assessing the Impact of Reverse Cross Subsidization Practice on Risk Selection

Hypothesis 3 states that the reverse cross subsidization practice of MA health plans is associated with the risk selection problem in MA. §6.1 develops an empirical strategy to examine this association. §6.2 presents the econometric model based on this empirical strategy.

6.1. Assessing the Impact of Reverse Cross Subsidization Practice through ACA Capitation Rate Cuts

One way to test Hypothesis 3 is to examine whether counties phased into the ACA benchmarks earlier had higher county average MA risk scores in comparison to other counties. The reason is that if risk selection in MA was associated with the misallocation problem, the ACA shock, which mitigated reverse cross subsidization (as shown in our main hypothesis), should also reduce risk selection in MA. In this section, we use the county-level data introduced in §4.2 to empirically investigate possible changes in county average MA risk scores associated with reverse cross subsidization.

To control for alternative channels that may also affect county average MA risk scores, we include a set of control variables in the econometric model to test Hypothesis 3. Specifically, since the ACA shock might affect risk selection via channels other than reverse cross subsidization, we explicitly control for changes in MA capitation payments related to other ACA programs (e.g. QBPs)¹⁹ through $Treatment_{c,t} \times \log(CapitationRate_{c,t})$ and $\log(CapitationRate_{c,t})$. Furthermore, we control for the difference in MA health plan quality across counties by the average star ratings of MA health plans in each county, i.e. $Star_{c,t}$.

6.2. The Econometric Model to Test Hypothesis 3

The econometric model to test Hypothesis 3 is specified as:

$$\begin{aligned} RiskScore_{c,t} = & \beta_0 + \delta_{DID} Treatment_{c,t} + \beta_{Cap(DID)} Treatment_{c,t} \times \log(CapitationRate_{c,t}) \\ & + \beta_{Cap} \log(CapitationRate_{c,t}) + \beta_{Star} Star_{c,t} + \alpha_c + \gamma_t + \epsilon_{c,t}, \end{aligned} \quad (4)$$

where δ_{DID} captures the impact of ACA shock on risk selection through the reverse cross subsidization channel. Therefore, testing Hypothesis 3 is equivalent to testing whether $\delta_{DID} > 0$.

Table 6 summarizes the identification strategy of (4). Specifically, time invariant changes of MA risk scores, i.e. α_c , are controlled by the first difference. Besides, pure temporal changes in MA risk scores, i.e. γ_t , are controlled by the second difference. The econometric model (4) further decomposes the ACA shock to those that were directly associated with risk selection through reverse

¹⁹ For example, MA health plans could earn back parts of the reduction in ACA benchmark payments through the Quality Bonus Payments (QBPs) program if they achieved certain quality ratings. Please refer to Appendix A for a detailed discussion of these programs.

Group	Time	Outcomes	First Difference	Second Difference
	Before	α_c		
Control			γ_t	
	After	$\alpha_c + \gamma_t$		
				$\beta_{Cap(DID)} + \delta_{DID}$
	Before	α_c		
Treatment			$\gamma_t + \beta_{Cap(DID)} + \delta_{DID}$	
	After	$\alpha_c + \gamma_t + \beta_{Cap(DID)} + \delta_{DID}$		

Table 6 The Identification Strategy of Econometric model (4)

cross subsidization, i.e. δ_{DID} , and those that were indirectly associated risk selection through other ACA programs like QBPs, i.e. $\beta_{Cap(DID)}$. Therefore, δ_{DID} identifies the impact of ACA shock on risk selection through the reverse cross subsidization channel.

6.3. Results for Hypothesis 3

	[\$29,\$31]	[\$28,\$32]	[\$27,\$33]
δ_{DID}	1.781***	1.714***	1.847***
β_{Cap}	0.031	0.165	0.233
$\beta_{Cap(DID)}$	-0.194***	-0.188***	-0.202***
β_{Star}	0.020**	0.027***	0.014
Observations	124	203	304
R ²	0.921	0.917	0.912
Adjusted R ²	0.884	0.882	0.876

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

Table 7 Estimation Results of Econometric Model (4)

Table 7 presents the estimation results of econometric model (4), which is estimated using two-way fixed effect estimators. The results confirm Hypothesis 3. Specifically, the parameter of interest, δ_{DID} , is statistically significant and positive in all three neighborhood specifications, implying that risk selection in MA was reduced as reverse cross subsidization was mitigated. The estimation results of the other parameters also have the expected signs. Specifically, β_{Cap} and $\beta_{Cap(DID)}$ capture the impact of capitation payments on risk selection in MA that is associated with other ACA programs. Consistent with the results by Fioretti et al. (2019), we find that MA health plans that received higher capitation payments after ACA through programs like QBPs tended to have worse risk selection, i.e. $\beta_{Cap(DID)} < 0$. Finally, as reported in the existing risk selection literature (c.f. Rahman et al. (2015), Meyers et al. (2018, 2019), Morrissey et al. (2013), Jacobson et al. (2019b)), we observe that the sicker patients were attracted to high quality MA health plans, as indicated

by the finding that counties with higher average MA health plan ratings tended to have sicker MA population, i.e. $\beta_{Star} > 0$.

As for the effect size, we find that the impact of reverse cross subsidization on risk selection tends to be more profound than other sources of risk selection previously documented in the literature. Specifically, Table 7 shows that counties in the treatment group, where the reverse cross subsidization practice is restricted, had at least a $\delta_{DID} = 1.714$ points increase in county average risk scores compared to those in the control group, where the reverse cross subsidization practice is not restricted. To put the increase in risk scores into context, consider Brown et al. (2014) who found that improving the risk adjustment formula increased the average risk score in MA by 0.106. Such a comparison of the effect sizes suggests that the misallocation of capitation payments have a more substantial impact on risk selection in MA, compared to the miscalculation of capitation payments. An important implication of this finding is that, to address the risk selection problem in MA, restricting the reverse cross subsidization practice of MA health plans might be notably more effective than correcting for the estimation errors in risk adjustment and capitation payments.

7. Robustness Checks

This section conducts several robustness checks to examine our findings. Specifically, we check whether results for the main hypothesis are robust against the panel attrition problem in §7.1 and measurement errors in the dependent variable in §7.2. In addition, we examine the parallel trend assumption of our DID analysis through a placebo test in §7.3.

7.1. 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 §5 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 §5 while the reverse 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 §5 is panel attrition, the narrowing spending-cost difference we observed after the ACA shock does not necessarily imply mitigated reverse cross subsidization in MA.

To address this concern, we conduct a subsample analysis focusing on the continuously enrolled MA patients throughout the 2009-2012 study period. By focusing on 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 8. We note that these results are qualitatively consistent with our main results in Table 5. Furthermore, the main parameter of interest, δ_{DID} in (3), becomes even more significant and has larger effect sizes than those in the main results. As such, this subsample analysis rules out the aforementioned alternative explanation of the observed mitigated reverse cross subsidization in MA.

	[\$29,\$31]	[\$28,\$32]	[\$27,\$33]
β_{RS}	-12.04***	-12.01***	-11.73***
δ_{DID}	1.15***	1.15***	0.96***
β_1	-3.01***	-2.56***	-2.16***
Observations	158,968	191,912	211,571
R ²	0.466	0.461	0.471
Adjusted R ²	0.295	0.288	0.300

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

7.2. 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 reverse cross subsidization in §5 could be mainly driven by several extreme medical claims, and thus was not representative of the MA market.

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

$$\begin{aligned}
 & \text{Log}[(\text{MA Spending}_{i,t} - \text{Estimated Cost}_{i,t})^*] \\
 & = \beta_0 + \delta_{DID} \text{RiskScore}_{i,t} \times \text{Treatment}_{i,t} + \beta_1 \text{Treatment}_{i,t} \\
 & + \beta_{RS} \text{RiskScore}_{i,t} + \alpha_i + \gamma_t + \epsilon_{i,t},
 \end{aligned} \tag{5}$$

where the dependent variable is defined as

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

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 (5) is shown in Table 9. We can see that the parameters of interest, i.e. δ_{DID} , have the expected sign as those in Table 5. In addition, these estimates remain statistically significant when we examine them in a smaller neighborhood around the \$30 cut-off, e.g. [\$28,\$32], [\$29,\$31].²⁰ As such, these findings indicate that reverse cross subsidization is not confounded by the fat-tailed distributions of spending-cost differences.

	[\$29,\$31]	[\$28,\$32]	[\$27,\$33]
β_{RS}	-0.057***	-0.022***	-0.022***
δ_{DID}	0.003***	0.001**	0.001*
β_1	-0.010***	-0.003***	-0.003***
Observations	285,879	368,844	453,612
R ²	0.571	0.580	0.598
Adjusted R ²	0.311	0.312	0.321

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

Table 9 Estimation Results of Econometric Model (5)

7.3. A Placebo Test of the 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 counterfactual 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.

The parallel trend assumption in DID is in principle difficult to test. A common approach in practice is to conduct an event-study test to examine the pre-treatment trends between the treatment and control groups (c.f. Angrist and Pischke (2008)). However, to have enough statistical power, one would typically need data from multiple time points before the treatment time to conduct this kind of event-study tests (Roth 2019). For the study period (2009-2012) of this paper, we only have one time point, namely 2009, before the treatment in 2010 when the ACA was passed, and thus cannot test the parallel trend assumption through this approach.

To alleviate these concerns, we conduct a placebo test (c.f. Cooper et al. (2011), Cunningham (2021), Finkelstein (2002)) to examine the parallel trend assumption. The idea of this placebo test is

²⁰ We remark that the estimated effects are expected to be less significant when we expand this neighborhood because counties in the expanded neighborhood would be less homogeneous.

that we restrict the DID regression (3) to a subsample of MA health plans (the placebo group) that were less likely to respond differentially to the ACA capitation payment cuts based on whether a patient was in a treatment or a control county, and see whether the treatment effect, i.e. δ_{DID} in (3), is statistically significant in this subsample. Ideally, we should not find any statistically significant treatment effect in this restricted DID regression. Otherwise, any treatment effects we find in this subsample would come from some underlying differences in parallel trends between treatment and control groups, instead of MA health plans' responses to the ACA capitation payment cuts.

We discuss our approach to placebo group selection next. We contend that MA health plans that operated in both treatment and control counties are ideal candidates for the placebo test. To see this, recall that MA health plans differ in their responsiveness to the ACA capitation payment cuts, depending on whether they operate in the treatment or control counties. In §5, we established that MA health plans operating in the treatment counties had stronger incentives to restrict their reverse cross subsidization practice in response to the ACA capitation payment cuts, compared to MA health plans operating in the control counties. However, for MA health plans operating in both treatment and control counties, it would be operationally difficult, if not impossible, for them to restrict their reverse cross subsidization practice only in treatment counties while continuing such practice without restriction in control counties. 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}	-13.82***	-13.38***	-12.97***
δ_{DID}	0.53	0.19	-0.50
β_1	-2.37***	-1.45*	-1.19*
Observations	130,850	196,664	253,657
R ²	0.636	0.635	0.650
Adjusted R ²	0.398	0.378	0.387

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

Table 10 Estimation Results of Econometric Model (3) for MA Health Plans in the Placebo Group

To operationalize this placebo test, we re-estimate the econometric model (3) for MA health plans in the placebo group. Specifically, since these MA health plans enrolled patients both in treatment and control counties, we would not expect to find any statistically significant effect based on treatment status if the parallel trend assumption held. As shown in Table 10, we do not find any statistically significant treatment effect (δ_{DID}) among patients enrolled in these placebo MA health

	[\$29,\$31]	[\$28,\$32]	[\$27,\$33]
β_{RS}	-12.41***	-12.74***	-12.74***
δ_{DID}	1.22**	1.43**	1.42***
β_1	-2.64***	-2.75***	-3.04***
Observations	155,029	172,180	199,955
R ²	0.587	0.598	0.644
Adjusted R ²	0.312	0.319	0.372

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

Table 11 Estimation Results of Econometric Model (3) for MA Health Plans not in the Placebo Group

plans, which supports the parallel trend assumption of our DID design. In addition, among patients enrolled in the non-placebo MA health plans, Table 11 shows statistically significant treatment effect (δ_{DID}), which is consistent with our main results in Table 5.

8. Conclusion

Healthcare capitation payment models, in particular the Medicare capitation program, have been extensively studied in recent decades. However, due to the limited access to MA health plans claims data, little is known about how MA health plans allocate capitation payments across various patient risk groups. More importantly, we have 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 benefits she gets from MA. Our 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.

Our study provides, to our knowledge, the first empirical evidence of reverse cross subsidization practice in MA, a major resource misallocation problem overlooked in the extant literature. Specifically, we empirically show that by reallocating parts of the capitation payments from the sick to cross subsidize the healthy, MA health plans strategically overspend on the low-risk patients and underspend on the high-risk patients. The resulting misallocation worsens the healthcare inequality problem in MA, where the low-risk population benefit at the expense of the more vulnerable high-risk population. By empirically documenting this perverse reverse cross subsidization practice in MA health plans, we hope to raise attention to this critical resource misallocation problem in one of the world's largest healthcare capitation programs.

While recent healthcare literature has reported various practices hinting at misallocation problems such as MA health plans' gym membership offers to healthier Medicare beneficiaries (c.f. [Cooper and Trivedi \(2012\)](#)) or lower quality nursing home care provision in MA compared to Traditional Medicare (c.f. [Meyers et al. \(2018\)](#)), these studies tended to have narrow scope with limited evidence base. By investigating patient level healthcare spending as well as expected cost for more than 2 million MA patients, our paper provides the first large scale evidence pointing out reverse cross subsidization in MA health plans. This finding is also important because by law MA health plans are required to provide “the same level of benefits” to their enrollees as provided by Traditional Medicare.²¹ However, our results suggest that MA health plans may not abide by this law in practice as they spend capitation payments and use healthcare resources in a way that is disproportionately benefiting the healthier while disadvantaging the sicker.

In addition, our study reveals that reverse cross subsidization practice of MA health plans is an important contributing factor to the risk selection problem in MA, where low-risk patients are more likely to enroll in MA compared to the high-risk patients. This is the first empirical evidence suggesting that the risk selection problem in MA can be attributed to the misallocation, instead of the miscalculation, problem of the MA capitation payments. In other words, our paper empirically demonstrates that improving risk adjustment formula is not sufficient to eliminate the risk selection problem in MA. To effectively address the risk selection problem, CMS should focus more on regulating how MA health plans allocate the capitation payments.

Our study also calls for more future research in this direction. First, as population-based payment models are increasingly adopted by healthcare payers, reverse cross subsidization practice is likely to exist outside of the MA capitation program. Further investigations are needed to understand the role of reverse cross subsidization practice in other population-based payment models. Second, provided that MA health plans consistently underspend on high-risk patients, these patients may experience adverse health events in the long run. Consequently, future work can assess health outcome implications of reverse cross subsidization practice with long-term follow-up data.

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.
- Angrist, J. D. and Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

²¹ *UnitedHealthcare Ins. Co. v. Azar*, Civil Action No. 16-157 (RMC) (D.D.C. Jan. 27, 2020): “A Medicare Advantage insurer must provide, at a minimum, the same level of benefits provided by traditional Medicare itself. See 42 U.S.C. § 1395w-22(a).” ([U.S. District Court 2020](#))

-
- 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 (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 August 24, 2019).
- Centers for Medicare & Medicaid Services (2011c). 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 (2011d). Risk Adjustment Data Validation (RADV) Medical Record Checklist and Guidance. Available at <https://www.cms.gov/Medicare/Medicare-Advantage/Plan-Payment/Downloads/radvchecklist.pdf> (Accessed October 18, 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 (2018a). FFS Data (2008-2016). Available at <https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/FFS-Data.html> (Accessed May 22, 2108).

- Centers for Medicare & Medicaid Services (2018b). Monthly MA Enrollment by State/County/Contract. Available at <https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/MCRAdvPartDEnrolData/Monthly-MA-Enrollment-by-State-County-Contract.html> (Accessed May 22, 2108).
- Centers for Medicare & Medicaid Services (2018c). Plan Payment Data. Available at <https://www.cms.gov/Medicare/Medicare-Advantage/Plan-Payment/Plan-Payment-Data.html> (Accessed May 22, 2108).
- Centers for Medicare & Medicaid Services (2018d). Ratebooks & Supporting Data. Available at <https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Ratebooks-and-Supporting-Data.html> (Accessed May 22, 2018).
- 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 January, 21, 2020).
- 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).
- Chen, C. and Savva, N. (2018). Unintended consequences of hospital regulation: The case of the hospital readmissions reduction program. *Available at SSRN 3236983*.
- Cooke, R. M., Nieboer, D., and Misiewicz, J. (2014). *Fat-tailed distributions: Data, diagnostics and dependence*, volume 1. John Wiley & Sons.
- 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.
- Cooper, Z., Gibbons, S., Jones, S., and McGuire, A. (2011). Does hospital competition save lives? evidence from the english nhs patient choice reforms. *The Economic Journal*, 121(554):F228–F260.
- Cunningham, S. (2021). *Causal inference: The mixtape*. Yale University Press.
- 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.
- 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.
- Finkelstein, A. (2002). The effect of tax subsidies to employer-provided supplementary health insurance: evidence from canada. *Journal of Public Economics*, 84(3):305–339.
- Fioretti, M., Wang, H., et al. (2019). Subsidizing inequality: Performance pay and risk selection in medicare. Technical report, Sciences Po Departement of Economics.

- Geruso, M. and Layton, T. J. (2017). Selection in health insurance markets and its policy remedies. *Journal of Economic Perspectives*, 31(4):23–50.
- 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.
- 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.
- 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.
- 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.
- 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.
- Medicare Payment Advisory Commission (2019). Report to congress: Medicare payment policy. Available at http://www.medpac.gov/docs/default-source/reports/mar19_medpac_entirereport_sec.pdf (Accessed December, 3, 2020).
- 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.
- 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.theactuarmagazine.org/bonus/> (Accessed May, 21, 2018).
- 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.
- Roth, J. (2019). Pre-test with caution: Event-study estimates after testing for parallel trends. *Unpublished manuscript, Department of Economics, Harvard University*.
- Savva, N., Tezcan, T., and Yildiz, O. (2018). Can yardstick competition reduce waiting times? *Management Science*.
- She, Z., Ayer, T., and Montanera, D. (2020). Can big data cure risk selection in healthcare capitation programs? Technical report, Georgia Institute of Technology.
- 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 (1999). Social Security Act §1853(a)(1)(B)(ii). Available at https://www.ssa.gov/OP_Home/ssact/title18/1853.htm (Accessed May, 23, 2021).
- 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. District Court (2018). *UnitedHealthcare Ins. Co. v. Azar*, 330 F. Supp. 3d 173. Available at <https://casetext.com/case/unitedhealthcare-ins-co-v-azar-1> (Accessed May, 23, 2021).
- U.S. District Court (2020). *UnitedHealthcare Ins. Co. v. Azar*, Civil Action No. 16-157 (RMC). Available at <https://casetext.com/case/unitedhealthcare-ins-co-v-azar-2> (Accessed May, 23, 2021).

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 §3.1, 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.²² As discussed in §3.2, 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 Groups 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 half of this section explains how MA capitation payments were determined when the ACA rates fully phased in. The second half of this section 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.

A.1. MA Capitation Payments with Fully Phased-in ACA Benchmark (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 benchmark 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 12, 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 $\text{FFS Rate}_{c,t} \times quartile_c$.

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

²² 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.

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

Table 12 Medicare Advantage County Quartile Payment ($quartile_c$)

service quality. Specifically, CMS assigned each MA health plan a star rating based on their quality in a list of monitored services²³, 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 ($FFS\ Rate_{c,t} \times bonus_{c,t}^h$) in (1). The exact bonus amount can be calculated based on Table 13. Furthermore, for a selected set of counties, they can receive a double bonus on their benchmark payments. For example, a MA health plan with star rating 3 in 2012 would have an additional bonus $bonus_{c,t}^h = 3\%$ ($bonus_{c,t}^h = 6\%$ 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 13 QBP Bonus Benchmark Payment ($bonus_{c,t}^h$)

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

$$ACA_Benchmark_{c,t} = FFS\ Rate_{c,t} \times (quartile_c + bonus_{c,t}^h). \quad (6)$$

Before taking bonus payments into account, $FFS\ Rate_{c,t} \times quartile_c$ were significantly lower than the pre-ACA benchmark payments ($\text{Max}\{Update_{c,t}, FFS\ Rate_{c,t}\}$). 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 QBPs program indeed offered incentives for MA health plans to reduce their losses in benchmark payments through providing better service quality.

²³ The list of monitored services and their targeted quality can be found at [Centers for Medicare & Medicaid Services \(2020\)](https://www.cms.gov/medicare/medicaid-support/quality/2020).

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

As discussed in §3, the ACA benchmark 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

$$\begin{aligned} Benchmark_{c,t} = & \text{Max}\{Update_{c,t}, \text{FFS Rate}_{c,t}\} \times (1 + bonus_{c,t}^h) \times (1 - PhaseInFactor_{c,t}) \\ & + \text{FFS Rate}_{c,t} \times (quartile_c + bonus_{c,t}^h) \times PhaseInFactor_{c,t}, \end{aligned} \quad (7)$$

where $PhaseInFactor_{c,t} \in [0, 1]$. Clearly, the actual MA benchmark payments during the 2012-2017 period was a weighted average of $\text{Max}\{Update_{c,t}, \text{FFS Rate}_{c,t}\} \times (1 + bonus_{c,t}^h)$ 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

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

(Centers for Medicare & Medicaid Services 2011c). 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.

B. Supplementary Figures and Tables

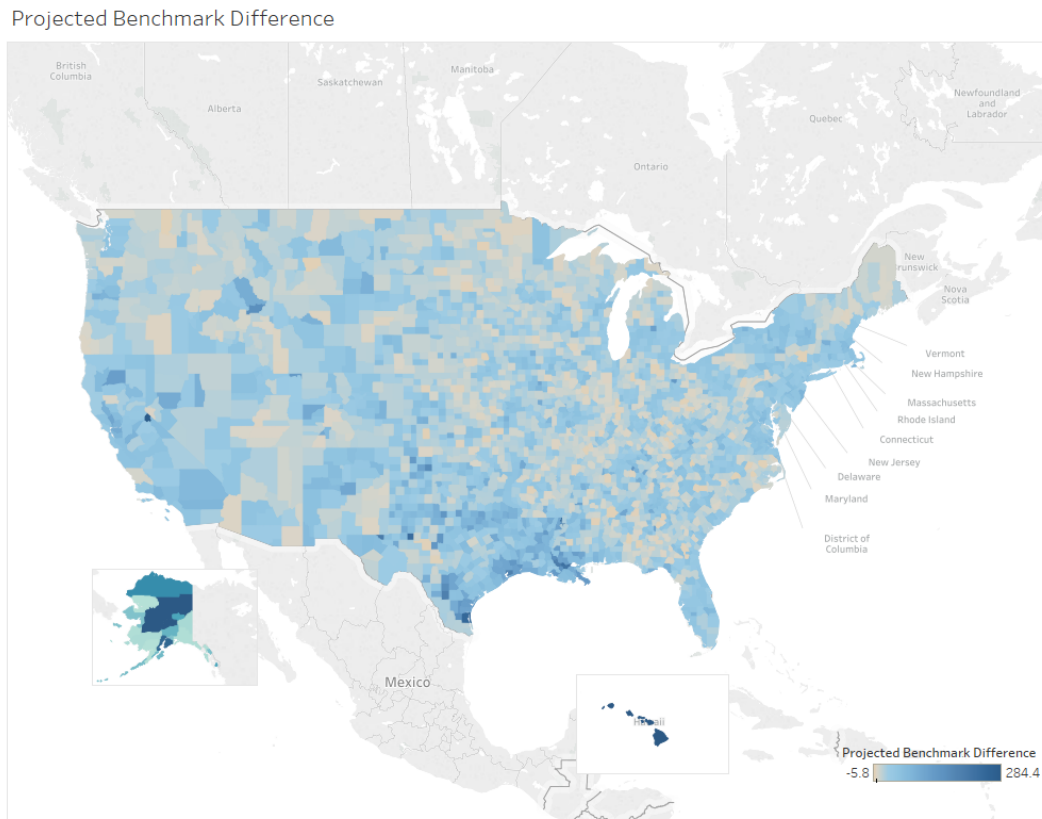


Figure 2 The Projected Difference between the ACA and Pre-ACA Benchmark Payments Per Beneficiary Per Month (PBPM)

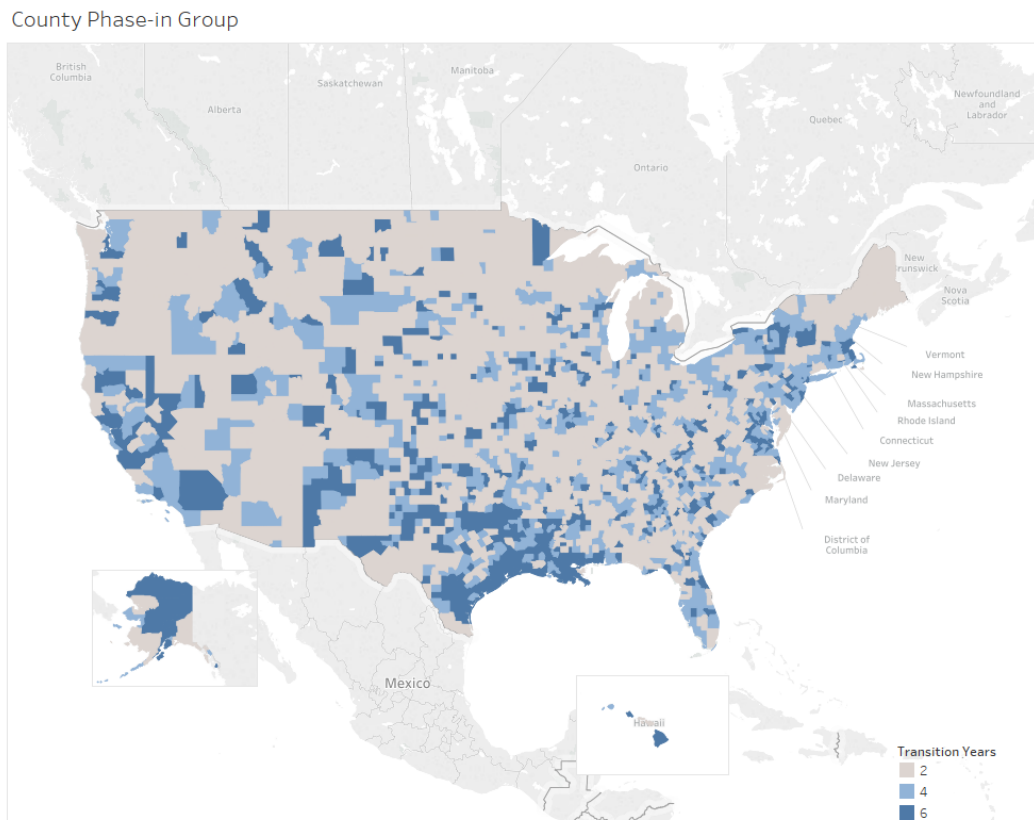


Figure 3 Phase-in Group of Counties across the U.S.: Counties in brown were in the Two Year Phase-in Group, who phased in the ACA benchmark by 2013. Counties in light blue were in the Four Year Phase-in Group, who phased in the ACA benchmark by 2015. Counties in dark blue were in the Six Year Phase-in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

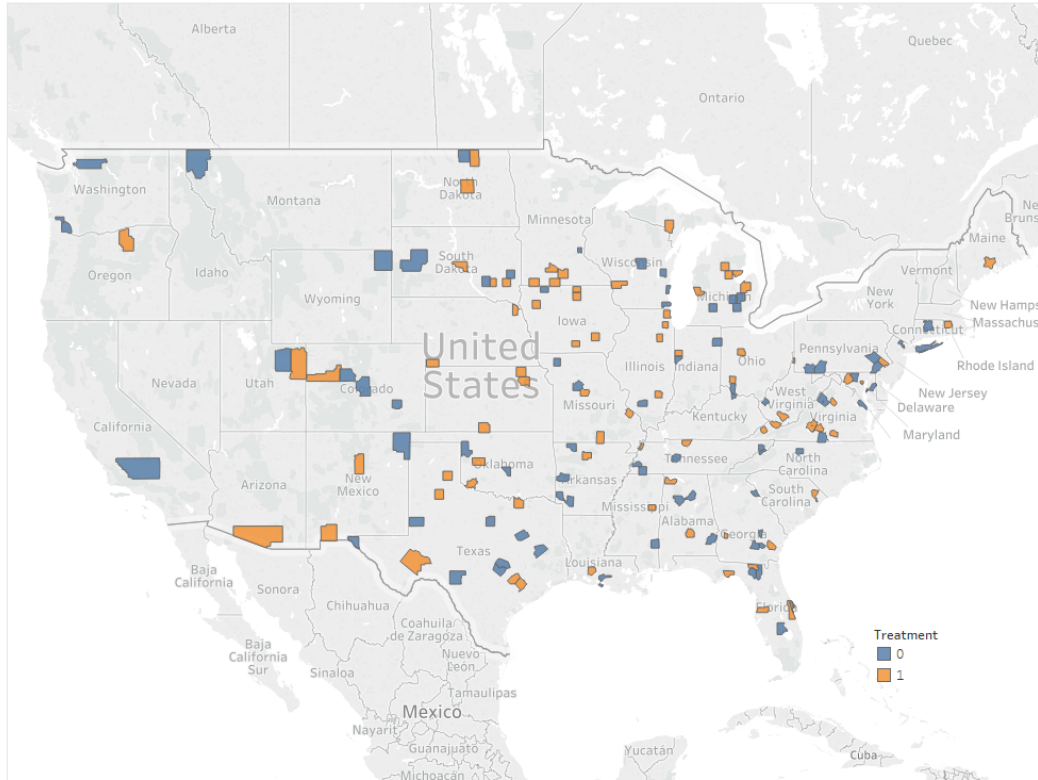


Figure 4 Counties in the DID Analysis with Projected Benchmark Difference in $[\$27, \$33]$: There were 155 of these counties in our data. All counties are in the control group before 2012. In 2012, 77 counties with projected benchmark difference in $[\$27, \$30]$ are assigned to the treatment group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

	HIV/AIDS		Cancer		Diabetes		Cardiovascular	
	Before	After	Before	After	Before	After	Before	After
Spending-Cost Difference	\$12,464.44	-\$4,193.97	\$10,171.48	\$1,416.14	\$6,568.65	\$555.34	\$11,156.95	\$25.63
Risk Score	1.38	2.87	1.11	1.93	0.98	1.63	1.03	2.10

Table 14 Case Study: Changes in mean spending-cost differences and risk scores when MA enrollees developed HIV/AIDS (HCC1), Cancer (HCC7-HCC10)²⁴, Diabetes (HCC15 - HCC19)²⁵ or Cardiovascular diseases(HCC79-HCC83, HCC92)²⁶

	IBD		Renal		Cerebrovascular		Hepatitis	
	Before	After	Before	After	Before	After	Before	After
Spending-Cost Difference	\$11,709.04	\$1,310.19	\$8,862.85	-\$1,890.74	\$13,514.72	\$182.52	\$11,512.05	\$2,191.49
Risk Score	1.30	1.99	1.23	2.32	1.41	2.42	1.30	2.10

Table 15 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)²⁷, Cerebrovascular diseases (HCC95, HCC96, HCC100, HCC101)²⁸ or Chronic Hepatitis (HCC27).

²⁴ 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).

²⁵ 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).

²⁶ 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 Infarction (HCC83) and Specified Heart Arrhythmias (HCC92).

²⁷ Renal diseases include Dialysis Status (HCC130) and Renal Failure (HCC131).

²⁸ Cerebrovascular diseases include Cerebral Hemorrhage (HCC95), Ischemic or Unspecified Stroke (HCC96), Hemiplegia/Hemiparesis (HCC100) and Cerebral Palsy and Other Paralytic Syndromes (HCC101).