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
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 Part C and Part D to analyze how Medicare beneficiaries responded to the option of switching to 5-star rated plans at anytime, in a setting where insurers adjusted premiums and benefit design to counterbalance the increased selection risk. We find that within-year switches to 5-star plans increased by 7-16% and that those who switch are advantageously selected. Furthermore, demand for 5-star plans across the years did not change

Francesco Decarolis
Department of Economics
Boston University
270 Bay State Road
Boston, MA 02215
and EIEF
and also NBER
francesco.decarolis@eief.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 market can lead to more desirable social outcomes. Several of these studies have involved the design of the Medicare system: with expenditures totaling $646.2 billion in 2015 and growing by 4.5 percent relative to the previous year, Medicare represents through its Part C (or Medicare Advantage) and Part D programs the largest existing case of a publicly founded, but privately provided health insurance system.

As typical in insurance markets, both Medicare Part C, 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 coverage under clearly defined contractual conditions. Among such conditions, there is the impossibility for the enrollee of switching plan at will during the coverage period. Open enrollment periods (OEP) play a key role for 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 into high generosity 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 not an option for insurers offering Part C or D plans who must accept all enrollment requests from Medicare beneficiaries. This makes all the more surprising that since 2012 the Part C and D markets have been able to continue working despite increased potential for adverse selection. This change was brought about by a reform of the OEP rules allowing enrollees to switch at anytime under the sole condition that the destination plan was 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), was introduced with the goal of increasing the enrollment in 5-star plans. It involves a large share of the 40 million Medicare beneficiaries as, for instance, in 2017 the 5-star SEP is available to 11.5 million individuals residing in areas with at least one 5-star plan.[2]

In a previous study, we analyzed how insurers responded to the 5-star SEP (Decarolis and

[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] 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 are available across all counties in 12 states and 5-Star rated Medicare Advantage plans are available in 261 counties across 18 states, see: `https://q1medicare.com/q1group/MedicareAdvantagePartDQA/FAQ.php?faq=What-is-the-5-star-Special-Enrollment-Period-&faq_id=558&category_id=125`.

Guglielmo, 2017). By exploiting the geographical variation in the offering of 5-star plans, we causally identified the 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 seeking 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 has been the impact on demand of the combined effects of free plan switching by enrollees and plan design changes by insurers. Answering this question is the main contribution of the current study and it represents key knowledge to understand the potential effectiveness of using open enrollment rules as a market design tool in environments where insurers can alter their plan design.

To identify how demand responded to the SEP reform, we use a similar approach to that in (Decarolis and Guglielmo, 2017). We exploit the geographical variation in 5-star plans to compare demand in markets with 5-star plans to that of similar markets, but where no 5-star plan is offered. Our difference-in-differences strategy is particularly effective when insurers have a limited scope to game the star rating system. Therefore, we focus on the first two years of the reform (2012 and 2013) when the insurers could alter the plan design, but not their star rating due to the lag in the timing of the specific measures composing the rating. We also focus mostly on Part C plans given the greater geographical dispersion of 5-star plans, relative to that of 5-star Part D plans. Furthermore, for the Part C plans, 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 for 5-star Part C plans are as follows. First, we estimate that the within-year increase in enrollment due to the 5-star SEP ranges between 7 and 16 percent of the enrollment base of the 5-star plans. This confirms a sizable response of 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 choices: enrollees do not take advantage of the possibility of staying outside the Part C program (or enrolling in the cheapest plans) during the open enrollment period, to then switch to 5-star plans only if hit by health shocks. Identical results in terms of within and across year demand changes are documented also for Part D. Third, the risk pools of Part C 5-star plans improves, albeit 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 be improving because they are bringing in enrollees that, despite being among the high risk ones in their plan of origin, are nevertheless of 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 Part C plan is negatively associated with measures of poor health status. 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 one. This is consistent with the supply response to the SEP involving changes to the plan characteristics that made them more appealing for most enrollees (though lower premiums), but less so for those in worse health (through lower benefit generosity for poor health enrollees).

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 demand increase to be 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. This evidence complements the very scarce evidence existing on the effects of open enrollment periods and that it is likely due to the lack of policy variations. Indeed, one of the few other papers 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 effective to induce consumers to enroll now rather than delay.

More generally, 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 incidence of Medicare has received considerable attention in the literature (see, for instance, Bhattacharya and Lakdawalla (2006), McClellan and Skinner (2006) and Duggan et al. (2016)). In these studies, quantifying the insurance value of Medicare serves a key role to assess its distributional impacts. In this respect, our findings reveal how even a "small" reform affecting directly just 5-star plans is able to trigger potentially vast changes in Medicare's insurance value by triggering supply and demand responses ending up in an equilibrium where the highest quality plans are less valuable for those enrollees in worse health.

Our results on the effects of the 5-star SEP are also relevant because, to the best of our knowledge, the research on this important reform is very limited with only two other papers looking at it. The first is our supply side study, (Decarolis and Guglielmo, 2017), discussed earlier. 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. He seeks to study whether, by removing the typical Part D enrollment deadline, the 5-star SEP could have induced consumers to switch plans less frequently by giving them the opportunity to procrastinate. The results, albeit preliminary,[3] suggest that switching rates (across the years) decrease as a result of the policy change, in a way consistent with a procrastination story. Our results complement and substantially extend these findings as they look directly at the main aspect of the policy (the within-year switches, instead of the across-years plan changes) and they do so by using not only 2012, but also 2013 data and focusing mostly on Part C, for which we observe nearly 180 treated markets, relative to the only 2 treated markets in Part D. As discussed below, this clearly impacts the reliability of the estimates.

The implications of our findings beyond the context of Medicare hinge on two key features regarding both demand and supply. Starting from the latter, Medicare private insurers have been extensively shown to have sufficient leeway to shape plan features (Cao and McGuire (2003), Batata (2004), McWilliams et al. (2012), Newhouse et al. (2013), Brown et al. (2014), Polyakova (2014), Carey (2016), Guglielmo (2016) and Decarolis and Guglielmo (2017)). Medicare, however, is not unusual in terms of the type of actions that insurers can take and various other important markets have been shown to be characterized by similar features. For instance, Kuziemko et al. (2014) shows how competition in the presence of risk selection in Medicaid managed care leads to a worsening of outcomes for enrollees in poorer health conditions, while Shepard (2016) shows evidence of how selection in the choice of plans' hospital networks leads to leave out of network the "star" hospitals preferred by high cost enrollees in the context of the Massachusetts subsidized health insurance exchange.

Regarding demand, our findings reflect features of the demand for Medicare plans and, in particular, switching cost and inertia. These phenomena have been extensively analyzed in the context of Part D and, to a lesser extent, Part C. These studies include Nosal (2012) and Miller et al. (2014) for Part C and Abaluck and Gruber (2011), Ketcham et al. (2012), Marzilli Ericson (2014), Ketcham et al. (2014), Abaluck and Gruber (2016), Ho et al. (2014), Polyakova (2014), Ho et al. (2017), Madeira (2015) and Heiss et al. (2016) for Part D. In line with these studies, we also document that Medicare enrollees are not acting in a sophisticated way, which in our context would amount to exiting expensive 5-star plans

[3] To the best of our knowledge this job market paper has not been updated since 01/31/2015.

during the regular open enrollment period to then join them through the 5-star SEP once hit by negative health shocks. The presence of this type of inertia is likely of great importance for the insurers' strategy as the introduction of the 5-star SEP in a frictionless environment might have required on the insurers' side even more drastic changes in the benefit design to counterbalance the increased adverse selection risk, possibly interfering with what the regulations allow insurers to do in terms of plan design. The presence of inertia, however, is by no means unique to the Medicare context. On the contrary, its presence in various insurance markets has been documented, for instance by Handel and Kolstad (2015) in employer sponsored health insurance, Honka (2014) in the auto insurance industry and Handel and Kolstad (2015) in pension plan choice.

Therefore, since both the demand and the supply features that play a key role in shaping the findings documented in this paper are likely to be present in several other markets as well, we believe that the analysis presented in this paper can be useful more broadly to understand the benefits and risks of using open enrollment rules as a design tool for insurance markets.

II Institutions: Medicare Open Enrollment Periods

The Medicare system consists of a series of interlinked programs aimed chiefly at US elderly aged 65 or older. Traditional Medicare (TM) is composed by 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 two privately provided programs that supplement TM: Part C and Part D. In both programs, private insurers offer a menu of plans to Medicare beneficiaries: Part C plans are alternative to TM and hence must cover Medicare Part A and B benefits (except hospice care), but can also offer additional benefits. Part D plans cover prescription drugs. The two programs are closely connected in many ways, the most evident being that almost all Part C plans also include Part D benefits. These Part C+D 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 solar year. The open enrollment period (OEP) is the window of time during which enrollees can enroll into these plans. It typically spans from October to December of the the year before the coverage period. Although enrollees are generally required to keep the same plan for the entire coverage year, exceptions to the OEP exist. Special Enrollment Periods (SEPs) permit enrollees to change plans when certain special circumstances occur. The most relevant SEPs involve individuals turning 65 during the coverage year, changing residency or transiting to "low income enrollee" status. Starting in 2012, an additional SEP was introduced: all Medicare eligibles residing in an area with one or more 5-star Part C or D plan offered, can switch from their plan (or from TM) to anyone of these 5-star plans during the coverage year, with the new coverage starting the first day of the month following the enrollment request.[4] 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 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 consumers as well as on its web site (see Figure 1).

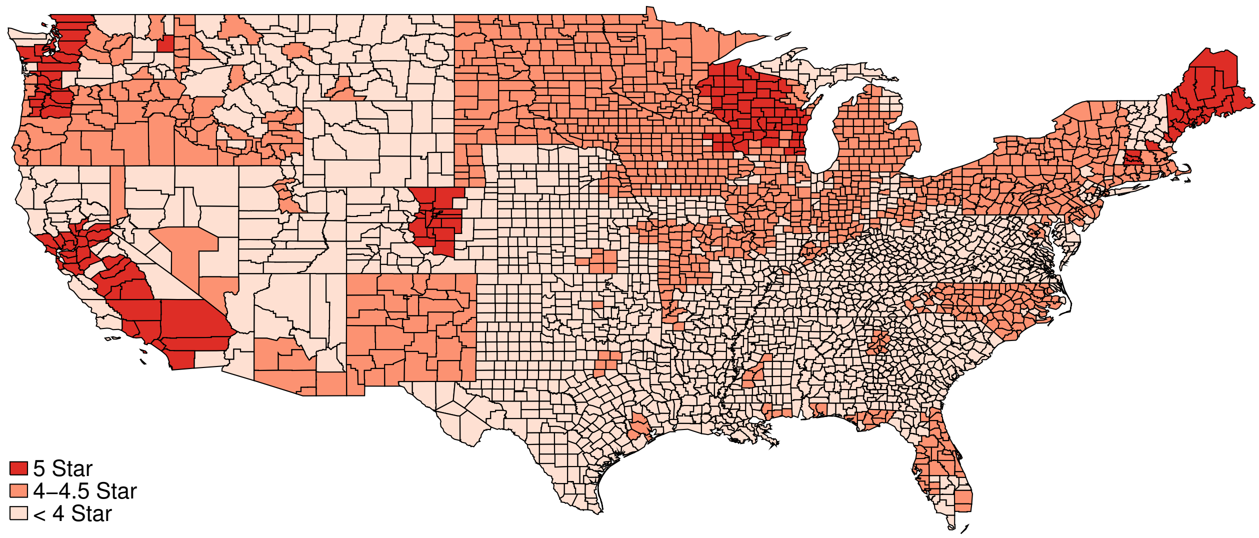
Figure 1: Screen Shot of the CMS Web Page on the 5-star SEP

Snapshot of the CMS web site taken on September, 2017. https://www.medicare.gov/sign-up-change-plans/when-can-i-join-a-health-or-drug-plan/five-star-enrollment/5-star-enrollment-period.html

The key novelty of this SEP is its linkage to the plans supply side. Contrary to all other Medicare SEPs, the possibility of a within-year plan switch is driven exclusively by the presence of a 5-star plan offered in the enrollee's area of residency. This area has a very

[4]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 - under Section 1851(e)(4)(D) of the Social Security Act - as part of CMS' overall effort effort to bolster plan quality and, hence, most notably with the quality bonus payment demonstration (see Layton and Ryan (2015) and Li and Doshi (2016)). See also http://www.cms.gov/Medicare/Prescription-Drug-Coverage/PrescriptionDrugCovContra/downloads/Announcement2012final2.pdf.

Figure 2: Maps of 5-Star Counties



Heat map: darkest colors indicate counties where the highest-rated MAPD have higher star ratings.

distinct size for Part C and 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). These MAPD were offered by 7 different insurers, mainly Kaiser and Humana, while the PDP were offered by 2 insurers, one of them being again Humana. Due to the greater heterogeneity among MAPD offerings, we will focus most of our analysis on MAPD. Figure 2 illustrates the spatial pattern of 5-star MAPD offerings. While not present in the Southern states, 5-star MAPD are present in the Eastern, Central and Western regions. The heat map reveals also the localization of counties whose highest rated MAPD were either 4 or 4.5 stars. As discussed below, while it would be inadequate to consider the localization of 5-star counties as exogenous, the identification strategy that we follow is based on the idea that within the union of counties with 4-4.5 and 5-star counties the assignment of the 5-star status is random. Under this assumption we can then compare the enrollment behavior in 4-4.5 vs 5-star counties to study the effects of the 5-star SEP reform.

The logic behind this assumption is straightforward if two facts are simultaneously considered. First, 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. An example is offered in Table 1 where we report the 17 individual measures entering the PDP rating in 2012. For MAPD, in addition to the measures used for PDP, about 30 more measures are used to evaluate their Part C component. 5-star plans are those whose overall score is at least 4.75, while 4.5-star plans have an overall rating below that, but above 4.25. Most 5-star plans fall short of having an overall continuous score of 5, reaching a score not much higher than 4.75.[5] This is thus reassuring regarding their comparability to lower star-rated plans.

Second, up until the enrollment year 2014, insurers could not respond to the 5-star SEP by altering the menu of 5-star contract offered. The calculation of the star rating is based on measures collected from different sources. For instance, the second column of Table 1 lists 9 different types of data ranging from survey to call center and administrative data. Crucially, 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 on November 2010, this implies that any action aimed at altering the star rating would not produce its 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. Clearly, insurers became able to alter the premium and benefit design of their plans much earlier, starting from their 2012 plan offerings. As discussed above, Decarolis and Guglielmo (2017) indeed find that insurers responded along both margins with 5-star plans lowering their premium, but also their generosity specifically for those features most valued by the least health enrollees. This implies that our analysis below must be interpreted as an assessment of how enrollment responded jointly to the 5-star SEP and the plan design changes.

III Data

The analysis combines several data sources. The analysis of enrollment patterns and risk score changes will be based on CMS data on plan monthly enrollment and characteristics (star rating, premiums and various features of the benefit design). To analyze more in details the enrollees switching under the 5-star SEP, we use consumer-level data based on a random sample of about 2.5 million Medicare enrollees per year followed from 2011 to 2013 in all their drug purchases and plan choices. Finally, the Area Health Resource File is used to control for county-level demographic, economic and health characteristics.

The three main outcome variables that we analyze are: *(i)* the within-year change in

[5]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.

Table 1: Rating Calculation for Part D - Year 2012

Individual Measures			Domain Measures	Summary Measures
Definition	Type of Data	Weights		
D01 Call Center - Hold Time	Call Center Monitored by CMS	1.5	Domain 1 Drug Plan Customer Service	Summary Rating
D02 Call Center - Foreign Language Interpreter	Call Center Monitored by CMS	1.5		
D03 Appeals Auto-Forward	Independent Review Entity	1.5		
D04 Appeals Upheld	Independent Review Entity	1.5		
D05 Enrollment Timeliness	Medicare Advantage Prescription Drug System (CMS)	1		
D06 - Complaints about the Drug Plan	Complaint Tracking System (CMS)	1.5	Domain 2 Member Complaints, Problems Getting Services, and Choosing to Leave the Plan	
D07 - Beneficiary Access and Performance Problems	CMS Administrative Data	1.5		
D08 - Members Choosing to Leave the Plan	Medicare Beneficiary Database Suite of Systems (CMS)	1.5		
D09 - Getting Information From Drug Plan	CAHPS Survey	1.5	Domain 3 Experience with Drug Plan	
D10 - Rating of Drug Plan	CAHPS Survey	1.5		
D11 - Getting Needed Prescription Drugs	CAHPS Survey	1.5		
D12 - MPF Composite	Prescription Drug Event, Medicare Plan Finder, Health Management Plan System and Medispan	1	Domain 4 Drug Pricing and Patient Safety	
D13 - High Risk Medication	Prescription Drug Event	3		
D14 - Diabetes Treatment	Prescription Drug Event	3		
D15 - Part D Medication Adherence for Oral Diabetes Medications	Prescription Drug Event	3		
D16 - Part D Medication Adherence for Hypertension (ACEI or ARB)	Prescription Drug Event	3		
D17 - Part D Medication Adherence for Cholesterol (Statins)	Prescription Drug Event	3		

Notes: The table reports the details of how the 2012 summary rating is calculated for Part D. There are three sets of measures: individual measures (17 measures, reported in the first column), domain measures (4 measures, reported in the fourth column) and the final summary rating (fifth column). The third column describes the weights associated with the individual measures in the calculation of the domain measures.

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, hence, 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.

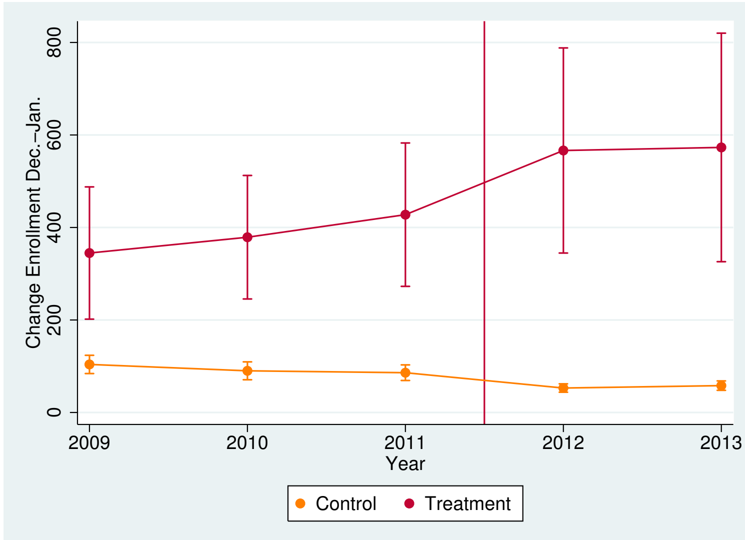
Table 2: Descriptive Statistics for Part C

	2009-2011							
	Control				Treament			
	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

	2012-2013							
	Control				Treament			
	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: "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 outside 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 (out of 117 most frequently purchased) 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 coverage. "Health Care Quality" is a star rating (1-5), over member's evaluation of health care quality (CAHPS). "Customer Service" is a star rating (1-5), over ability of the health plan to provide information or help when members need it (CAHPS). "Drug Access" is a star rating (1-5) over 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 contract level. "Premium Part C", "Premium Part D", "In Network MOOP", "Deductible Part D", "N. Top Drugs", "N. Unrestricted Drug", "Part C OOPC Excellent (Poor)", " Drug OOPC Excellent (Poor)", "Risk Score Part C" and "Risk Score Part D" are measured at plan level and aggregated at contract level as weighted average, with enrollment as weights.

Figure 3: MAPD Contracts - Within Year Enrollment Change



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

More precisely it is calculated as $Enrollment_{1/t} - Enrollment_{12/t-1}$. This variable can capture a strategic response by consumers: 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 yearly level and separately for Part C and D. The risk score is the key statistic mapping how enrollment composition impacts expected plan costs. In the 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 in not the case for the control group. To better explore this feature of the data, Figure 3 reports the evolution of the average within-year enrollment change for both treatment and control plans. The figure confirms both the presence of a relatively large increase for the treatment group and the lack of any increase for the control group. Below we offer a more precise quantification of the causal effect of the 5-star SEP, controlling for several confounding factors that might be present in the averages displayed in Figure 3.[6]

[6] Another key feature of the analysis that is revealed by Table 2 is the change in the plan offerings following the 5-star SEP. As discussed above, Decarolis and Guglielmo (2017) find that plans altered both premiums and generosity and, indeed, the data in Table 2 show that Part C premiums tend to decline more for the treatment than for the control group, while those features valued by the least healthy enrollees worsen (as illustrated for, instance, by the increase in the maximum out of pocket, or the expected out of pocket for enrollees in poor health). These changes are driven by changes in the existing set of contracts and not by a

IV Empirical Strategy

Our strategy to identify 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 2, that 5-star contracts are offered in only a subset of the US counties. We consider all contracts that achieve the 5-star rating in the period 2012-2013 as the DID treatment group (dark red areas in in Figure 2) and all contracts that achieve a 4 or 4.5 rating in the same period and are offered in counties without any 5-star contract as the control group (light red areas in in Figure 2). 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 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 of this baseline model are presented below.

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 our setting, 4 and 4.5 star contracts offered in counties that do not have any 5-star plan are a nearly ideal control group. Clearly, both the control and the treatment contracts are similar as they are the top quality contracts offered in their respective counties. Furthermore, contracts in the control group face similar financial incentives of those in the treatment group as all payments linked to the star rating are very similar for these two groups of plans.footnoteSee a discussion of the link between the star rating and financial incentives in Decarolis and Guglielmo (2017).

As shown in Table 2, however, treatment and control groups differ along several observable characteristics, like size of the enrollment base and features of the enrollment pool. Indeed, although Figure 2 reveals that the 5-star plans are scattered across many different counties, this does not ensure their assignment to counties is random. We have two arguments to address this concern, the first is that, for the three reasons explained in section 2, it is hard for insurers to perfectly control their rating so that the difference between a 4-4.5 and a 5-star plan is likely quasi-random, at least for the period object of analysis. Therefore, our identification strategy rests upon the fact that the assignment of the treatment relative to the control status is quasi-random within the union of the counties marked in dark and

different composition of the set of contracts.

light red in Figure 2. Since the regulation separates the geographical markets, an additional benefit of this strategy is that, by selecting the treatment and control groups from different counties, it avoids contamination issues. Second, to the extent that the selection into the treatment state is based on observable characteristics, we have a rich set of covariates that permits us to control for this threat. Thus, as a robustness check for our baseline estimates we use a matching DID strategy, where the control group observations are selected to match the characteristics of the treatment group.

V Results

This section presents the results separately in five parts. The first two sections regard respectively within and across years enrollment changes for Part C plans. Then we present together all plan selection estimates for Part D plan. Finally, the two final sections report results for the Part C and D risk scores as well as the evidence from the micro-level data.

A. Within-Year Enrollment for MAPD

Table 3 displays our baseline 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 4 models: odd numbered columns include county and year fixed effects; 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 contracts with 5-star in 2012 or 2013, while the control group contains contracts with 4 or 4.5 star in 2012 or 2013 in counties without 5-star contracts. The next two panels report two robustness checks involving a matched-DID estimator (Panel B) and a placebo test (Panel C).

Table 3: 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
	Panel C: Placebo							
	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

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. 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. Panel C, treatment group contracts with 5-star in 2012 or 2013; control group contracts with 4 or 4.5 star in both 2012 and 2013 in counties without 5-star contracts. Panel D, placebo test, over the year 2009-2011 with a simulated policy introduced in 2011 (same sample as Panel A). Standard errors in parentheses clustered at county level *** p<0.01, ** p<0.05, * p<0.1.

The baseline estimates in Panel A show that the 5-star SEP has a large and statistically significant effect on the within year change in enrollment. In our baseline specifications, columns 1 and 2, 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 to normalize the enrollment changes by the existing enrollment base. The estimates that we obtain range from 7% to 9% in the baseline specifications and from 15% to 16% when including time trends.

The results in the baseline estimates are broadly confirmed by the two sets of robustness checks presented in the remaining panels of Table 3. To assess the sensitivity of our estimates to the choice of the 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-economical, 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 belonging to the common support of the propensity score between the treatment and the control groups.[7] The estimates obtained are similar in terms of both magnitude and significance to the baseline ones. Not all coefficients of the matched-DID, however, lie within the 95 percent confidence interval of the baseline estimates. In particular, the matched-DID indicates 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 in the data, we take the baseline estimate of a 15% as a more conservative estimate.

In Panel C, to further assess the robustness of our estimates, we conduct a placebo test. In particular, we repeat our analysis as if the 5-star SEP was 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. Panel D 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.

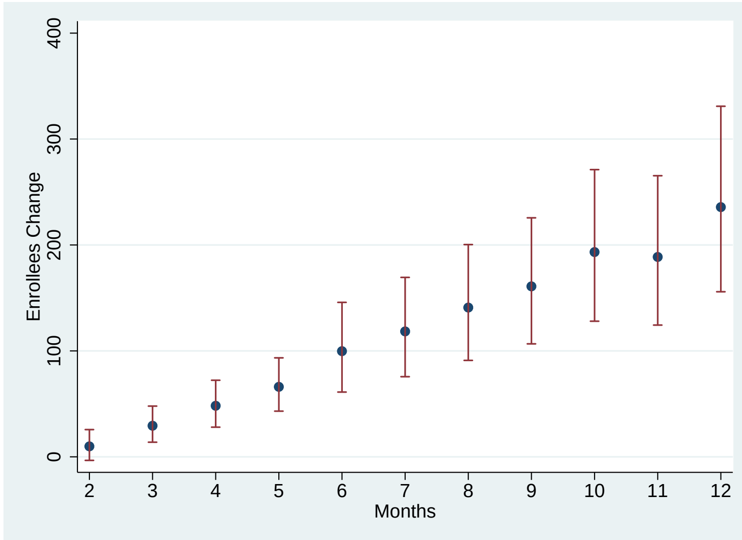
Finally, it is informative to know in which month of the year enrollees use the SEP. Thus, we consider complementing the above estimates of the December minus January enrollment change with analogous estimates for the other months preceding December. In Figure 4, we plot the estimates obtained for the same specification as in model (2) of Table 3. The effect on enrollment of the SEP appears linearly increasing 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

In Table 4, we repeat the previous analysis using as dependent variable the enrollment change across the years. The effect of the 5-star SEP is ex ante ambiguous in this case. A negative

[7] 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) that represent 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 3. All other estimates are obtained using the same specification.

effect of the policy would be compatible with consumers acting strategically by enrolling in cheaper, but less generous plans with a lower star rating and then switching to a more expensive 5-star plan only if hit by a health shock during the year. The previous estimates in Table 3 indicate that within the year switches do occur. However, this is not incompatible with increases across the years that might be driven, for instance, by the promotion of 5-star plans by CMS. The findings in Decarolis and Guglielmo (2017) also suggest that the decline in 5-star plan premiums could have bolstered the demand of enrollees in good health.

The estimates in Table 4 reveal that demand for 5-start plans did not decline across the years. No specification leads to finding a negative and significant effect. Statistical significance is achieved only for the estimates involving percentage increase and, within these cases, only the for the specifications including time trends (models 7-8). 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 7-8, we might conclude that there is evidence in favor of an enrollment increase 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 in interpreting the finding as conclusive in terms of any positive effect on across the year-years demand.

Table 4: MAPD Contracts - Across Years 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
	Panel C: Placebo							
	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. 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. Panel C, treatment group contracts with 5-star in 2012 or 2013; control group contracts with 4 or 4.5 star in both 2012 and 2013 in counties without 5-star contracts. Panel D, placebo test, over the year 2009-2011 with a simulated policy introduced in 2011 (same sample as Panel A). Standard errors in parentheses clustered at county level *** p<0.01, ** p<0.05, * p<0.1.

In any case, all estimates agree in indicating that any strategic consideration for consumers leaving 5-star plans was muted by the forces inducing a stronger demand. This finding indicates 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 also if we rule out this channel by excluding 2013 data. The placebo estimates in Panel C further confirm that the 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.

C. Across and Within-Year Enrollment for PDP

The analysis of PDP demand effects presents different challenges relative to 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 rating no lower than 4, in principle it could be less problematic for PDP to have a broader definition of the control group.

Table 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 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.

D. Part C and D Risk Scores for MAPD

The final set of results regards 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, an increased across-year enrollment can imply improved selection for 5-star plans, especially to the extent that its driving force is the demand by relatively healthy enrollees attracted by lower premiums. This same force can also be the trigger behind within-year plan switches driven by healthy enrollees looking for high quality plans, irrespective of their decreased financial generosity.

The two dependent variables on which we focus are the yearly average MAPD risk scores that CMS releases separately for 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 6.

Panel A in Table 6 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 control group is identical to what described for Table 3 and 4. All the estimates in this panel agree in showing 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 6 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 an yearly average. Could it be that this variable is unable to capture in a timely manner the high risk of those joining 5-star plans? The annual average risk score for a plan is built up by taking all of the individual-level risk scores and averaging them. So, when new enrollees join during year t, the risk scores of those enrollees will be factored into the year t average risk score. We also know from Geruso and Layton (2015) that insurers are proactive in adjusting upward the risk score of their enrollees. All this makes our measure adequate.

Nevertheless, there is a lag in how often the individual-level risk scores are updated. 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 be able to capture his higher risk only a year after the switch. This feature makes the current risk score system inadequate to deal with selection driven by within-year plan changes and this problem is especially severe in Part C where no ex post

Table 6: 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.

adjustment measures (like the Part D risk corridors and reinsurance) exist.

Moreover, a more subtle problem could, in principle, involve new Medicare enrollees. Enrollees who are enrolling in Medicare for the first time (either FFS or MA) have no diagnoses, so their risk scores are based on age/gender only and are not particularly indicative of health status. 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 since they could join any plan during any month of the year (as long as it is the first month they enroll), so this should not be a concern for our analysis.

To account for these issues, we exploit the fact that we observe 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 exclusively 2013 as the post-policy year should lead us to find less negative, if not even 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 6. 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 in magnitude than those in Panel B. Hence, these results confirm that the risk pool of 5-star plans improved and it is not a spurious correlation driven by a lagged response in the risk score measures.

E. 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?

Figure 5 shows the average risk score separately for MAPD that lose and that gain enrollees during the year.[8] 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.[9] Since 5-star plans were characterized by high risk enrollees pre 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 worst 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 not of worse risk (no selection) or even healthier (advantageous selection) than those in non-5-star plans.

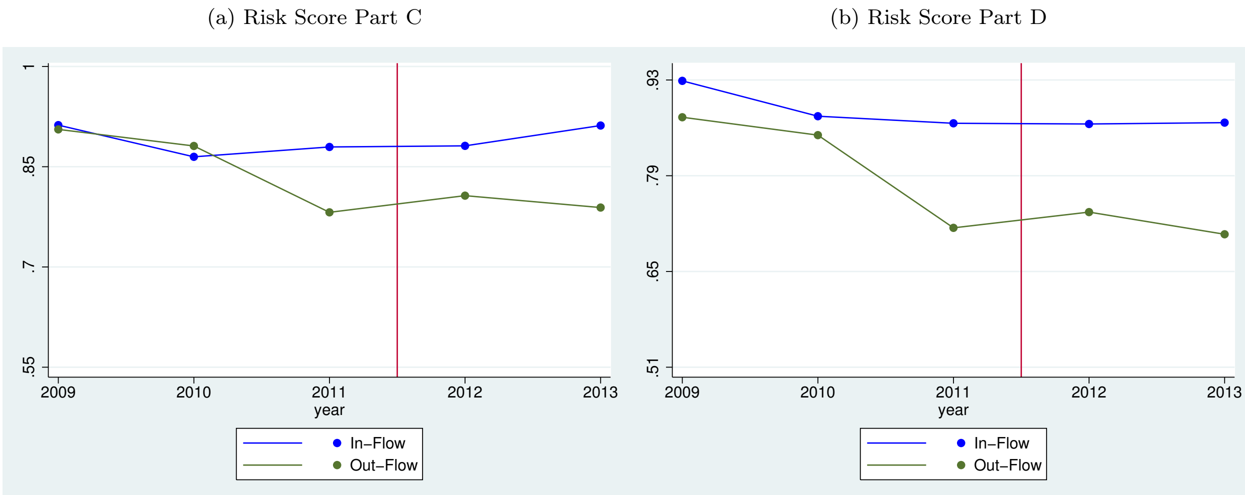
To resolve this ambiguity, we resort to the CMS Part D claims-level data to construct a a random sample of 4 million Medicare beneficiaries observed from 2011 to 2013.[10] To study within-year switching behavior to 5-star MAPD, we first exclude from the sample individuals enrolled in any 5-star MAPD continuously for all 12 months of any enrollment year. We also exclude beneficiaries who never purchase any drug under Part D. The resulting sample

[8]The average is calculated 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 2, 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).

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

[10]See Polyakova (2016) and Ho et al. (2017) for a detailed discussion of the content of these data.

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



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 the 5-star SEP.[11]

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 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 on the risk of switchers relative to non-switchers. Indeed, 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,

[11]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.

strokes, fractured hips. *Chronic Low* records the presence of chronic maintenance conditions that are not debilitating (asthma, diabetes, hyperlipidemia, etc). *Chronic High* indicates the existence of debilitating chronic conditions (osteoporosis, cancer, etc). Finally, *Mental* covers alzheimer and depression. Since a flag is recorded even if there is just 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 remaining three.

Table 7 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. The idea of this latter specification is to check whether the enrollees switching in 2012 are more likely to be those in worse health status the following year.

The main result is that, across all models, all the four *HealthStatus* variables have a negative and significant effect on the probability of switching. The estimates are highly statistically significant and their magnitude is rather robust to the inclusion of additional controls in the model specification and to the two different subsamples considered. The marginal effects are reported in Table A.3 in the appendix. This evidence is thus revealing that the enrollees most likely to switch are likely healthier than those not switching. The relatively larger coefficient for *Chronic High* which is twice that for *Chronic Low* is compatible with the anecdotal evidence that patients that are particularly sick might find harder to switch plans as their conditions make more problematic to change physicians, hospital networks or drugs. Evidence consistent with this interpretation is offered in Miller et al. (2014).

This evidence is also in line with the estimates obtained for some of the additional controls and, specifically, with the 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 PDP program or from TM without any Part D coverage. The positive estimate on Age and $Female$, instead, runs against what expected under advantageous selection. Their magnitude, however, are smaller

Table 7: 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.

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 for switching plans within the year. Although the data does 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 is reassuring 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 5-star rated plans is a rare example of a change in open enrollment rules. Therefore it represents a valuable natural experiment to learn about the effects that this kind 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 are 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 premiums and benefit generosity of the 5-star plans, while our study illustrates how enrollees responded to the combined changes in plan characteristics and within-year switching possibilities. 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 plan.

These results suggest the relevance of two main avenues for future research. First, enrollees inertia in plan choices makes prominent 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" allowing plan switches.[12]

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

[12]This is also related to Ericson et al. (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. Although convincing evidence on the roles of these two forces in the ACA exchanges is still missing, 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, but, after the carriers claimed widespread abuse of the SEP by ineligible people, CMS tightened the requirements for SEP applicants by requesting to document their eligibility. It would thus be interesting to quantify how 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, J. and Gruber, J. (2011). Choice inconsistencies among the elderly: Evidence from plan choice in the medicare part d program. *American Economic Review*, 101(4):1180–1210.

Abaluck, J. and Gruber, J. (2016). Improving the quality of choices in health insurance markets. *No. w22917. National Bureau of Economic Research*.

Batata, A. (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, J. and Lakdawalla, D. (2006). Does medicare benefit the poor? *Journal of Public Economics*, 90(1):277 – 292.

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

Cao, Z. and McGuire, T. G. (2003). Service-level selection by hmos in medicare. *Journal of Health Economics*, 22(6):915–931.

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

Conley, T. G. and Taber, C. R. (2011). Inference with difference in differences with a small number of policy changes. *The Review of Economics and Statistics*, 93(1):113–125.

Cutler, D. M. and Reber, S. J. (1998). Paying for health insurance: The trade-off between competition and adverse selection. *The Quarterly Journal of Economics*, 113(2):433–466.

Cutler, D. M. and Zeckhauser, R. J. (1998). Adverse selection in health insurance. In Garber, A. M., editor, *Frontiers in Health Policy Research, volume 1*. MIT.

Decarolis, F. and Guglielmo, A. (2017). Insurers' response to selection risk: Evidence from medicare enrollment reforms. *Journal of Health Economics*, 56:383–396.

Dorn, S. (2016). Helping special enrollment periods work under the affordable care act. Urban institute, research report.

Duggan, M., Starc, A., and Vabson, B. (2016). Who benefits when the government pays more? pass-through in the medicare advantage program. *Journal of Public Economics*, 141(Supplement C):50 – 67.

Ellis, R. P. and Savage, E. (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, K. M., Geissler, K., and Lubin, B. (2017). The impact of partial-year enrollment on the accuracy of risk adjustment systems: A framework and evidence. Working Paper 23765, National Bureau of Economic Research.

Geruso, M. and Layton, T. (2015). Upcoding: Evidence from medicare on squishy risk adjustment. Technical report, National Bureau of Economic Research.

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

Guglielmo, A. (2016). Competition and costs in medicare advantage. mimeo.

Handel, B. R. and Kolstad, J. T. (2015). Health insurance for "humans": Information frictions, plan choice, and consumer welfare. *American Economic Review*, 105(8):2449–2500.

Heiss, F., McFadden, D., Winter, J., Wuppermann, A., and Zhou, B. (2016). Inattention and switching costs as sources of inertia in medicare part d. mimeo.

Ho, K., Hogan, J., and Morton, F. S. (2017). The impact of consumer inattention on insurer pricing in the medicare part d program. *RAND Journal of Economics*, forthcoming.

Ho, K., Hogan, J., and Scott Morton, F. (2014). The impact of consumer inattention on insurer pricing in the medicare part d program. mimeo.

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

Ketcham, J., Lucarelli, C., and Powers, C. A. (2014). Paying attention or paying too much in medicare part d. *American Economic Review*, forthcoming.

Ketcham, J. D., Lucarelli, C., Miravete, E. J., and Roebuck, M. C. (2012). Sinking, swimming, or learning to swim in medicare part d. *American Economic Review*, 102(6):2639–73.

Kuziemko, N., Meckel, K., and Rossin-Slater, M. (2014). Do insurers risk-select against each other? evidence from medicaid and implications for health reform. mimeo.

Layton, T. J. and Ryan, A. (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, P. and Doshi, J. A. (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, T. (2015). The cost of removing deadlines: Evidence from medicare part d. mimeo.

Marzilli Ericson, K. 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, M. and Skinner, J. (2006). The incidence of medicare. *Journal of Public Economics*, 90(1):257 – 276.

McWilliams, J. M., Hsu, J., and Newhouse, J. P. (2012). New risk-adjustment system was associated with reduced favorable selection in medicare advantage. *Health Affairs*, 31(12):2630–2640.

Miller, K., Petrin, A., Town, R., and Chernew., M. (2014). Health plan changes in response to medicare advantage payment rates. *mimeo.*

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

Nosal, K. (2012). Estimating switching costs for medicare advantage plans. *Mimeo*.

Polyakova, M. (2014). Regulation of insurance with adverse selection and switching costs: Evidence from medicare part d. mimeo.

Polyakova, M. (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, M. and Stiglitz, J. (1976). Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *Quarterly Journal of Economics*, 90(4):629Ð649.

Shepard, M. (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 CMS (Center for Medicare and Medicaid Services). This is the same dataset as use in Decarolis and Guglielmo (2017). In particular, data on monthly enrollment for the years 2009-2013 at plan level was 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 network are from the FRF (*Formulary Reference Files*):

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

Part C and D performance data determining the star ratings were obtained from:

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

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

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

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.1 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 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.1. 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.2 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.

Table A.1: 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.2: 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			
treat_overall	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.3: 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