NBER WORKING PAPER SERIES

OPEN ENROLLMENT PERIODS AND PLAN CHOICES

Francesco Decarolis
Andrea Guglielmo
Calvin Luscombe

Working Paper 24156
http://www.nber.org/papers/w24156

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2017

Decarolis gratefully acknowledges financial support from the National Science Foundation (SES-1357705) and the Sloan Foundation (2011-5-23 ECON). Earlier drafts of this study benefited from comments by Pierre Andre Chiappori, Mark Duggan, Randall Ellis, Joshua Gottlieb, Kate Ho, Pietro Tebaldi and Robert Town. Andrea Guglielmo is an associate at Analysis Group, Inc. Research for this article was undertaken when he was a student at University of Wisconsin - Madison. The views presented in this article are those of the authors and do not reflect those of Analysis Group. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w24156.ack

Open Enrollment Periods and Plan Choices
Francesco Decarolis, Andrea Guglielmo, and Calvin Luscombe
NBER Working Paper No. 24156
December 2017, Revised August 2019
JEL No. I11,I18,L22

ABSTRACT

Open enrollment periods are pervasively used in insurance markets to limit adverse selection risks resulting when enrollees can switch plans at will. We exploit a change in the open enrollment rules of Medicare Advatage to analyze how beneficiaries responded to the option of switching to a 5-star rated plan at anytime, in a setting where insurers adjusted premiums and benefit design to counterbalance the increased selection risk. We present three findings: within-year switches to 5-star plans increase by 7-16%; demand for 5-star plans across the years does not change; the enrollees who switch to a 5-star plan during the year are in better health status than those who do not switch.

Francesco Decarolis
Bocconi University
Via Sarfatti 25
Milan, 20136
Italy
and EIEF
francesco.decarolis@unibocconi.it

Andrea Guglielmo
Analysis Group, Inc
Prudential Center Plaza III
111 Huntington Ave, B
Boston, MA 02199
aguglielmo@wisc.edu

Calvin Luscombe
Boston University
calvinl@bu.edu

I Introduction

The growing economic importance of health insurance markets has driven the flourishing of research into what features of these markets can lead to more desirable social outcomes. Several of these studies have involved the design of the Medicare system. With expenditures totalling $646.2 billion in 2015 and growing by 4.5 percent relative to the previous year, Medicare represents, through its Medicare Advantage and Part D programs, the largest existing case of a publicly founded, but privately provided health insurance system.

As is typical in insurance markets, both Medicare Advantage, covering hospital stays and physician visits, and Part D, covering prescription drugs, have an "open enrollment period" during which consumers select a plan that will subsequently provide them with coverage under clearly defined contractual conditions. Among these conditions is the inability for the enrollee to switch plan at will during the coverage period. Open enrollment periods (OEPs) play a key role in the stability of health insurance markets as they limit the perverse dynamics produced by adverse selection: beneficiaries that can remain uninsured (or choose cheap, low-coverage plans) when they are healthy and then switch to generous plans when sick pose the risk of sending high-coverage plans into an "adverse selection death spiral" of increasing costs and increasing premiums, ultimately leading to the collapse of the market.[1]

In private insurance markets, insurers can often refuse to sell, but this is typically not an option for publicly subsidized health insurers programs like Medicare Advantage or the Affordable Care Act (ACA) exchanges. In these programs, all enrollment requests from eligible beneficiaries must be accepted. This feature implies that increasing the flexiblity for enrollees to change plans during the coverage period might come at the cost of worsening adverse selection. Recent work by Diamond et al. (2018) shows that the ACA exchanges are currently experiencing this type of problem: attrition is widespread, with more than half of all new enrollees dropping coverage before the end of the plan year. This paper looks at the same phenomenon within the context of a 2012 reform of Medicare Advantage under which the OEP rules were changed to allow enrollees to switch at anytime under the sole condition that the destination plan is rated 5-star (the highest score in the Medicare plan quality rating system). This reform, known as the "5-star Special Enrollment Period" (or 5-star SEP), aimed at increasing enrollment in 5-star contracts.[2] It involves a large share of

[1] For a well known discussion of a case of adverse selection death spiral involving the health insurance plans offered to Harvard University employees see Cutler and Reber (1998) and Cutler and Zeckhauser (1998).

[2] In Medicare Advantage enrollees choose plans but the star rating system applies to contracts. Contracts typically include multiple plans. A contract is a particular product type (HMO, PPO or Private FFS) covering a specific service area (i.e., county or group of counties). Within a contract, different plans typically have differences in their benefit package (type of coverage, premium, copayment, etc.). In this paper, we use

the Medicare beneficiaries – in 2017 the 5-star SEP was available to 11.5 million individuals residing in areas with at least one 5-star plan.[3]

Various demand and supply forces present in the market are likely to limit the possibility of this reform triggering an "adverse selection death spiral." On the demand side, both consumer inertia in choosing an insurance plan and the inherent complexity of changing a Medicare Advantage plan – which, as discussed below, implies changing provider network – are likely to reduce plan-switching behavior. On the supply side, there are at least two mechanisms at play. First, Medicare Advantage plans receive risk-adjusted payments. Although it may not be possible to perfectly compensate for all cost, risk adjustment serves to compensate plans receiving an influx of less healthy beneficiaries. Second, insurers can modify both the premiums and other elements of their plans' menus.

In a previous study, we analyzed this latter feature by studying how insurers responded to the 5-star SEP (Decarolis and Guglielmo, 2017). By exploiting the geographical variation in the availability of 5-star plans, we identified the causal effect of the 5-star SEP on the distribution of plan characteristics in the markets affected by the reform. We found strong empirical evidence in support of the theoretical predictions of models à la Rothschild and Stiglitz (1976) and Glazer and McGuire (2000) in which plans alter their product in an attempt to attract good risks: relative to the distribution of competing plans, 5-star plans lower both their premium and their generosity, especially on those margins most valued by the enrollees in worst health conditions. That study, however, left open the question of what the impact on demand has been of the combined effects of free plan switching by enrollees and changes to plan designs by insurers. Answering this question is the main contribution of the current study and is key to understanding the potential effectiveness of using open enrollment rules as a tool to regulate insurance markets with managed competition.

To identify how demand responded to the SEP reform, we use a similar approach to that of (Decarolis and Guglielmo, 2017). We exploit the geographical variation in 5-star Medicare Advantage plans to compare demand in markets with 5-star plans to that in similar markets where no 5-star plan is offered. Our difference-in-differences strategy is particularly effective when insurers have limited scope to game the star rating system. Therefore, we focus on the first two years of the reform (2012 and 2013), when insurers could alter the plan design but not their star rating due to the lag in the timing of the specific measures that compose

both terms "contract" and "plan," depending on which of the two is most appropriate.

[3]As discussed in section 2, the reform was introduced as part of the quality bonus payment demonstration. The 11.5 million figure is from Q1 Medicare and is based on the fact that in 2017, 5-star rated Medicare Part D plans were available across all counties in 12 states and 5-Star rated Medicare Advantage plans were available in 261 counties across 18 states.

the rating. We also restrict the control group to plans with a rating no lower than 4 stars to account for the different financial incentives created by the bonuses for higher rated plans introduced by the Patient Protection and Affordable Care Act (see Layton and Ryan (2015)).

Our main findings are as follows. First, we estimate that the within-year increase in enrollment due to the 5-star SEP ranges from 7 to 16 percent of the enrollment base of the 5-star plans. This indicates a sizable response by consumers to the new SEP. Second, we estimate either an insignificant or a positive effect (depending on the model specification) of the reform on enrollment changes across the years. This is indicative of inertia in plan choice: enrollees do not take advantage of the possibility to stay outside the Medicare Advantage program (or to enroll in the cheapest plans) during the open enrollment period and to switch to 5-star plans only if hit by a health shock. Third, the risk pools of 5-star plans improve, albeit only by a small amount.

The latter finding is not indicative of advantageous selection by itself. Before the reform, 5-star plans tended to have particularly high-risk enrollees. Therefore, their average risk score might have improved because they are bringing in enrollees that, despite being among the high risk enrollees in their original plan, still represent a lower risk than the average 5-star enrollee. Using detailed claim-level data, however, we estimate that the probability of switching to a 5-star plan is negatively associated with measures of poor health. In particular, this is what we obtain for four measures accounting for nearly all the major conditions characterizing poor health for acute, chronic and mental health pathologies. Therefore, we conclude that the increased demand for 5-star plans resulting from within-year switches is not associated with greater adverse selection, but with advantageous selection. This is consistent with the supply response to the SEP involving changes to the plan characteristics that made them more appealing to most enrollees (though lower premiums), but less so to those in worse health (through lower benefit generosity for enrollees in poor health).

The two, closely connected implications deriving from these three results are that the 5-star SEP was effective in steering enrollees toward 5-star plans, and that insurers offering 5-star plans were effective in preventing this increase in demand from being driven by high cost enrollees. These results are therefore informative of the usefulness of designing special enrollment periods as a tool to guide the functioning of health insurance markets. Moreover, they indicate that using this tool requires taking into account both supply and demand responses. Although there is no theoretical literature guiding the design of optimal enrollment periods, as health insurance is increasingly organized in the form of markets with regulated competition, we expect increasing attention to be paid to a market design approach to these markets. A recent example of a study taking this perspective is Einav,

Finkelstein and Tebaldi (2019), who compare subsidies for consumers and risk adjustment for insurers and find that the former are a more effective regulatory tool in markets with adverse selection. With a better understanding of their effects, open enrollment periods rules might also become a useful tool to steer the market towards socially desirable goals.

This study is most closely connected to two other existing works on the effects of the 5-star SEP. The first is our supply-side study, (Decarolis and Guglielmo, 2017), discussed earlier.[4] The second, by Madeira (2015), is an early attempt to study behavioral biases among Part D enrollees exposed to the 5-star SEP in 2012. The 5-star SEP reform involved not just Medicare Advantage, but also Medicare Part D – the voluntary program for prescription drug insurance plans. Madeira (2015) studies whether, by removing the typical Part D enrollment deadline, the 5-star SEP could have induced consumers to switch plans less frequently by allowing them to procrastinate. His results suggest that switching rates (across the years) decrease as a result of the policy change in a way that is consistent with a procrastination bias. Our results complement and substantially extend these findings as they look directly at the main aspect of the policy (within-year switches, instead of across-years plan changes) and they do so by using data not only from 2012, but also from 2013.[5] More crucially, by focusing on Medicare Advantage instead of Part D, our analysis benefits from more comprehensive geographical variation of the policy, which involved nearly 180 Medicare Advantage counties but only 2 Part D regions.[6]

The evidence in this study also complements the very scarce evidence that exists on the effects of open enrollment periods in other markets. In the ACA exchanges, we have already mentioned the study by Diamond et al. (2018) documenting how high attrition rates undermine market stability, leading to insurers' exit and higher premiums for enrollees who do not drop out. The only other paper that we are aware of in this area is Ellis and Savage (2008), which looks at a reform by the Australian government aimed at increasing private health insurance coverage by introducing selective age-based premium increases for those enrolling after a deadline. They find the introduction of the deadline to be effective in inducing consumers to enroll right away rather than delay.

[4]Related evidence showing insurers' strategic choice of plan features in environments different from the one studied here is offered in Cao and McGuire (2003), Batata (2004), McWilliams, Hsu and Newhouse (2012), Newhouse et al. (2013), Brown et al. (2014), Polyakova (2014), Carey (2016) and Shepard (2016).

[5]These results also complement the growing literature on demand for insurance. Related works include Nosal (2012) and Miller et al. (2014) for Medicare Advantage and Abaluck and Gruber (2011), Ketcham et al. (2012), Marzilli Ericson (2014), Ketcham, Lucarelli and Powers (2014), Abaluck and Gruber (2016), Ho, Hogan and Morton (2017), Polyakova (2014), Ho, Hogan and Morton (2017) and Heiss et al. (2016) for Part D. Our analysis is also related to the studies on inertia in employer sponsored health insurance (Handel and Kolstad, 2015), in auto insurance (Honka, 2014) and in pension plans (Handel and Kolstad, 2015).

[6]Our own analysis on Part D is reported in the web appendix for completeness.

Finally, our micro-level evidence on how different groups of consumers are differentially affected by the 5-star SEP is a clean example of the distributional consequences of a recent Medicare reform. Due to its size and organization, the question of the distributional impact of Medicare has received considerable attention in the literature (see, for instance, Bhattacharya and Lakdawalla (2006), McClellan and Skinner (2006) and Duggan, Starc and Vabson (2016)). In these studies, quantifying the insurance value of Medicare plays a key role in assessing its distributional impacts. In this respect, our findings reveal how even a "small" reform affecting directly just 5-star plans is able to trigger multiple changes in Medicare Advantage, by inducing both supply and demand responses.

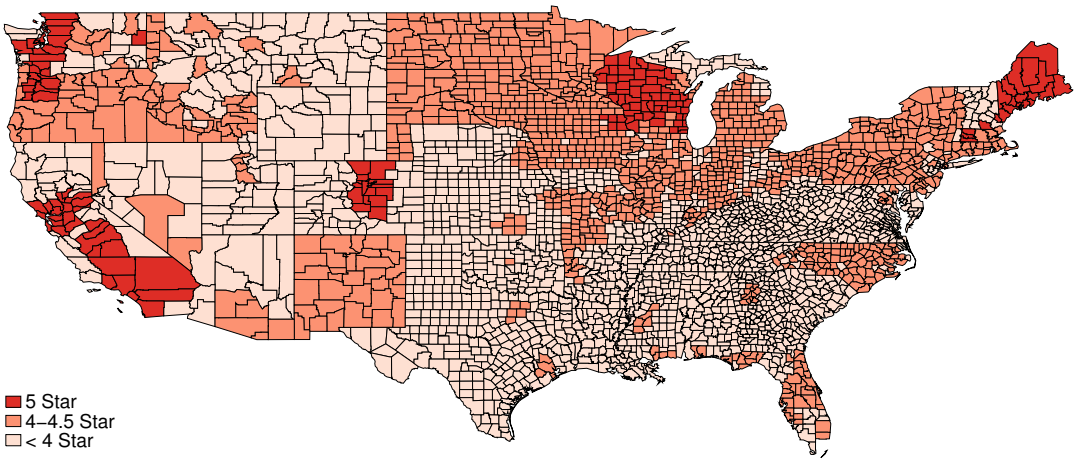
II Institutions: Medicare Open Enrollment Periods

The Medicare system consists of a series of interlinked programs aimed chiefly at those aged 65 or older in the US. Traditional Medicare (TM) is composed of Medicare Part A, covering inpatient hospital, skilled nursing, and some home health services, and Medicare Part B, covering physicians' services, outpatient care, and durable medical equipment. This study focuses on the privately provided programs that coexist with TM: Medicare Advantage and Part D. In both programs, private insurers offer a menu of plans to Medicare beneficiaries: Medicare Advantage plans are an alternative to TM and so must cover all Medicare Part A and B benefits (except hospice care), but can also offer additional benefits. Part D plans complement TM by covering prescription drugs. The two programs are closely connected in many ways, the most evident being that almost all Medicare Advantage plans also include Part D benefits. These latter plans will be denoted below as MAPD. As an alternative to MAPD, enrollees opting for TM, but who want to access the (voluntary) Part D program can purchase stand-alone Prescription Drug Plans (PDP).

Both MAPD and PDP offer one-year, renewable coverage coinciding with the calendar year. The open enrollment period (OEP) is the window of time during which people can enroll in these plans. It typically spans from October to December of the year before the coverage period. Although enrollees are generally required to stay on the same plan for the entire year of coverage, exceptions to the OEP exist. Special Enrollment Periods (SEPs) permit enrollees to change plans when certain circumstances occur. The most common SEPs involve individuals turning 65 during the coverage year, changing residency or entering "low-income enrollee" status. Starting in 2012, an additional SEP was introduced: people eligible for Medicare residing in an area where one or more 5-star Medicare Advantage or Part D plan is offered can switch from their plan (or from TM) to a 5-star plan during the coverage

year, with the new coverage starting on the first day of the month following the enrollment request.[7] We will refer to this reform as the 5-star SEP. The 5-star plans cannot deny enrollment. Beneficiaries can use this SEP only once per year and can also use it to switch from one 5-star plan to a different 5-star plan. To promote this policy, CMS has extensively advertised this new SEP rule in its communications to potential enrollees.[8]

Figure 1: Highest Rated MAPD Contract by County - Year 2012



Heat map: darkest colors indicate counties where the highest-rated MAPD has an higher star rating.

In contrast to all other Medicare SEPs, the possibility of a within-year plan switch is entirely dependent on a 5-star plan being offered in the enrollee's area of residency. This "area" differs between Medicare Advantage and Part D plans: for MAPD it corresponds to a county, while for PDP it corresponds to one of the 34 macro-regions partitioning the US. In 2012 and 2013, nearly 180 counties belonging to 17 different states had at least one 5-star MAPD, while only 2 regions had a 5-star PDP (New York and a macro region formed by 7 midwest states). Given the importance of cross-market variation for identifying the effects of the 5-star SEP, the remaining part of this study will focus exclusively on Medicare Advantage, leaving the analysis for Part D to the web appendix. Figure 1 illustrates the spatial pattern of 5-star MAPD offerings. While not present in the South, 5-star MAPD are present in all other regions. The heat map also reveals the location of counties whose highest rated MAPD were either 4 or 4.5 stars.

In the post-reform period, 7 insurers offer 5-star plans. The main ones are Kaiser, Hu-

[7]Although not directly included among the provisions in the PPACA, the 5-star SEP is linked to it as CMS created it through its statutory authority - Section 1851(e)(4)(D) of the Social Security Act - as part of its efforts to bolster plan quality (see Layton and Ryan (2015) and Li and Doshi (2016)).

[8]See, for instance: `https://www.medicare.gov/sign-up-change-plans/when-can-i-join-a-health-or-drug-plan/5-star-special-enrollment-period`.

mana and Group Health.[9] The case of Kaiser offers another reason why it is important to keep analysis of the 5-star SEP for Medicare Advantage and Part D distinct: switching plans in response to a health shock is inherently different and less complicated for a PDP than for a MAPD plan. In the case of switching from a different provider to Kaiser, an enrollee would need to change the whole network of primary, secondary, and hospital care. It might also imply changing how the enrollee relates to healthcare provisions, as Kaiser extensively uses care pathways and electronic medical records. While the case of Kaiser is extreme, it is clear that for some enrollees changing care providers might be undesirable, even though they might benefit from some of the high-quality features provided by 5-star plans.

III Empirical Strategy

Our strategy for identifying the effect of the 5-star SEP on plan enrollment is based on a difference-in-differences (DID) approach. For MAPD plans, this strategy exploits the fact, documented in Figure 1, that 5-star contracts are offered in only a subset of US counties. We consider all contracts that achieved a 5-star rating in the period 2012-2013 as the DID treatment group (dark red areas in in Figure 1) and all contracts that achieved a rating of 4 or 4.5 stars in the same period and are offered in counties without any 5-star contracts as the control group (light red areas in in Figure 1).

Specifically, the regression model that we estimate is:

$$Y_{ict} = \beta D_{it}^{5S} + \alpha_c + \gamma_t + \delta_i + \varepsilon_{ict}, \tag{1}$$

where i indicates the contract, c the county and t the year. The outcome variable is one of the three variables described in the previous section (*(i)* the within-year change in enrollment, *(ii)* the across-year change in enrollment, and *(iii)* the plan average risk score). The coefficient of interest is β, the effect on the dependent variable of a dummy equal to one for 5-star contracts after 2011, conditional on fixed effects for the county (α_c), time (γ_t) and contract (δ_i). Various extensions are presented below.

In an ideal scenario, this identification strategy allows the causal effects of the 5-star SEP to be estimated through the random assignment of counties between treatment and control groups. Clearly, however, the observational data that we use fall short of this ideal scenario, and so there are challenges to interpret β as the causal effect of the policy change. As usual in any DID study, the first and foremost concern is to select an adequate control group. In

[9] In contrast, only 2 insurers (one being Humana) offer 5-star PDP in the same period.

our setting, 4- and 4.5-star contracts offered in counties that do not have any 5-star plans are a natural control group. First of all, since the regulation separates the geographical markets, a benefit of the proposed DID strategy is that, by selecting the treatment and control groups from different counties, it avoids contamination issues. Furthermore, in the period analyzed, the CMS payment demonstration made the contracts in the control group face similar financial incentives to those in the treatment group, as all payments linked to the star rating were similar for these two groups of plans (Decarolis and Guglielmo, 2017).

Although Figure 1 reveals that the 5-star plans are scattered across many different counties, this clearly does not ensure their assignment to counties is random. At least two problems related to selection can emerge: ratings could be subject to manipulations, implying self selection by the insurers; or, even absent manipulations, counties in the treatment and control groups might differ.

Regarding the first problem, we note that it is hard for insurers to perfectly control their rating. This is due to the institutional features of the Medicare rating system. In particular, the star rating is derived from a combination of several (about 50) individual measures.[10] The use of such a large number of measures, together with the fact that both the exact set of measures and the scoring formula change from year to year, implies that insurers do not have full control over their rating. Furthermore, details about the timing of the measures are crucial for understanding why the 5-star SEP should not trigger rating manipulations for 2012 and 2013. Several of these measures enter with a two-year lag. Since insurers must define their plan offerings in June of the year before the enrollment and since the 5-Star SEP was announced in November 2010, any action aimed at altering the star rating would not take effect before the 2014 enrollment year. This fact is also consistent with the fact that the 2012 and 2013 offering of 5-star plans remained nearly unaltered relative to 2011 in terms of counties served and insurers involved.

Graphical evidence in support of the claim of no rating manipulation is offered in Figure 2. Underlying the discrete scores (appearing in 0.5 increments) that CMS discloses to enrollees and that determine the applicability of the 5-star SEP, there is a continuous measure which summarizes multiple indicators.[11] 5-star plans are those whose overall score is at least 4.75, while 4.5-star plans have an overall rating below 4.75, but above 4.25. If we look at this continuous measure in Figure 2, two elements are suggestive of the adequacy of the proposed

[10] In Table A.1, we report the 2 areas (managed care and prescription drugs) and the associated 9 domain measures grouping the 47 individual measures used to determine the 2012 scores. These measures are from different sources – from survey to call center and administrative data – and at different timings.

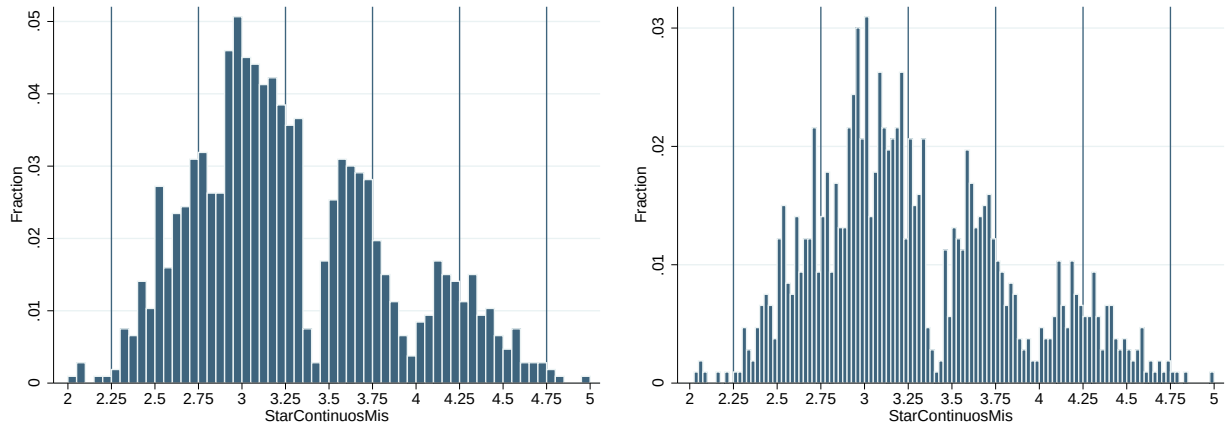
[11] Since CMS does not disclose the continuous measure, this remark is based on the continuous summary score measure that we constructed by combining the individual measures and the period-specific aggregation rules. We successfully match the CMS discrete score for 95 percent of the 1,284 contracts in 2011 and 2013.

identification strategy. First, there is no clear jump in the plan density at the relevant cutoff points, either before or after the reform.[12] Second, most 5-star plans fall short of having an overall continuous score of 5, reaching a score not much higher than 4.75. This is thus reassuring regarding their comparability to lower-rated plans.

Figure 2: Distribution of Star Rating across Contracts

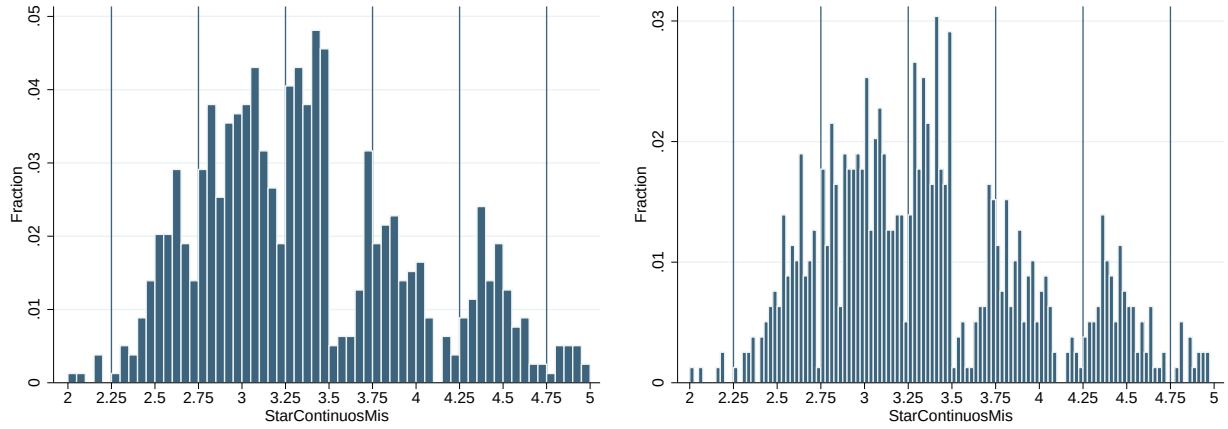
(a) Years 2009-2011 - Bin 0.05

(b) Years 2009-2011 - Bin 0.025



(c) Years 2012-2013 - Bin 0.05

(d) Years 2012-2013 - Bin 0.025



Regarding the second selection problem, counties in the proposed treatment and control groups might differ irrespective of the lack of rating manipulations. Our strategy for addressing this issue involves controlling for observable differences. Furthermore, to the extent that we can control for both fixed and time-varying unobservables, we gradually expand the model specification to incorporate both contract fixed effects and linear time trends,

[12]This is further supported by McCrary tests reported in Figure A.1 in the appendix. Moreover, while the distribution in Figure 2 is not smooth and appears to have three modes, this is due to the scoring formula and, specifically, to the fact that, for some measures, plans earn points only if they are above certain thresholds.

separately for treatment and control counties.[13] As discussed in the next section, when presenting the summary statistics in Table 1, treatment and control groups do indeed differ along several observable characteristics, such as size of the enrollment base and features of the enrollment pool. Hence, to the extent that the selection into the treatment state is based on these observable characteristics, we can address this threat to identification. Therefore, as a robustness check, we also present a matching DID strategy. In this case, the control group observations will be selected to match the characteristics of the treatment group.

IV Data

The analysis combines several data sources. In the first part of the analysis, we focus on plan- and contract-level data, while in the second we exploit claims-level data.

The data for the first part of the analysis are publicly available data from CMS. In particular, we obtained monthly enrollment data for the years 2009-2013 at the plan level, as well as plan characteristics and risk scores (both at the yearly level). Also from the CMS files we obtained the scores that each contract received on each individual measure, which we used to compute the continuous score. Table 1 shows summary statistics for these data. The three main outcome variables that we analyze are: *(i)* the within-year change in enrollment, *(ii)* the across-year change in enrollment, and *(iii)* the plan average risk score. The first variable is calculated as the difference in the contract enrollment in the last and first month of the year (i.e., $Enrollment_{12/t} - Enrollment_{1/t}$, with j/t indicating the j-th month of the year). It captures changes in plan enrollment within-year, and thus it measures the most direct effect that the policy produces in terms of increased within-year plan switches.

The second outcome variable considers the possibility of plan switching across years. We calculate it as the difference in the contract enrollment in two consecutive years. More precisely, it is calculated as $Enrollment_{1/t} - Enrollment_{12/t-1}$. This variable can capture a strategic response by consumers, namely, greater plan switching during the regular open enrollment period driven by the possibility of switching to a 5-star plan later. The third outcome variable is a proxy for the plan's risk pool. More precisely, we use the mean contract risk score, available from CMS at the yearly level and separately for the managed care (Part C) and prescription drug (Part D) components of the MAPD plans. The risk score is the key statistic mapping how enrollment composition impacts expected plan costs. In the

[13] Our unit of observation is a contract-year-county. In most of the control counties, we observe multiple contracts with 4-4.5 stars. To a lesser extent, we observe multiple contracts also in the treated counties (i.e., we observe 7 counties for which there were more than one 5-star plan in either 2012 or 2013).

final part of our analysis, we will look at the demographic characteristics of the switchers to better understand what drives the findings on plan-level risk scores. The summary statistics are immediately suggestive of interesting patterns in the data. In particular, we see that the within-year change in enrollment into treatment plans increases after the 5-star SEP. This is not the case for the control group.

In addition to the three outcome variables described above, the dataset allows us to observe a rich set of plan characteristics, as described in Table 1.[14] This allows us to observe changes in the plan offerings following the 5-star SEP. Table 1 reveals that premiums tend to decline after the reform more for the treatment than for the control group. Moreover, those features valued by the least healthy enrollees worsen (as illustrated for, instance, by the increase in the maximum out of pocket). For both premiums and generosity, the evidence in (Decarolis and Guglielmo, 2017) indicates that these changes do indeed reflect a strategic response by insurers to the 5-star SEP. This implies that our analysis below is best interpreted as an assessment of how enrollment responded jointly to the 5-star SEP and plan design changes.

The analysis involving claims-level data, instead, uses the NBER - Medicare Part D Research Identifiable Files. These data are a 20 percent random sample of all beneficiaries. For each beneficiary, we observe information on age, sex, place of residence, health conditions, and plan enrollment. To study within-year switching to 5-star MAPD behavior, we focus on those enrollees residing in counties where 5-star MAPD were offered in 2012 or 2013. The resulting sample has 2.4 million enrollees for 2012 and 2.5 million enrollees for 2013. In each year, about 0.25 percent of these enrollees switch to a 5-star MAPD during the year under

[14]In particular, the description of each variable appearing in the table is as follows: "Tot. Enrollment" is the contract enrollment measures as January. "Change Enrollment Dec.-Jan" is the change in enrollment from January to December. "% Change Enrollment Dec.-Jan" is the percentage change in enrollment from January to December. "Premium Part C" is the annual premium for Part C. "Premium Part D" is the annual premium for Part D. "In Network MOOP" is the maximum out-of-pocket expenditure for in-network service, excluding Part D drugs (we observe it starting from 2011). "Deductible Part D" is the maximum annual amount of initial out-of-pocket expenses for Part D drugs. "N. Top Drugs" is the number of top drugs included in the plan formulary. "N. Unrestricted Drug" is the number of drugs without restriction on utilization included in the plan formulary. "Risk Score Part C" is the average risk score measure for Part C coverage. "Risk Score Part D" is the average risk score measure for Part D. "Part C OOPC Excellent (Poor)" is the average yearly out-of-pocket for individuals with excellent (poor) heath status for Part C coverage. "Drug OOPC Excellent (Poor)" is the average yearly out-of-pocket for individuals with excellent (poor) heath status for Part D coverage. "Health Care Quality" is a star rating (1-5) for member's evaluation of health care quality (CAHPS). "Customer Service" is a star rating (1-5) for the ability of the health plan to provide information or help when members need it (CAHPS). "Drug Access" is a star rating (1-5) for the ease of getting prescriptions filled when using the plan (CAHPS Survey). "Tot. Enrollment", "Change Enrollment Dec.-Jan", "% Change Enrollment Dec.-Jan', "Health Care Quality", "Customer Service" and "Drug Access" are measured at the contract level, all other variables are measured at the plan level and aggregated at the contract level as (enrollment) weighted averages.

Table 1: Descriptive Statistics

	Control				Treament			
Pre Reform Sample: Years 2009-2011	Mean	s.d.	Median	N	Mean	s.d.	Median	N
Tot. Enrollment	1338.7	4176.5	196.3	4796	7129.7	17910.4	888	409
Change Enrollment Dec.-Jan.	92.38	378.3	27	4796	386.0	863.7	117.5	409
% Change Enrollment Dec.-Jan.	0.350	0.743	0.147	4796	0.301	0.721	0.068	409
Premium Part C	497.3	467.0	435.5	4796	754.9	408.6	838.9	409
Premium Part D	333.9	210.9	348.8	4796	232.9	140.7	255.6	409
In Network MOOP	3838	1084.3	3400	1696	2781.4	604.8	2682	148
N. Top Drugs	95.20	5.973	94	4765	83.17	14.92	90	409
N. Unrestricted Drug	532.6	130.5	520	4765	641.4	102.4	641	409
Deductible Part D	44.59	94.41	0	4796	21.34	61.12	0	409
Risk Score Part C	0.965	0.229	0.908	4796	0.925	0.109	0.965	409
Risk Score Part D	0.934	0.111	0.915	4796	0.882	0.044	0.880	409
Part C OOPC Excellent	823.2	197.7	807.9	4425	800.2	110.8	801.2	409
Part C OOPC Poor	1763.5	529.9	1730.2	4425	1632.6	393.2	1643.3	409
Drug OOPC - Excellent	592.2	145.8	597.2	4425	720.7	151.0	777.3	409
Drug OOPC - Poor	1974.9	645.2	1972.9	4425	2455.9	687.5	2552	409
Health Care Quality	4.048	0.788	4	4658	4.748	0.435	5	397
Customer Service	3.809	1.128	4	3660	4.698	0.492	5	397
Drug Access	4.163	0.838	4	4654	4.952	0.214	5	397

	Control				Treament			
Post Reform Sample: Years 2012-2013	Mean	s.d.	Median	N	Mean	s.d.	Median	N
Tot. Enrollment	1265.5	3753.6	236	4300	8636.0	21040.4	1320	263
Change Enrollment Dec.-Jan.	55.68	228.7	13	4300	569.6	1364.1	122.1	263
% Change Enrollment Dec.-Jan.	0.133	0.327	0.066	4300	0.101	0.110	0.0674	263
Premium Part C	427.7	423.1	374.3	4300	632.1	349.8	647.1	263
Premium Part D	310.3	223.8	306	4300	213.1	165.8	210.4	263
In Network MOOP	3755.6	991.6	3400	4026	3362.9	1124.3	3400.0	263
N. Top Drugs	87.05	3.757	88	4274	89.31	3.132	88	263
N. Unrestricted Drug	415.2	123.5	409.4	4274	415.6	75.30	389	263
Deductible Part D	40.54	89.19	0	4300	30.68	73.59	0	263
Risk Score Part C	0.953	0.196	0.900	4299	0.907	0.0913	0.930	263
Risk Score Part D	0.909	0.0967	0.893	4299	0.857	0.043	0.854	263
Part C OOPC Excellent	979.0	192.5	998.2	4033	989.8	121.2	1009.2	263
Part C OOPC Poor	2225.2	412.7	2286.9	4033	2172.4	372.3	2121.5	263
Drug OOPC - Excellent	624.8	130.9	618.0	4033	629.7	207.5	524.8	263
Drug OOPC - Poor	2399.0	546.6	2367.9	4033	2312.6	989.2	2163.6	263
Health Care Quality	4.236	0.622	4	4267	4.817	0.387	5	263
Customer Service	3.926	1.033	4	4219	4.319	1.225	5	263
Drug Access	3.908	1.015	4	4272	4.669	0.929	5	263

Note: The unit of observation is Contract/County/Year. The top panel includes observations from 2009 to 2011. The bottom panel includes observations from 2012 to 2013. On the left there are statistics for observations in the control group: the contracts offered in the 1084 counties with no 5-star plans, but at least one 4 or 4.5-star plan in 2012 or 2013. On the right there are statistics for observations in the control group: the contracts offered in 160 counties with at leas one 5-star plan in 2012 or 2013.

the 5-star SEP.[15]

[15]We observe 5,502 switching cases in 2012 and 5,667 cases in 2013. To ensure these are all due to the 5-star SEP, we had excluded from the sample individuals changing residency or turning 65 during the year. We exclude from the sample those individuals who cannot switch because they are already enrolled in a

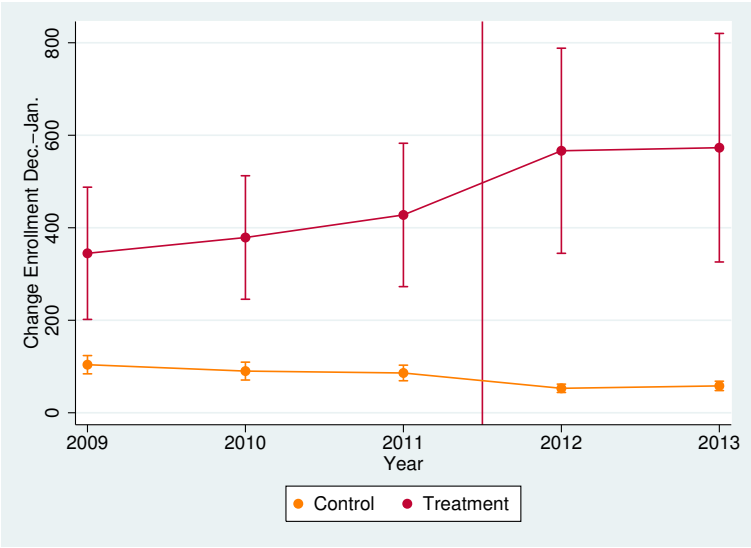
V Results

This section presents the results separately in four parts. The first three sections look at contract-level outcomes, while the fourth presents evidence from the claims-level data.

A. Within-Year Enrollment for MAPD

The evolution of the average within-year enrollment change is described in Figure 3 for both treatment and control plans. In line with the statistics in Table 1, it shows both a relatively large increase for the treatment group after the 5-star SEP (represented by the straight, vertical line) and a lack of any increase for the control group.

Figure 3: MAPD Contracts - Within-Year Enrollment Change



Notes: Evolution of the within-year enrollment variable for both treatment and control contracts.

5-star MAPD plan as well as those who never purchase any drug. We also drop individuals with missing values for race in the Medicare files.

Even before 2012, there is a trend for growth in the treated group, compared to a declining path for the control group. Although for both groups these year-to-year changes are not statistically significant, thus limiting potential bias in the estimate of β, we will also report estimates including group-specific linear time trends in the DID model specification.[16] Nevertheless, since the statistical evidence in favor of a differential trend is rather weak and since it is conceivable that trends might obfuscate the effects of the 5-star SEP if consumers learn over time to exploit the new enrollment flexibility, we will also report estimates from a more parsimonious model without time trends and describe both sets of estimates. As discussed below, the main findings for the contract-level analysis will remain qualitatively the same.

Table 2 displays our DID estimates for the within-year enrollment in MAPD. The dependent variable is thus the within-year enrollment change both in levels (Columns 1-4) and in percentage terms relative to January enrollment (Columns 5-8). We estimate four models. Odd numbered columns include county and year fixed effects, and even numbered columns add contract fixed effects. Columns 3, 4, 7 and 8 add a linear trend at state/treatment level. Panel A reports the estimates for the baseline sample: the treatment group has 5-star contracts in 2012 or 2013, while the control group contains 4 or 4.5 star contracts in 2012 or 2013 in counties without any 5-star contracts. The next panel reports the robustness check involving a matched-DID estimator.

The estimates in Panel A show that the 5-star SEP has a large and statistically significant effect on the within-year change in enrollment. The effect reported in columns 1 and 2 implies that the number of enrollees increases on average by 225-235 enrollees. This effect is quite substantial, if, for instance, we compare it to an average value of the dependent variable in the pre-treatment period of 386 enrollees. When including time trends, the effect is still present, but its magnitude is attenuated. Columns 5-8 report analogous estimates for the percentage enrollment change. This variable allows the enrollment changes to be normalized by the existing enrollment base. The estimates that we obtain range from 7% to 9% in the more parsimonious specifications and from 15% to 16% when including time trends.

To assess the sensitivity of our estimates to the choice of control group, in Panel B, we use a matched-DID estimator by constructing a sample of comparable contracts through propensity score matching. In particular, we use an extensive list of socio-economic, demographic and health indicators to predict the probability that a county has a 5-star contract in the 2012-13 period. Then, we restrict the control group to those contracts in counties belong-

[16]The presence of an upward trend for the treatment group, can be explained by several factors. CMS has been strongly advertising star ratings as a measure of quality, possibly impacting the evolution of enrollment over time. There is no evidence of trends for the other outcome variables used below.

ing to the common support of the propensity score between the treatment and the control groups.[17] The estimates obtained are similar in terms of both magnitude and significance to those in Panel A. Not all coefficients of the matched-DID, however, lie within the 95 percent confidence interval of those in Panel A. In particular, the matched-DID indicates a larger percentage increase, amounting roughly to a 20% effect, when including trends. While these estimates are likely the preferable ones as they fully exploit the richness of the data, we take the Panel A estimate of a 15% effect as a more conservative estimate.

Table 2: MAPD Contracts - Within-Year Enrollment Change

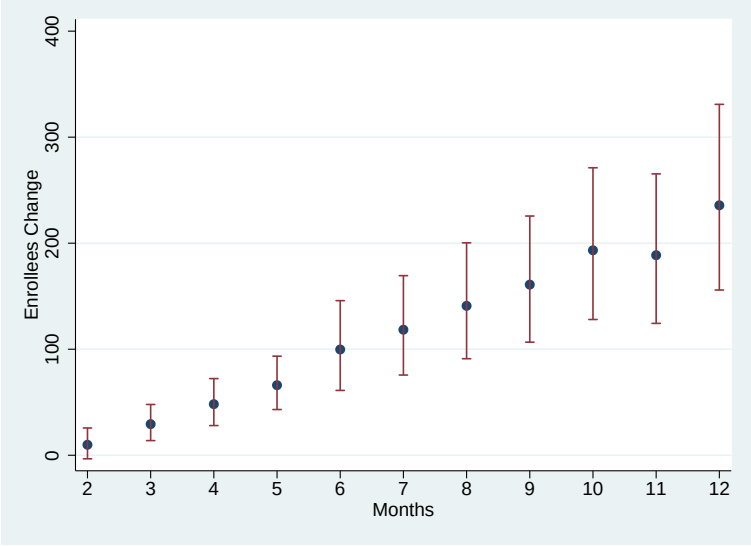
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Baseline Sample							
	Dec.-Jan. Enrollment Change				Dec.-Jan. Enrollment % Change			
5 Star	224.327***	235.741***	86.860**	86.131**	0.074*	0.089**	0.165**	0.155**
	(50.125)	(48.533)	(39.527)	(37.405)	(0.044)	(0.042)	(0.075)	(0.070)
Observations	9,768	9,768	9,768	9,768	9,768	9,768	9,768	9,768
R-squared	0.553	0.620	0.564	0.630	0.196	0.281	0.229	0.313
	Panel B Matched Sample							
	Dec.-Jan. Enrollment Change				Dec.-Jan. Enrollment % Change			
5 Star	145.972***	153.032***	63.519**	60.888**	0.089*	0.099**	0.219***	0.202***
	(25.732)	(25.236)	(25.683)	(24.662)	(0.046)	(0.046)	(0.079)	(0.075)
Observations	7,616	7,616	7,616	7,616	7,616	7,616	7,616	7,616
R-squared	0.461	0.548	0.475	0.562	0.185	0.272	0.220	0.308
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES	YES	YES	YES
Contract FE	NO	YES	NO	YES	NO	YES	NO	YES
Time Trend	NO	NO	YES	YES	NO	NO	YES	YES

Notes: The table reports the DID estimates of the effect of the 5-star SEP. The outcome variable is the difference in the contract enrollment between December and January (of the same year) calculated either in levels (first four columns) or as a percentage (latter four columns). The four model specifications considered for each dependent variable differ in the set of controls used, as reported in the block at the very end of the table. Panel A reports the estimates for the baseline sample: treatment group contracts with 5-star in 2012 or 2013; control group contracts with more 4 or 4.5 star in 2012 or 2013 in counties without 5-star contracts. Panel B reports the estimates for a sample matched using a propensity score. The probability that a county has a 5-star contract is estimated over a range of socio-economic, demographic and health indicators of the counties. Only the county on common support of the propensity score between the treatment and the control groups are included. Standard errors in parentheses clustered at county level *** p<0.01, ** p<0.05, * p<0.1.

Finally, it is informative to know in which month of the year enrollees use the SEP. Thus, we consider complementing the above estimates of the December minus January enrollment change with analogous estimates for the other months preceding December. In Figure 4, we plot the estimates obtained for the same specification as in model (2) of Table 2. The

[17] We tried various specification for the propensity score and results were broadly comparable to those in Panel B. Further details as well as the probit estimates are shown in Table A.1 and A.2 in the web appendix.

Figure 4: MAPD Contracts - Monthly Enrollment Change Relative to January



Notes: Estimate of the effect of the 5-star SEP on within year enrollment change, calculated at all months. The last value on the horizontal axis (12) represents the Dec. minus Jan. enrollment, the next value (11) represents the Nov. minus Jan. enrollment, and so on until (2), which represents the Feb. minus Jan. enrollment. The value for the Dec. minus Jan. enrollment is the same reported in the second column of Panel A in Table 2. All other estimates are obtained using the same specification.

effect on enrollment of the SEP appears to increase linearly over time up until October and then it flattens out. Thus enrollees seem to use the new SEP uniformly over most of the year.

B. Across-Years Enrollment for MAPD

Next we explore the behavior of consumers across years. In Table 3, we therefore repeat the previous analysis using as a dependent variable the change in enrollment across years.

The effect of the 5-star SEP is, however, ex ante ambiguous in this case. A decrease in demand for 5-star plans is compatible with consumers acting strategically: that is, enrolling in cheap, low-coverage non 5-star plans, but with the intention of switching to a more expensive 5-star plan if hit by a health shock during the year. The previous estimates in Table 2 indicate that within-year switches do occur. But, this is not enough to also imply that consumers will act strategically in their choice of switching plan across years. In fact, increases in enrollment in 5-star plans across years might be driven, for instance, by their enhanced promotion by CMS. The findings in Decarolis and Guglielmo (2017) also suggest that their lower post-reform premiums could have bolstered demand for such plans.

Furthermore, inertia in plan choice might imply a lack of changes in across-year enrollment. Empirical evidence by Handel (2013) and related work has revealed how relevant inertia is for health insurance plan choices. But it is noteworthy that, in our setting, inertia

Table 3: MAPD Contracts - Across-Year Enrollment Change

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Baseline Sample							
	Jan.-Dec. Enrollment Change				Jan.-Dec. Enrollment % Change			
5 Star	-2.072	0.272	21.254	22.616	0.044	0.039	0.186***	0.204***
	(15.362)	(15.002)	(26.370)	(24.777)	(0.037)	(0.033)	(0.054)	(0.052)
Observations	8,823	8,823	8,823	8,823	8,823	8,823	8,823	8,823
R-squared	0.079	0.121	0.088	0.130	0.143	0.219	0.148	0.225
	Panel B: Matched Sample							
	Jan.-Dec. Enrollment Change				Jan.-Dec. Enrollment % Change			
Star 5	8.495	10.458	8.776	8.914	0.065	0.057	0.243***	0.261***
	(13.164)	(12.988)	(19.117)	(18.275)	(0.040)	(0.036)	(0.056)	(0.055)
Observations	7,094	7,094	7,094	7,094	7,094	7,094	7,094	7,094
R-squared	0.138	0.190	0.167	0.220	0.118	0.204	0.124	0.212
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES	YES	YES	YES
Contract FE	NO	YES	NO	YES	NO	YES	NO	YES
Time Trend	NO	NO	YES	YES	NO	NO	YES	YES

Notes: The table reports the DID estimates of the effect of the 5-star SEP. The outcome variable is the difference in the contract enrollment between January and December (of consecutive years) calculated either in levels (first four columns) or in percentage (latter four columns). Panel A reports the estimates for the baseline sample: treatment group contracts with 5-star in 2012 or 2013; control group contracts with more 4 or 4.5 star in 2012 or 2013 in counties without 5-star contracts. Panel B reports the estimates for a sample matched using a propensity score. The probability that a county has a 5-star contract is estimated over a range of socio-economical, demographic and health indicators of the counties. Only the county on common support of the propensity score between the treatment and the control groups are included. Standard errors in parentheses clustered at county level *** p<0.01, ** p<0.05, * p<0.1.

might interact in complex ways with the 5-star reform. In the presence of present-biased beneficiaries, it might drive a drop in 5-star plan enrollment across years due to procrastination (as argued in Madeira (2015)). Furthermore, while inertia is sometimes observable through the comparison of the choices of new and continuing enrollees,[18] this is not necessarily the case for the 5-star SEP. In fact, inertia could lead even new enrollees to voluntarily ignore the possibility of "gaming" the system: selecting a plan during the open enrollment period with the idea of switching to a 5-star plan during the coverage period might be undesirable for those consumers who are aware that their inertial behavior is driven by features like search costs, switching costs, or psychological costs. As discussed earlier, these costs might be severe since switching MAPD typically implies multiple changes (insurer, hospitals, doctors, and - possibly - more). Below we do not attempt to distinguish between the distinct sources and consequences of inertia or between new and continuing enrollees, but merely to analyze

[18]Handel (2013) proposes a test for inertia based on the comparison of new and continuing enrollees.

whether in aggregate the total enrollment in 5-star plans changes across years in response to the 5-star SEP.

The estimates in Table 3 reveal that demand for 5-star plans did not decline across years: there is no specification that results in a negative and significant effect. Statistical significance is achieved only for the estimates involving percentage increase and, within these cases, only for the specifications including time trends (models 7 and 8). This finding emerges for both the baseline estimates (Panel A) and the matched-DID (Panel B). Since we tend to prefer the more complete specifications of models 7 and 8, we might conclude that there is evidence in favor of an increase in enrollment across years. However, contrary to the within-year demand estimates that systematically lead to very consistent estimates in terms of sign and significance, the lack of stability in the across-years demand estimates suggest caution should be taken in interpreting the finding as conclusive in terms of any positive effect on across-years demand.

In any case, all estimates indicate that any strategic consideration for consumers to leave 5-star plans was muted by the forces inducing a stronger demand. This finding suggests that the reform was successful in shifting enrollees to 5-star plans in a stable way. Although a positive coefficient can mechanically result from the combination of increased within-year switches in 2012 and the presence of plan switching cost, our estimates remain qualitatively identical if we rule out this channel by excluding 2013 data.

C. Part C and D Risk Scores for MAPD

The final set of results concern the effects of the 5-star SEP on the contracts' risk pools. The earlier results offer conflicting predictions on what effect we should expect. On the one hand, increased within-year enrollment might be a sign of worsening selection for 5-star plans. On the other hand, increased across-year enrollment could imply improved selection for 5-star plans, especially to the extent that it is driven by demand from relatively healthy enrollees attracted by lower premiums. This same force could also be the trigger for within-year plan switches driven by healthy enrollees looking for high-quality plans, irrespective of their reduced financial generosity.

The two dependent variables on which we focus are the yearly average MAPD risk scores that CMS releases separately for the two components of MAPD plans, Part C and D. Each one of these two measures is normalized to 1 for the average risk of a TM enrollee; the higher the risk score the higher the risk (and the expected cost) of the enrollee. The estimates are reported in Table 4.

Panel A in Table 4 presents the baseline estimates, separately for Part C (first 4 columns) and Part D (latter 4 columns). Both the model specifications and the construction of the

Table 4: MAPD Contracts - Risk Score Part C and D

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: Baseline Sample				
	Risk Score Part C				Risk Score Part D			
5 Star	-0.024***	-0.029***	-0.016***	-0.014***	-0.007***	-0.010***	-0.010***	-0.008***
	(0.005)	(0.004)	(0.005)	(0.004)	(0.002)	(0.002)	(0.003)	(0.002)
Observations	9,767	9,767	9,767	9,767	9,767	9,767	9,767	9,767
R-squared	0.349	0.949	0.354	0.953	0.349	0.930	0.354	0.935
				Panel B: 2012 Effect				
	Risk Score Part C				Risk Score Part D			
5 Star	-0.014***	-0.023***	-0.025***	-0.022***	-0.004**	-0.009***	-0.015***	-0.013***
	(0.004)	(0.004)	(0.005)	(0.003)	(0.002)	(0.001)	(0.003)	(0.002)
Observations	7,372	7,372	7,372	7,372	7,372	7,372	7,372	7,372
R-squared	0.355	0.954	0.361	0.959	0.363	0.937	0.368	0.942
				Panel C: 2013 Effect				
	Risk Score Part C				Risk Score Part D			
5 Star	-0.042***	-0.042***	-0.032***	-0.013	-0.013***	-0.013***	-0.020***	-0.012**
	(0.007)	(0.006)	(0.011)	(0.010)	(0.003)	(0.003)	(0.005)	(0.005)
Observations	7,600	7,600	7,600	7,600	7,600	7,600	7,600	7,600
R-squared	0.356	0.951	0.361	0.955	0.366	0.928	0.371	0.934
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES	YES	YES	YES
Contract FE	NO	YES	NO	YES	NO	YES	NO	YES
Time Trend	NO	NO	YES	YES	NO	NO	YES	YES

Notes: The table reports the DID estimates of the effect of the 5-star SEP. The outcome variable is the risk score for Part C (first four columns) and Part D (latter four columns). The four model specifications considered for each dependent variable differ in the set of controls used, as reported in the block at the very end of the table. Panel A reports the estimates for the baseline sample: treatment group contracts with 5-star in 2012 or 2013; control group contracts with more 4 or 4.5 star in 2012 or 2013 in counties without 5-star contracts. Panel B reports estimates from a sample without observation from 2013. Panel C reports estimates from a sample without observation from 2012. Standard errors in parentheses clustered at county level *** p<0.01, ** p<0.05, * p<0.1.

control group is identical to those described for Table 2 and 3. All the estimates in this panel show a negative and significant effect on both risk scores. The magnitude of the estimated coefficients is small, but not negligible. Relative to the summary statistics reported earlier, the estimates for the effect on Part C of the 5-star SEP roughly correspond to one fifth of a standard deviation of the dependent variable. The analogous figure for the Part D risk score is one fourth of a standard deviation.

The next two panels in Table 4 aim to assess whether the observed improvements in risk scores might be driven by the timing with which risk scores are measured. The measure that we use is a yearly average, and a potential worry is whether this variable is unable to capture

in a timely manner the high risk of those joining 5-star plans. However, the annual average risk score for a plan is built up by taking all of the individual-level risk scores and averaging them. Hence, when new enrollees join during year t, the risk scores of those enrollees will be factored into the year t average risk score.[19]

Nevertheless, there is a lag in updating the individual-level risk scores. In 2013, an individual's risk score is based on his health status (diagnoses) from 2012. Thus, if an enrollee who used to be healthy switches to a 5-star plan immediately after becoming sick, our measure might only be able to capture his higher risk a year after the switch. This feature means the current risk score system is inadequate for dealing with selection driven by within-year plan changes, and this problem is especially severe in Part C where no ex post adjustment measures (like the Part D risk corridors and reinsurance) exist.

Moreover, a more subtle problem could, in principle, involve new Medicare enrollees. First-time enrollees in Medicare have no diagnoses, so their risk scores are based on age/gender only and are not particularly indicative of health status. First-time enrollees in After they have been in Medicare for a full calendar year, their risk scores switch to being based on diagnoses instead. However, since new Medicare enrollees aren't actually affected by the reform as they could join any plan during any month of the year (as long as it is the first month they enroll), this should not be a concern for our analysis.

To account for these issues, we exploit the fact that we have two years of data since the inception of the policy and repeat the DID estimates iteratively dropping from the sample one of the two post-policy years. Our expectation is that, if the negative estimate in the risk score regressions is driven by a lag in how the score is recorded, we will likely find that using 2013 exclusively as the post-policy year should lead us to find less negative, if not positive, estimates relative to when we use only 2012 as the post-policy year. The new estimates are reported in the latter two panels of Table 4. In Panel B we drop 2013, while in Panel C we drop 2012. The findings are rather surprising. Both sets of estimates confirm the negative sign of the coefficient. Moreover, although the magnitudes are similar, there is a tendency for the Panel C estimates to be larger than those in Panel B. Hence, these results confirm that the risk pool of 5-star plans effectively improved.

D. Additional Evidence from Claims Data

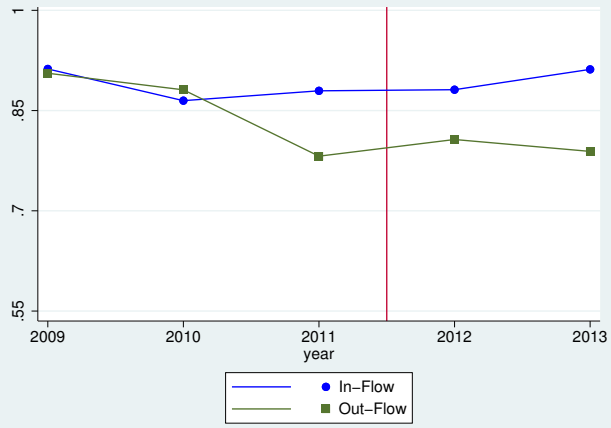
This last section analyzes the key question associated with the earlier findings of declining risk scores: is the lowered risk score in 5-star MAPD due to switchers who are healthier relative to the whole Medicare population or only relative to the risk pool of 5-star plans?

[19] We also know from Geruso and Layton (2015) that insurers are proactive in adjusting upward the risk score of their enrollees.

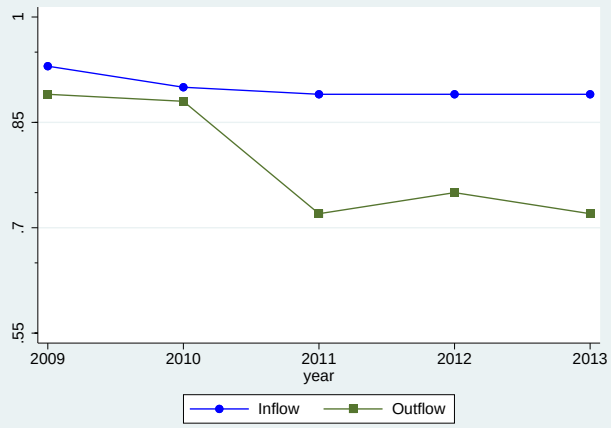
Figure 5 shows the average risk score separately for MAPD that lose enrollees and those that gain enrollees during the year.[20] The figure shows that there is a change between 2011 and the two previous years, which is likely the result of a revised approach to calculating risk scores starting in 2011. Regarding the effects of the 5-star SEP, instead, we see that both before and after the reform, switching tends to be from low-risk plans to high-risk ones.[21] Since 5-star plans were characterized by high risk enrollees prior to 2012, the patterns in Figure 5 are ambiguous as to whether the 5-star SEP produced adverse selection for the 5-star plans. They might have attracted the highest-risk enrollees from the non-5-star plans (adverse selection) who happen to be, however, lower risk than the average risk in 5-star plans. But they might also have attracted enrollees that are no more of a risk (no selection) or even healthier (advantageous selection) than those in non-5-star plans.

Figure 5: Average Risk Score of Contracts with Net Inflow or Outflow of Enrollees

(a) Risk Score Part C



(b) Risk Score Part D



To resolve this ambiguity, we resort to the CMS Part D claims-level data described in section 4. We estimate the probability of this type of switch through the following logit model:

$$Pr(Switch_{it}) = \Phi[\alpha + \sum_{z \in Z} \beta_z HealthStatus_{itz} + \sum_{j \in J} \gamma_j X_{itz} + \tau_t],$$

where i indexes the enrollee and t the year. Φ is the CDF of the logistic distribution. *HealthStatus* contains Z measures of the health conditions of enrollee i in year t. X contains

[20]The average is calculated by weighting contracts by their share of switchers in-flow or out-flow. The fact that both for out flow and in flow the average risk score is below 1 is explained by the fact that our analysis excludes the southern US regions, as illustrated in Figure 1, where risk scores tend to be higher. As discussed earlier, within-year switches occurring before 2012 are due to the presence of other SEP (see section 2).

[21]Switching in the pre 5-Star SEP period is driven by the presence of the other SEP listed in section 2.

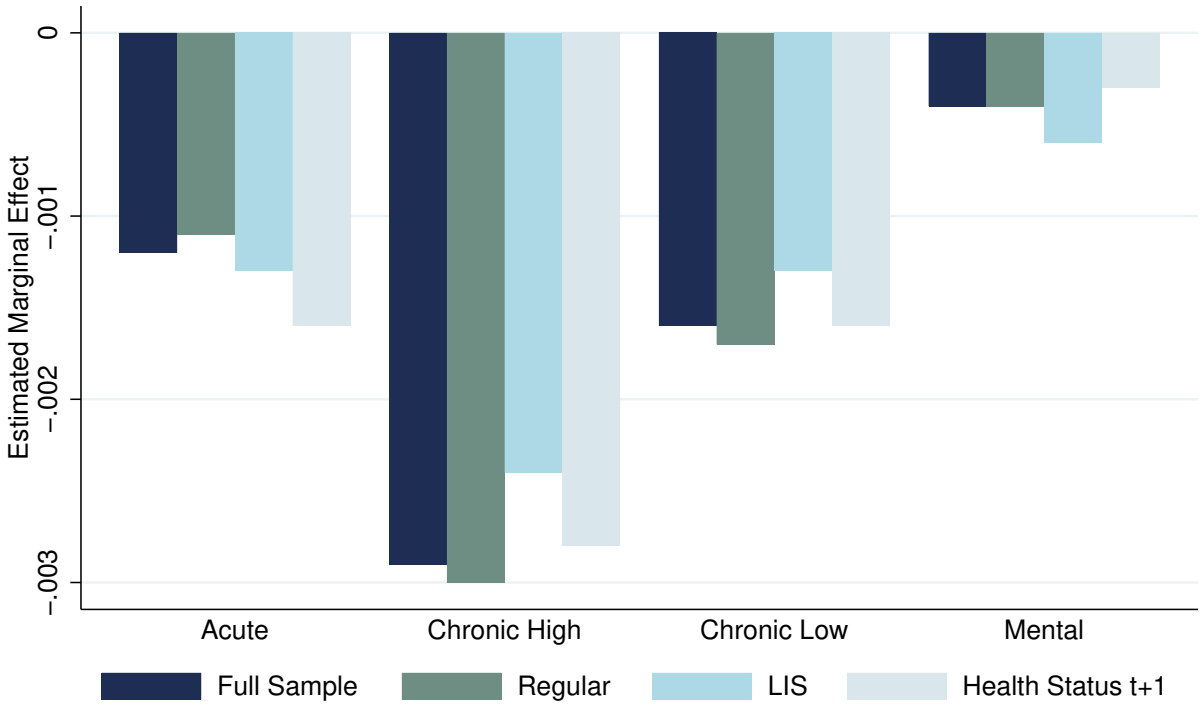
various additional controls that we will group in three main categories: *Demographics* (sex, age and race), *Financials* (current and last year OOPC) and *Programs* (indicator variables for whether in January of year t enrollee i is in MAPD, PDP or in TM without any Part D plan). We are particularly interested in the estimates of the β_z coefficients as they can provide direct evidence regarding the risk of switchers relative to non-switchers.

Although we cannot replicate exactly the CMS risk score measures used in the earlier section, the four variables that we use for *HealthStatus* capture most of the health conditions behind the determination of the risk scores. In particular, we consider four variables (*Acute High*, *Chronic Low*, *Chronic High* and *Mental*) which are constructed as follows. Each variable is a dummy variable for the existence of a flag for any of the relevant medical conditions in the chronic conditions component of the master beneficiary summary file. Together they act as a rough proxy of CMS' risk adjustment. *Acute High* accounts for any severe acute conditions such as heart attacks, strokes, or fractured hips. *Chronic Low* records the presence of chronic conditions that are not debilitating (asthma, diabetes, hyperlipidemia, etc). *Chronic High* indicates the existence of debilitating chronic conditions (osteoporosis, cancer, etc). Finally, *Mental* indicates Alzheimer and depression conditions. Since a flag is recorded even if there is only one event in the year triggering one of the diagnoses we consider, this implies that our measures are likely to capture any change in health status that could be also associated with switches to high-coverage, 5-star plans. The means (and standard deviations) for these dummy variables are around 0.8 (0.4) for *Chronic Low* and 0.7 (0.5) for the other three.

In Figure 6, we show the marginal effects for the logit regressions reported in Table 5. For each of the four health conditions, the figure shows the marginal effect estimated for four different samples. The bars in dark blue refer to the first sample, which is the baseline full sample of 2012 and 2013 switchers. All the four measures of adverse health conditions are clearly associated with a decline in the probability of switching. Moreover, the magnitude of the effect associated with *Chronic High* is about twice that of *Chronic Low*. The lower propensity to switch during the year to a 5-star plan among enrollees in worse health is also confirmed across the three subsamples represented with different colors in Figure 6.[22] This evidence is most obviously compatible with the greater difficulties that enrollees suffering from severe chronic disease would face were they to change provider network. As argued in Decarolis and Guglielmo (2017), insurers can enhance these difficulties by reducing the

[22]These are: regular enrollees (light green), LIS receivers (light blue) and the 2012 switchers only (grey). For the latter subsample, the model is estimated by replacing the concurrent *HealthStatus* measures with their values in 2013. The idea of this specification is to check whether the enrollees switching in 2012 are more likely to be those in worse health status the following year.

Figure 6: Heterogeneity Across Health Groups



financial generosity of their plans. But even more crucial for chronically ill patients might be the provider network. In this respect, the evidence in Shepard (2016), albeit in a different health insurance context, clearly shows the impacts on these enrollees of insurers' choice to leave out of their network "star" hospitals preferred by these enrollees.

Table 5 reports the logit estimates for different model specifications and sample restrictions. Model (1) includes only the *HealthStatus* measures, while the following three models gradually expand the specification to include *Demographics* (model (2)), *Financials* (model (3)) and *Programs* (model (4)). All models also include a constant and a dummy for 2013, both not reported in the table. Models (5) and (6) estimate the same specification of model (4) for two different subsamples: one excluding LIS enrollees (model (5)) and one including only LIS enrollees (model (6)). Finally, model (7) uses exclusively 2012 switching data, but replaces the concurrent *HealthStatus* measures with their values in 2013. Models (4)-(7) are those producing the marginal effects shown in Figure 6.

In addition to what has already been discussed concerning Figure 6, the estimates in Table 5 show a few interesting results. In particular, we find negative coefficients on both *black* race indicator and the two OOPC measures. The estimates are also indicative that switchers are more likely to originate from within the MAPD program rather than from the

Table 5: Logit Estimates for 5-Star SEP Switches

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Regular and LIS Enrollees				Regulars	LIS	Health_{t+1}
Health Status							
Acute High	-0.58***	-0.60***	-0.60***	-0.53***	-0.51***	-0.52***	-0.70***
	(0.051)	(0.05)	(0.05)	(0.05)	(0.06)	(0.08)	(0.08)
Chronic Low	-0.69***	-0.73***	-0.71***	-0.71***	-0.78***	-0.52***	-0.69***
	(0.026)	(0.03)	(0.03)	(0.03)	(0.03)	(0.05)	(0.04)
Chronic High	-1.44***	-1.41***	-1.44***	-1.29***	-1.35***	-0.95***	-1.25***
	(0.06)	(0.05)	(0.05)	(0.05)	(0.07)	(0.09)	(0.08)
Mental	-0.26***	-0.23***	-0.23***	-0.18***	-0.20***	-0.22***	-0.11*
	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	(0.07)	(0.07)
Demographics							
Female		-0.06***	-0.05***	-0.07***	-0.10***	-0.02	-.07**
		(0.02)	(0.02)	(0.02)	(0.02)	(0.05)	(0.03)
Age		0.02***	0.02***	0.02***	0.02***	0.02***	0.02***
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Black		-0.17***	-0.19***	-0.20***	-0.10**	-0.68***	-0.23***
		(0.04)	(0.04)	(0.04)	(0.04)	(0.08)	(0.06)
Latino		0.55***	0.51***	0.50***	0.17**	0.47***	0.41***
		(0.05)	(0.05)	(0.05)	(0.07)	(0.06)	(0.07)
Asian		1.07***	1.03***	1.06***	1.10***	.60***	1.00***
		(0.04)	(0.04)	(0.04)	(0.05)	(0.07)	(0.06)
Other		0.66***	0.63***	0.65***	0.65***	0.44***	0.62***
		(0.04)	(0.04)	(0.04)	(0.05)	(0.10)	(0.06)
Financials							
OOP			-0.28***	-0.30***	-0.28***	-0.44	0.00
			(0.08)	(0.08)	(0.08)	(0.69)	(0.00)
OOP_{lag}			-0.21***	-0.27***	-0.17***	-1.99***	-0.00***
			(0.05)	(0.05)	(0.05)	(0.46)	(0.00)
Programs							
PDP				-0.52***	-0.71***	-0.54***	-0.39***
				(0.04)	(0.05)	(0.07)	(0.06)
No Plan				-0.26***	-0.23***	0.55***	-0.34***
				(0.02)	(0.03)	(0.08)	(0.03)
Observations	4,934,656	4,934,656	4,934,656	4,934,656	4,125,297	809,359	2,211,384
Prob. Chi-Squared	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Logit regressions for the probability that an enrollee not enrolled in a 5-star plan in January of 2012 or 2013 switches during the year to a 5-star MAPD under the 5-star SEP. All regressions include a constant and a dummy equal to 1 if the year is 2013 and zero if it is 2012. For readability OOP and OOP_{lag} are rescaled by 1,000. Standard errors in parentheses clustered at county level *** p<0.01, ** p<0.05, * p<0.1.

PDP program or from TM without any Part D coverage. The positive estimate on *Age* and *Female*, instead, runs contrary to what would be expected under advantageous selection. Their magnitudes, however, are smaller if compared, for instance, to the effect of the *Black* indicator variable and, in the case of *Female*, the effect is not significant in model (6). The relevance of the subsampling results in models (5) and (6) derives from the fact that LIS enrollees have special rights to switch plan within the year. Although the data do not allow

us to separately identify the motive of the LIS request of plan switch, observing that the estimates are nearly identical for the two subsamples reassures us that our results are not driven by the mere presence of switches by LIS enrollees.

VI Conclusions

The 5-star SEP reform that, beginning in 2012, allowed Medicare enrollees to switch at any point in time to a 5-star rated plans is a rare example of a change in open enrollment rules. It therefore represents a valuable natural experiment to learn about the effects that these kinds of policies can produce and, hence, to what extent they can be used as a tool to guide health insurance markets toward socially desirable outcomes. In the context of Medicare, where as of 2017 more than 11 million beneficiaries were exposed to the effects of the 5-star SEP, this reform appears to have accomplished its intended effects of promoting enrollment into high-quality, 5-star plans without generating an adverse selection death spiral.

The analysis is based on a clean identification strategy exploiting the geographical distribution of plans with different star ratings in the years 2009-2013. Its focus on demand-side questions complements the supply-side analysis of the 5-star SEP presented in Decarolis and Guglielmo (2017). That paper showed a strategic response by the insurers who lowered both the premiums and the benefit generosity of the 5-star plans, while our study illustrates how enrollees responded to the combined changes in plan characteristics and the possibility to switch within-year. We find that switching within-year does increase, but that this is not associated with a worsening of selection. Indeed, enrollees in poor health are less likely to switch and this explains the reduction in risk scores observed for the 5-star plans.

These results suggest the relevance of two main avenues for future research. First, enrollees' inertia in choosing a plan emphasises the need to better understand the drivers of plan switching behavior and their interactions with the frequency and length of the open enrollment periods. Second, effective risk adjustment systems need to take into account plan switching behavior associated with the presence of special enrollment periods. This is a factor that should be preeminent in any discussion of SEP reforms involving changes to the set of "life qualifying events" that allow plan switches.[23]

Finally, the external validity of our results will be greater for those markets that, like

[23]This is also related to Ericson, Geissler and Lubin (2017) which acknowledges that partial-year enrollment is common and analyzes the problems that this poses to risk adjustment due to missing diagnoses. It then proposes a new adjustment for partial-year enrollment scaling up payments for partial-year enrollees' observed diagnoses.

Medicare, entail both consumers' inertia and insurers' ability to alter the product design. For instance, it would be interesting to consider how our results could contribute to the understanding of recent reforms of the ACA special enrollment periods. In fact, as discussed in Dorn (2016), the SEP in the ACA were designed to allow people who, due to job loss or other factors, needed to obtain Marketplace coverage outside of the standard open enrollment period. After the carriers claimed widespread abuse of the SEP by ineligible people, however, CMS tightened the requirements for SEP applicants by requesting to document their eligibility. It would thus be interesting to quantify whether this reform affected both premium and enrollment decisions in the ACA exchanges along the same lines that the 5-star SEP affected Medicare.

References

Abaluck, Jason, and Jonathan Gruber. 2011. "Choice Inconsistencies among the Elderly: Evidence from Plan Choice in the Medicare Part D Program." *American Economic Review*, 101(4): 1180–1210.

Abaluck, Jason, and Jonathan Gruber. 2016. "Improving the Quality of Choices in Health Insurance Markets." *No. w22917. National Bureau of Economic Research.*

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

Bhattacharya, Jay, and Darius Lakdawalla. 2006. "Does Medicare benefit the poor?" *Journal of Public Economics*, 90(1): 277 – 292.

Brown, Jason, Mark Duggan, Ilyana Kuziemko, and William Woolston. 2014. "How Does Risk Selection Respond to Risk Adjustment? New Evidence from the Medicare Advantage Program." *American Economic Review*, 104(10): 3335–64.

Cao, Zhun, and Thomas G McGuire. 2003. "Service-level selection by HMOs in Medicare." *Journal of Health Economics*, 22(6): 915–931.

Carey, Colleen. 2016. " Technological Change and Risk Adjustment: Benefit Design Incentives in Medicare Part D." *American Economic Journal: Economic Policy*, 9(1).

Conley, Timothy G., and Christopher R. Taber. 2011. "Inference with Difference in Differences with a Small Number of Policy Changes." *The Review of Economics and Statistics*, 93(1): 113–125.

Cutler, David M., and Richard J. Zeckhauser. 1998. "Adverse Selection in Health Insurance." In *Frontiers in Health Policy Research, volume 1.* , ed. Alan M. Garber. MIT.

Cutler, David M., and Sarah J. Reber. 1998. "Paying for Health Insurance: The Trade-Off between Competition and Adverse Selection." *The Quarterly Journal of Economics*, 113(2): 433–466.

Decarolis, Francesco, and Andrea Guglielmo. 2017. "Insurers' Response to Selection Risk: Evidence from Medicare Enrollment Reforms." *Journal of Health Economics*, 56: 383–396.

Diamond, Rebecca, Michael J Dickstein, Timothy McQuade, and Petra Persson. 2018. "Take-Up, Drop-Out, and Spending in ACA Marketplaces." National Bureau of Economic Research Working Paper 24668.

Dorn, Stan. 2016. "Helping Special Enrollment Periods Work under the Affordable Care Act." Urban Institute, Research Report.

Duggan, Mark, Amanda Starc, and Boris Vabson. 2016. "Who benefits when the government pays more? Pass-through in the Medicare Advantage program." *Journal of Public Economics*, 141(Supplement C): 50 – 67.

Einav, Liran, Amy Finkelstein, and Pietro Tebaldi. 2019. "Market Design in Regulated Health Insurance Markets:Risk Adjustment vs. Subsidies." mimeo mimeo.

Ellis, Randall P, and Elizabeth Savage. 2008. "Run for cover now or later? The impact of premiums, threats and deadlines on private health insurance in Australia." *International Journal of Health Care Finance and Economics*, 8(4): 257–277.

Ericson, Keith Marzilli, Kimberley Geissler, and Benjamin Lubin. 2017. "The Impact of Partial-Year Enrollment on the Accuracy of Risk Adjustment Systems: A Framework and Evidence." National Bureau of Economic Research Working Paper 23765.

Geruso, Michael, and Timothy Layton. 2015. "Upcoding: Evidence from Medicare on Squishy Risk Adjustment." National Bureau of Economic Research.

Glazer, Jacob, and Thomas McGuire. 2000. "Optimal Risk Adjustment in Markets with Adverse Selection: An Application to Managed Care." *American Economic Review*, 1055–1071.

Guglielmo, Andrea. 2016. "Competition and Costs in Medicare Advantage." mimeo.

Handel, Ben. 2013. "Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts." *American Economic Review*, 103(7): 2643–2682.

Handel, Benjamin R., and Jonathan T. Kolstad. 2015. "Health Insurance for "Humans": Information Frictions, Plan Choice, and Consumer Welfare." *American Economic Review*, 105(8): 2449–2500.

Heiss, Florian, Daniel McFadden, Joachim Winter, Amelie Wuppermann, and Bo Zhou. 2016. "Inattention and switching costs as sources of inertia in Medicare Part D." mimeo.

Ho, Kate, Joseph Hogan, and Fiona Scott Morton. 2017. "The Impact of Consumer Inattention on Insurer Pricing in the Medicare Part D Program." *RAND Journal of Economics*, forthcoming.

Honka, Elisabeth. 2014. "Quantifying search and switching costs in the US auto insurance industry." *The RAND Journal of Economics*, 45(4): 847–884.

Ketcham, Jonathan, Claudio Lucarelli, and Christopher A. Powers. 2014. "Paying Attention or Paying Too Much in Medicare Part D." *American Economic Review*, forthcoming.

Ketcham, Jonathan D., Claudio Lucarelli, Eugenio J. Miravete, and M. Christopher Roebuck. 2012. "Sinking, Swimming, or Learning to Swim in Medicare Part D." *American Economic Review*, 102(6): 2639–73.

Kuziemko, NIlyana, Katherine Meckel, and Maya Rossin-Slater. 2014. "Do Insurers Risk-Select Against Each Other? Evidence from Medicaid and Implications for Health Reform." mimeo.

Layton, Timothy J., and Andrew Ryan. 2015. "Higher Incentive Payments in Medicare Advantage's Pay-for-Performance Did Not Improve Quality, But Did Increase Plan Offerings." *Health Services Research*, 50(6): 1810–28.

Li, Pengxiang, and Jalpa A. Doshi. 2016. "Impact of Medicare Advantage Prescription Drug Plan Star Ratings on Enrollment before and after Implementation of Quality-Related Bonus Payments in 2012." *PLOS One*, May.

Madeira, Tarso. 2015. "The Cost of Removing Deadlines: Evidence from Medicare Part D." mimeo, `https://www.ocf.berkeley.edu/~tmadeira/jmp.pdf`.

Marzilli Ericson, Keith M. 2014. "Consumer inertia and firm pricing in the Medicare Part D prescription drug insurance exchange." *American Economic Journal: Economic Policy* 6, no. 1: 38-64.

McClellan, Mark, and Jonathan Skinner. 2006. "The incidence of Medicare." *Journal of Public Economics*, 90(1): 257 – 276.

McWilliams, J Michael, John Hsu, and Joseph P Newhouse. 2012. "New risk-adjustment system was associated with reduced favorable selection in Medicare Advantage." *Health Affairs*, 31(12): 2630–2640.

Miller, Keaton, Amil Petrin, Robert Town, and Michael Chernew. 2014. "Health Plan Changes in Response to Medicare Advantage Payment Rates." *mimeo.*

Newhouse, Joseph P, J Michael McWilliams, Mary Price, Jie Huang, Bruce Fireman, and John Hsu. 2013. "Do Medicare Advantage plans select enrollees in higher margin clinical categories?" *Journal of health economics*, 32(6): 1278–1288.

Nosal, Kathleen. 2012. "Estimating Switching Costs for Medicare Advantage Plans." *Mimeo.*

Polyakova, Maria. 2014. "Regulation of insurance with adverse selection and switching costs: Evidence from Medicare Part D." mimeo.

Polyakova, Maria. 2016. "Regulation of Insurance with Adverse Selection and Switching Costs: Evidence from Medicare Part D." *American Economic Journal: Applied Economics*, 8(3): 165–95.

Rothschild, Michael, and Joseph Stiglitz. 1976. "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information." *Quarterly Journal of Economics*, 90(4): 629Ð649.

Shepard, Mark. 2016. "Hospital Network Competition and Adverse Selection: Evidence from the Massachusetts Health Insurance Exchange." mimeo.

For Publication on the Authors' Web Pages

Web Appendix

1) Data

The dataset was assembled from data made publicly available by the Center for Medicare and Medicaid Services (CMS). This is the same dataset as used in Decarolis and Guglielmo (2017). In particular, data on monthly enrollment for the years 2009-2013 at plan level were downloaded from:

`http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/MCRAdvPartDEnrolData/index.html`.

The *Crosswalk Files* available from the same web site were used to link plans through the years. Premiums and plan financial characteristics are from the *Premium Files*:

`http://www.cms.gov/Medicare/Prescription-Drug-Coverage/PrescriptionDrugCovGenIn/index.html`.

Plans' formulary and pharmacy networks are from the FRF (*Formulary Reference Files*):

`https://www.cms.gov/PrescriptionDrugCovContra/03_RxContracting_FormularyGuidance.asp`

Demographic characteristics for the geographic areas are the only ancillary data source and were obtained from:

`http://ahrf.hrsa.gov/download.htm`.

Contract performance data determining the star ratings were obtained from:

`https://www.cms.gov/medicare/prescription-drug-coverage/prescriptiondrugcovgenin/performancedata.html`

An example for 2012 of the how the nearly 50 individual measures are grouped into 9 domains and 2 components (managed care and prescription drugs) is presented in Table A.1. Evidence on the lack of score manipulation at the threshold in the post-reform period is presented through the McCrary tests reported in Figure A.1.

2) Matched Sample Results

The first set of additional results reported concerns the probit estimates used for the construction of the matched DID estimates in the enrollment analysis. Table A.2 reports the estimates for four model specifications (i.e., columns 1-2, 3-4, 5 and 6) where we gradually increase the set of controls. All controls are county-level demographic characteristics collected from the AHRF files of the Health Resources and Services Administration. The estimates reported in column 2 and 4 differ from those in columns 1 and 3, respectively, for the sample of counties included: due to missing data for some characteristics, for columns 2 and 4 we use a smaller sample than that used for columns 1 and 3. The sample used for columns 2 and 4 is the same as that used for columns 5 and 6. The matched DID reported in the main text are based on the estimates in column 6 of Table A.2. Although this table clearly shows that estimates are fairly stable across models, to further assess the robustness of the DID in the main text we report in Table A.3 matched DID estimates based on the outcomes of the three other probit models (i.e. model 1, 3 and 5). Overall, the results are broadly in line with what is reported in the main text.

3) Placebo Analysis

In Table A.6 we report the results of a placebo analysis. In particular, we repeat our analysis as if the 5-star SEP were introduced in 2011 instead of 2012. To avoid potential spillovers from the true SEP, we narrowed our exercise to the enrollment periods from 2009 to 2011. For the within-year changes, Panel A shows that, in our first two specifications, the simulated SEP has a positive and statistically significant effect on the within-year enrollment change, but this effect vanishes once we control for time trends. Furthermore, we do not find a statistically significant effect of the placebo SEP on the percentage change in enrollment. Panel B reports the analogous estimates for the across-year estimates. These estimates indicate that plan choices post the 5-star SEP are indeed different from those in the earlier period: with the placebo simulating that the policy were implemented in 2011, we obtain negative (albeit not always significant) estimates.

4) Across- and Within-Year Enrollment for PDP

The analysis of PDP demand effects presents different challenges from the MAPD case. A major concern is that only 2 regions are treated. Even with consumer-level data, this would limit the ability to conduct inference as the asymptotic conditions underlying the DID estimator cannot be satisfied. This problem can be solved by exploiting the large number of control group observations (32 regions) through the method of Conley and Taber (2011) if one is willing to assume that any random shock that might have hit the two treated regions simultaneously with the 5-star SEP reform belongs to the same distribution of shocks

affecting the regions in the control group. In this respect, a strength of the PDP market relative to the MAPD one is that no payment reforms occurred simultaneously with the 5-star SEP. Thus, while we follow an approach analogous to that used for MAPD and include in the control group only plans with a rating no lower than 4, in principle it could be less problematic for PDP to have a broader definition of the control group.

Table A.5 reports the estimates of the enrollment analysis for Part D. Within-year enrollment is the dependent variable in columns 1-4. In each case, the first two sets of estimates regard the variable in levels, while the next two involve enrollment as a percentage calculated as for the Part C case. Across-years enrollment is the dependent variable in the following 4 columns. The Part D estimates are broadly in line with the earlier Part C findings. There is a positive and significant effect of the 5-star SEP on within-year enrollment change. The magnitude is also similar to what found for Part C amounting to roughly 10 percent of the enrollment base. The across-years enrollment of 5 star contracts declines, but in a way that is not statistically significant.

Table A.1: Domain Measures for Medicare Advantage Plan Rating - Year 2012

Managed Care		**Prescription Drugs**	
Staying Healthy: screenings, tests, vaccines	12	Drug Plan Customer Service	3
Managing Chronic (long-term) Conditions	9	Member Complaints, problems getting services, and improvement in the drug plan's performance	3
Member Experience with the Health Plan	5	Member Experience with the Drug Plan	3
Member Complaints, problems getting services, and improvement in the health plan's performance	3	Patient safety and accuracy of drug pricing	6
Health Plan Customer Service	2		

Notes: The table reports the list of the domain measures used to calculate the summary ratings in 2012. There are 5 domain measures for the managed care component and 4 cor the prescription drugs component. The numbers in the table that follow the description of each domain measure indicate the number of underlying individual measures.

Table A.2: Probit Results - Probability of County Having 5 Star Plan

	(1) 5 Star County	(2) 5 Star County	(3) 5 Star County	(4) 5 Star County	(5) 5 Star County	(6) 5 Star County
MA Enrollees	2.981***	2.334***	2.858***	2.268***	2.234***	2.255***
	(0.448)	(0.484)	(0.454)	(0.487)	(0.513)	(0.518)
Pop. Male > 65	0.000951***	0.00126***	0.000896***	0.00120***	0.00100*	0.00105*
	(0.000333)	(0.000461)	(0.000317)	(0.000456)	(0.000555)	(0.000600)
Pop. Female > 65	-0.000787***	-0.000973***	-0.000747***	-0.000921***	-0.000836**	-0.000878**
	(0.000245)	(0.000328)	(0.000236)	(0.000324)	(0.000392)	(0.000430)
Pop. White-Male > 65	-0.000890**	-0.00119**	-0.000851**	-0.00114**	-0.00111*	-0.00118*
	(0.000361)	(0.000489)	(0.000344)	(0.000484)	(0.000592)	(0.000645)
Pop. White-Female > 65	0.000573**	0.000780**	0.000542**	0.000739**	0.000653	0.000705
	(0.000255)	(0.000348)	(0.000242)	(0.000344)	(0.000413)	(0.000451)
Medicare Eligibles	8.13e-05***	6.55e-05***	8.25e-05***	6.47e-05**	0.000149***	0.000150***
	(2.38e-05)	(2.53e-05)	(2.46e-05)	(2.62e-05)	(3.80e-05)	(4.09e-05)
Unemployment			0.0519**	0.0488*	0.0305	0.0289
			(0.0254)	(0.0267)	(0.0285)	(0.0289)
Poverty Rate			-0.0321**	-0.0241	-0.0110	-0.0104
			(0.0148)	(0.0155)	(0.0159)	(0.0162)
# Medicare Cert Hosp.					0.216***	0.110
					(0.0660)	(0.256)
# Hosp. Med Patients					-2.32e-05***	-2.63e-05***
					(4.15e-06)	(4.87e-06)
# Outpatients Visits					1.50e-07	1.03e-07
					(2.17e-07)	(2.41e-07)
Hosp. Util. Rate 0-39						-0.0999
						(0.270)
Hosp. Util. Rate 40-59						0.144
						(0.262)
Hosp. Util. Rate 60-79						0.296
						(0.262)
Hosp. Util. Rate >80						0.330
						(0.283)
Constant	-1.762***	-1.588***	-1.756***	-1.681***	-1.960***	-1.922***
	(0.109)	(0.120)	(0.241)	(0.268)	(0.291)	(0.295)
Observations	987	841	987	841	841	841

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A.3: MAPD Contracts - Within-year Enrollment Change - Matched Sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Model 1							
	Dec.-Jan. Enrollment Change				Dec.-Jan. Enrollment % Change			
5 Star	212.267***	222.275***	79.771**	78.749**	0.078*	0.089**	0.154**	0.140*
	(49.064)	(48.172)	(38.894)	(37.235)	(0.044)	(0.043)	(0.075)	(0.072)
Observations	8,486	8,486	8,486	8,486	8,486	8,486	8,486	8,486
R-squared	0.635	0.686	0.647	0.697	0.193	0.272	0.224	0.305
	Panel B: Model 3							
	Dec.-Jan. Enrollment Change				Dec.-Jan. Enrollment % Change			
5 Star	210.904***	221.579***	80.953**	80.095**	0.073*	0.087**	0.156**	0.144**
	(49.086)	(48.160)	(38.891)	(37.213)	(0.044)	(0.043)	(0.075)	(0.072)
Observations	8,734	8,734	8,734	8,734	8,734	8,734	8,734	8,734
R-squared	0.628	0.682	0.640	0.694	0.188	0.273	0.219	0.305
	Panel C: Model 5							
	Dec.-Jan. Enrollment Change				Dec.-Jan. Enrollment % Change			
5 Star	154.346***	161.143***	66.955**	66.349***	0.089*	0.100**	0.222***	0.205***
	(26.869)	(26.381)	(26.612)	(25.548)	(0.046)	(0.046)	(0.079)	(0.076)
Observations	7,533	7,533	7,533	7,533	7,533	7,533	7,533	7,533
R-squared	0.440	0.523	0.453	0.536	0.183	0.271	0.219	0.307
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES	YES	YES	YES
Contract FE	NO	YES	NO	YES	NO	YES	NO	YES
Time Trend	NO	NO	YES	YES	NO	NO	YES	YES

Notes: The table reports the DID estimates of the effect of the 5-star SEP. The outcome variable is the difference in the contract enrollment between December and January (of the same year) calculated either in levels (first four columns) or in percentage (latter four columns). The four model specifications considered for each dependent variable differ in the set of controls used, as reported in the block at the very end of the table.

Table A.4: Logit Estimates for 5-Star SEP Switches: Marginal Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Regular and LIS Enrollees				Regulars	LIS	Health_{t+1}
Health Status							
Acute High	-0.0013	-0.0014	-0.0014	-0.0012	-0.0011	-0.0013	-0.0016
Chronic Low	-0.0016	-0.0017	-0.0016	-0.0016	-0.0017	-0.0013	-0.0016
Chronic High	-0.0032	-0.0032	-0.0032	-0.0029	-0.0030	-0.0024	-0.0028
Mental	-0.0006	-0.0005	-0.0005	-0.0004	-0.0004	-0.0006	-0.0003
Observations	4,934,656	4,934,656	4,934,656	4,934,656	4,125,297	809,359	2,211,384

Marginal effects calculated at the means for the logit regressions presented in the main text

Table A.5: PDP Plans - Within and Across Year Enrollment Changes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Within-Year				Across-Year			
	Change		% Change		Change		% Change	
5 Star	2,419**	2,416**	0.128**	0.130**	-17,835	-17,582	-0.0786	-0.0803
	(919.3)	(900.4)	(0.0567)	(0.0550)	(12,233)	(11,985)	(0.0896)	(0.0873)
Observations	499	499	499	499	497	497	372	372
R-squared	0.018	0.026	0.204	0.251	0.186	0.202	0.074	0.097
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Region FE	YES	YES	YES	YES	YES	YES	YES	YES
Contract FE	NO	YES	NO	YES	NO	YES	NO	YES

Columns (1)-(4) report estimates of the 5-star SEP dummy on within-year PDP enrollment changes; columns (5)-(8) report the effect on across-year changes. Significance level *** p<0.01, ** p<0.05, * p<0.1.

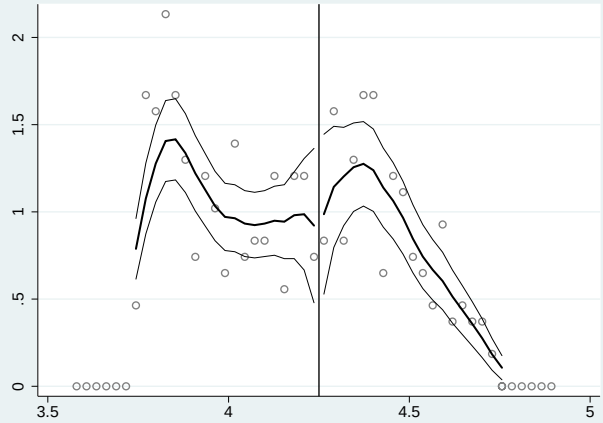
Table A.6: MAPD Contracts - Placebo Analysis

	Panel A: Within-Year Switches							
	Dec.-Jan. Enrollment Change				Dec.-Jan. Enrollment % Change			
5 Star	108.113***	116.613***	15.924	15.130	-0.038	0.015	0.155	0.102
	(33.168)	(30.406)	(46.024)	(42.162)	(0.065)	(0.059)	(0.110)	(0.099)
Observations	5,205	5,205	5,205	5,205	5,205	5,205	5,205	5,205
R-squared	0.469	0.618	0.478	0.630	0.277	0.428	0.311	0.464
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES	YES	YES	YES
Contract FE	NO	YES	NO	YES	NO	YES	NO	YES
Time Trend	NO	NO	YES	YES	NO	NO	YES	YES
	Panel B: Across-Years Switches							
	Jan.-Dec. Enrollment Change				Jan.-Dec. Enrollment % Change			
Star 5	-16.327	-13.821	-95.281**	-90.711***	-0.176***	-0.094**	0.009	0.027
	(19.684)	(18.303)	(37.429)	(32.715)	(0.054)	(0.048)	(0.106)	(0.087)
Observations	4,636	4,636	4,636	4,636	4,636	4,636	4,636	4,636
R-squared	0.090	0.172	0.092	0.174	0.197	0.391	0.204	0.395
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES	YES	YES	YES
Contract FE	NO	YES	NO	YES	NO	YES	NO	YES
Time Trend	NO	NO	YES	YES	NO	NO	YES	YES

Notes: The table reports the DID estimates of the effect of the 5-star SEP. Analogous to those in the main text, but using a placebo treatment group. The sample covers the year 2009-2011 with a simulated policy introduced in 2011. The estimates in Panel A, correspond to those in Table 2, while those in Panel B correspond to those in Table 3. Standard errors in parentheses clustered at county level *** p<0.01, ** p<0.05, * p<0.1.

Figure A.1: Distribution Star Rating Post Reform - McCrary Test

(a) Discontinuity - 4/4.5 Star



(b) Discontinuity - 4.5/5 Star