NBER WORKING PAPER SERIES

INSURERS' RESPONSE TO SELECTION RISK:
EVIDENCE FROM MEDICARE ENROLLMENT REFORMS

Francesco Decarolis
Andrea Guglielmo

Working Paper 22876
http://www.nber.org/papers/w22876

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2016

Insurers' Response to Selection Risk: Evidence from Medicare Enrollment Reforms
Francesco Decarolis and Andrea Guglielmo
NBER Working Paper No. 22876
December 2016, Revised February 2017
JEL No. I1,I13,I18,L22

ABSTRACT

Evidence on insurers' behavior in environments with both risk selection and market power is largely missing. We fill this gap by providing one of the first empirical accounts of how insurers adjust plan features when faced with potential changes in selection. Our strategy exploits a 2012 reform allowing Medicare enrollees to switch to 5-star contracts at anytime. This policy increased enrollment into 5-star contracts, but without risk selection worsening. Our findings show that this is due to 5-star plans lowering both premiums and generosity, thus becoming more appealing for most beneficiaries, but less so for those in worse health conditions.

Francesco Decarolis
Department of Economics
Boston University
270 Bay State Road
Boston, MA 02215
and EIEF
and also NBER
fdc@bu.edu

Andrea Guglielmo
Analysis Group, Inc
Prudential Center Plaza III
111 Huntington Ave, B
Boston, MA 02199
andrea.guglielmo.85@gmail.com

I Introduction

The behavior of insurers is a crucial component of the functioning of any insurance market. Understanding such behavior is thus key to evaluate reforms like the creation of the health-care marketplaces under the Patient Protection and Affordable Care Act (ACA) and the growingly privatized provision of Medicare throughout the Part C and Part D programs.[1] The question of how competition works in environments with potential risk selection (either advantageous or adverse) is, however, still unsettled from a theoretical perspective and there is still much to be learned on the complex interaction between market power and selection.

More specifically, nearly all the recent literature on selection markets focuses on pricing distortions while abstracting from how selection affects the broader set of characteristics of the contracts offered. The supply-side analysis presented in this paper is, instead, about how health plans respond in terms of both premiums and benefits to additional opportunities for beneficiaries to move among plans, possibly in response to health shocks. Therefore, our contribution follows in the tradition initiated by the seminal theoretical studies of Rothschild and Stiglitz (1976) and Glazer and McGuire (2000) in which plans alter their product seeking to attract good risks. Although a handful of earlier studies have already shown evidence of insurers taking actions to attract good and deter bad health risk,[2] our contribution is to provide a particularly clean identification strategy to quantify how both premiums and benefits respond to a potential change in selection driven by a policy reform stimulating consumers' mobility between plans and to do so in a context, that of Medicare Advantage, that is characterized by the presence of market power.[3]

Reliable evidence on this type of behavior is hard to collect because it is rare to observe

---

[1]Part C, also known as Medicare Advantage, is a substitute for the traditional Medicare services (i.e., Part A covering in-hospital services and Part B covering physicians, surgeons and other outpatient hospital services. Part D is a program offering prescription drug insurance.

[2]Several of these contributions, from the early study of Ellis and McGuire (2007) to the more recent contributions of Carey (2016) and Shepard (2016), will be discussed next.

[3]Using Part C data for 2006-2011, Curto et al. (2014) estimate that plan margins are on the order of 16 percent above their (variable) costs of coverage. Similar estimates are found by Guglielmo (2016) with a shorter dataset (2008-2011): by looking separately at plan types, he finds that HMOs and LPPOs generate the most profit per enrollee, amounting to a markup of 12 percent, while PFFS plans' markup is approximately 9.5 percent. These estimates are also broadly consistent with the average markup of 13 percent reported the MedPAC annual report 2010.

changes in selection risk within a market. Furthermore, even when selection risk changes for a subset of plans, it is often impossible to find a set of plans that can serve as a valid comparison group since the equilibrium in the whole market is affected. Our analysis overcomes this difficulty by exploiting the combined effects of a Medicare reform that altered the potential selection risk of the highest quality (5-star) Part C and D plans and the geographical dispersion of such plans over the US. This allows us to separately observe treated and control geographical markets both before and after this policy change, thus allowing a quantile differences-in-differences approach that we use to study distributional changes in contract features in the treated markets relative to the control ones. Our main finding is that the policy triggered a response that involved not only changing premiums, but also adjusting benefits. This made 5-star plans more appealing for most beneficiaries through lower premiums, but less so for those in worse health through higher out of pocket costs.

The starting point of our analysis is a Medicare reform changing the enrollment rules, most notably allowing enrollment outside the open enrollment period for a subset of plans. As in most insurance markets, beneficiaries select their Part C or D plan for coverage year t during a window of time in the fall of year $t-1$.[4] However, starting with the enrollment year 2012, a reform allowed enrollees to switch to 5-star Part C or D plans at any point during the year. Despite the official motivation for this reform (known as "5-star Special Enrollment Period" or "5-star SEP"), which was to foster enrollment into high quality plans, the reform exposes 5-star plans to an evident selection risk: enrollees could initially select cheap plans and then move to expensive 5-star plans with generous coverage only after being hit by health shocks. The selection risk associated with within-year plan changes is different from the typical selection problem studied in the existing Medicare literature involving choices made in the open enrollment period and is potentially more severe as people select plans after learning their health status. Limiting this type of selection is typically seen as important for the proper functioning of managed care markets and, indeed,

---

[4]While the open enrollment period length can very from 2 months to 2 weeks in employer sponsored health insurance, the open enrollment period in Medicare is fixed. Specifically in Medicare, as well as in the ACA exchanges, the open enrollment period is from October 15th to December 7th. This open enrollment period only applies to those who are already Medicare beneficiaries and not to individuals turning 65 who become eligible for Medicare.

this the logic behind the penalties for waiting beyond age 65 to join Part D and Medigap, for the individual mandate in the ACA and for insurers's resistance to expand the set of "qualifying life events" allowing plan changes.[5] Moreover, institutional remedies for selection that exist in both privatized Medicare and the ACA exchanges are currently not arranged to deal with selection originating from within-year plan changes.[6]

To study the impact of this reform, we exploit the heterogenous presence of 5-star plans across geographical markets. Due to regulatory reasons, the US is segmented into geographically separated markets both for Part C - where insurers offer plans at the county level - and for Part D - where insurers offer plans at regional level. Since not all geographical markets have 5-star plans, some markets were affected by the reform while others were not. Our empirical strategy exploits this difference, together with the robustness to manipulations of the star rating in the first two years after the policy change, to identify the causal effect of the policy on various features of the plans supplied. In particular, the methodology that we use is a quantile-based difference-in-differences analysis (Chetverikov, Larsen and Palmer (2015)) that we use to estimate distributional changes in the treated markets (those with at least one 5-star plan) and compare them to control markets (those with at least one 4 or 4.5-star plan, but no 5-star plans). Since, during our sample period, we observe 160 treated counties for Part C, but only 2 treated regions for Part D, we focus our analysis on the plans active in Part C, most of which also bundle together Part D benefits. These plans are usually referred to as MA-PD plans.[7]

We analyze how the distribution of both premiums and generosity changes in response to the 5-star SEP treatment and find a tendency for premiums to increase in the medium-low end of the premium distribution and to decrease in the medium-high end of the distribution, where 5-star plans are located. To measure the effects on plan generosity, we look at three measures of the out of pocket cost which have the benefit of aggregating all non-premium and non-customer service benefits into expected cost measures. The first is the Part C maximum out of pocket (MOOP). We find that the MOOP remains unchanged for plans in the high end

---

[5]In theACA exchanges, for instance, these events include marriage, release from prison, and childbirth.

[6]For instance, the enrollees' risk score is recalculated only on a yearly basis.

[7]We do not analyze, instead, the market for Prescription Drug Plans (PDP) active exclusively in Part D.

of the MOOP distribution, but tends to worsen for plans at the low and medium end of the distribution. Since 5-star plans are among those with a low MOOP before the reform, this result implies a worsening of their generosity. We find the same result when looking at our second proxy for benefits, the Part C plan simulated out of pocket cost (OOPC) of enrollees in poor health. For the third measure, the simulated OOPC of enrollees in excellent health, instead, we find that the 5-star SEP did does not cause changes in the OOPC distribution.

Among the additional coverage generosity measures that we observe, an interesting one for which we observe the opposite pattern relative to what described above (i.e., 5-star plans becoming more generous) is the Part D deductible. Given the importance of the deductible for beneficiaries switching to 5-star plans during the year, we argue that this is consistent with a strategic response by insurers. We also use the same empirical strategy to study several "soft" quality measures behind the star rating and show that 5-star plans do not worsen on those. We conclude that the insurers' response entailed making 5-star plans more appealing than competing plans for most consumers (by lowering premiums and deductibles), but less so for the less healthy enrollees (by worsening generosity for enrollees in poor health).

Finally, to better understand the interaction between competition and the effects of the 5-star SEP, we repeat the analysis separately for markets where there is a monopolist insurer for 5-star plans and for markets where there is competition (duopoly) in the supply of 5-star plans. The most interesting result is that competition among 5-star insurers seems to exacerbate the extent to which these insures try to cream skim the market by worsening their plan generosity. Consumers in duopoly markets are more likely to be negatively affected by the 5-star SEP: while the premium changes in the two cases are similar, the increase in the OOPC for poor health enrollees is about twice in duopoly relative to monopoly markets.

A simultaneous reform that, starting in 2012, bolstered plan payments in proportion to their star rating requires particular care on how the earlier results should be interpreted. Indeed, the evidence on declining premiums might be in part due to the pass-through of the extra payments for 5-star plans. Nevertheless, by comparison with 4/4.5-star plans that received similar, albeit smaller payment increases we illustrate how 5-star plans experienced a substantially stronger premium decline. Furthermore, the evidence on sharper declining

benefits for 5-star plans cannot be explained by the payment reform as a pass-through effect of these higher payments would have gone in the direction of higher, not lower benefits. Finally, the payment reform cannot explain the heterogenous decline in benefits for enrollees in different health status or the difference we find between monopoly and duopoly markets.

From a policy perspective, our results offer several contributions. First, they are one of the first comprehensive assessments of a complex, but little analyzed piece of regulation. The adoption of the 5-star SEP to boost 5-star plans enrollment was a risky choice from an ex ante perspective due to its potential to trigger substantial shifts in plan risk pools. It is therefore of great policy relevance to document both what it produced and what this implies for other possible policy reforms. Regarding the latter, our main insight is that insurers have the ability to design plan features even in the context of the tightly regulated Medicare market by changing not only easily observable features - like premiums - that a regulator can target, but also harder to measure financial generosity measures and soft quality features. Clearly, while the sophisticated reaction by insurers might have helped making the 5-star SEP reform successful in terms of improving 5-star plans enrollment without worsening their selection, it also underscores the complexity of designing rules capable of steering the market toward the goals set by the regulator. Indeed, even for the 5-star SEP, the different effect that we estimate for enrollees in different health status highlights a drawback of this policy and, more generally, a difficulty of relying on competition in selection markets.

Related literature - This study contributes to different strands of the literature on health insurance, especially within the context of insurers' response to potential selection risk. Following Layton, Ellis and McGuire (2015), it is useful to categorize the economic analysis of insurers' behavior into two broad groups, originating from the two seminal studies of Akerlof (1970) and Rothschild and Stiglitz (1976).[8] The first group of studies, descending from Akerlof (1970), looks at insurers' pricing choices when health plans have fixed characteristics. Within this framework, the early empirical work of Cutler and Reber (1998) has been recently complemented by a large number of studies proposing empirical methods to estimate

---

[8]Several recent studies theoretical studies, including Mahoney and Weyl (2014), Azevedo and Gottlieb (2015), Farinha Luz (2015), Lester et al. (2015) and Veiga and Weyl (2016), exemplify well how the theoretical literature is still hotly debating between these two approaches.

welfare and evaluate counterfactual policies through the formulation of structural models of plan demand and supply (see Einav, Finkelstein and Cullen (2010), Einav and Finkelstein (2011), Curto et al. (2014), Bundorf, Levin and Mahoney (2012), Lustig (2012), Starc (2014), Guglielmo (2016)).[9] The second group of studies, originating from Rothschild and Stiglitz (1976) (henceforth RS), endogenize not only premiums, but also benefits. Glazer and McGuire (2000) extended this theoretical framework to the setting of managed care health insurance where premiums are regulated, but insurers can use "service-level selection" to make their plans relatively more appealing to enrollees of different risk type. Our paper contributes to the empirical literature based on this endogenous contract framework which includes: Ellis and McGuire (2007) who apply to Medicare Part C the insights from Frank, Glazer and McGuire (2000) showing that services that are predictable (i.e., enrollees can foresee their future usage) are those rationed tightly; again in Part C, evidence of strategic benefit design in order to exploit the imperfect risk scoring is presented in Cao and McGuire (2003) and Batata (2004), but more recent studies have argued that risk adjustment drastically reduced it (McWilliams, Hsu and Newhouse (2012), Newhouse et al. (2013) and Brown et al. (2014)); in the Part D context, the study of Carey (2016) illustrates how insurers design more favorable benefits for drugs that treat profitable diagnoses as compared to unprofitable diagnoses, where the variation in diagnoses' profitability is driven by technological change after risk adjustment re-calibration (for new drug entry and generic competition); again in Part D Polyakova (2014) and Ho, Hogan and Scott Morton (2014) find evidence of selection and discuss how that interacted with the plan offerings by insurers.[10]

In addition to the literature on selection in insurance markets, our study also contributes to the analysis of how insurers respond to regulation. Thus, it is also related to other recent empirical studies that address this issue in the context of Medicare, like Decarolis (2015)

---

[9] See Einav, Finkelstein and Levin (2010) for a more complete survey of this literature.

[10] A closely related analysis is also that of Kuziemko, Meckel and Rossin-Slater (2014) on how competition in the presence of risk selection in Medicaid managed care leads to a worsening of outcomes for enrollees in poorer health conditions. Furthermore, Shepard (2016) presents in the context of the Massachusetts subsidized health insurance exchange an analysis of how selection interacts with the choice of the plan's hospital network, illustrating how the preference of high cost enrollees for "star" hospital would lead insurers to drop such hospitals from their network. Our broader focus on contract characteristics is also related to Hendel and Lizzeri (2003) and Crocker and Moran (2003) who argue that greater ex-ante commitment may reduce adverse selection and, thus, may increase insurance provision.

for Part D and Geruso and Layton (2015) for Part C. Finally, our analysis of how insurers affect soft quality measures of the offered plans is related to the issue of the public disclosure of quality measures analyzed in Glazer and McGuire (2006).[11] At a very general level, our findings about how firms adjust product features different than premium is an important contribution to the growing empirical literature on endogenous product characteristics (Crawford (2012); Fan (2013); Wollman (2014)). In most of the industrial organization literature, product characteristics are taken to be exogenous because it is to difficult to analyze when they are chosen, but our study isolates a clean setting in which it is possible to analyze multiple endogenous product characteristics. This is especially relevant in insurance markets where products are contracts characterized by multiple, simultaneously determined features.

Finally, a few demand-related papers have already stressed the relevance of the Medicare star rating system for plan choices (see Abaluck and Gruber (2015), for Part D, and Reid et al. (2013) and Darden and McCarthy (2014), for Part C). The specific impacts of the 5-star SEP on the demand for plans is analyzed in Madeira (2015) and Decarolis, Guglielmo and Luscombe (2016). The former study, uses consumer-level data in the Part D market to study plan switching with regard to the presence of behavioral biases in enrollee choices and finds that at least some Medicare beneficiaries are present-biased. For these enrollees tending to procrastinate choices, the 5-star SEP leads to a drop in enrollment in 5-star plans, driven by an overall increase in inertia. Decarolis, Guglielmo and Luscombe (2016), instead, uses both Part C and D data. For Part C, it finds that the introduction of the 5-star SEP caused an increase in within-year enrollment of 5-star plans amounting to 7% to 9% of their enrollment base at the beginning of the enrollment year (January), but it did not significantly affect plan switching across years. The increased enrollment into 5-star plans is not associated with worsening of the risk pools for these plans. The following analysis complements these findings with evidence on the insurers's response to the 5-star SEP reform.

---

[11]Related applications involve the cases of how cardiac surgery report cards led to selection by providers David Dranove and Satterthwaite (2003) in New York and Pennsylvania and the similar evidence on the Nursing Home Quality Initiative by Werner et al. (2009) and Lu (2012).

II Institutions: Rating System and Policy Changes

The Medicare Part C and D programs share several organizational features. Both programs entail Medicare beneficiaries choosing a plan from a menu of plans offered by private insurers. Detailed regulations, mostly from the Center for Medicare and Medicaid Services (CMS), contribute to the determination of both the types of plans offered and their premiums. The two programs, however, differ along many dimensions: Part C is a privately provided alternative to traditional Medicare, TM. Thus, plans must cover Medicare Part A and Part B benefits (except hospice care), but can offer additional benefits.[12] Part D, instead, is a program with voluntary enrollment that provides coverage for prescription drugs. For Part C, nearly all Medicare Advantage (MA) plans also include Part D benefits (i.e, these are MA-PD plans).[13] TM enrollees can obtain Part D benefits by enrolling in stand alone PDP plans. This section describes three key regulatory aspects: the plan rating systems and the two reforms linking ratings with enrollment periods and subsidies, respectively.[14]

A. Rating Systems for Part C and D

To help beneficiaries select plans and to monitor the market, CMS rates plans on a 1 to 5 scale, with 5-stars indicating the highest quality. More precisely, CMS assigns ratings at the contract level and so every plan covered under the same contract receives the same rating.[15] Information about plan performance has been collected since 1999, but the introduction of the star rating system started only in 2006 for Part D and in 2007 to Part C.

The details concerning the rating system are fairly complex and have changed over time. The essential aspect is that different data sources (enrollees surveys as well as CMS administrative data, and data from plans and other CMS contractors) are used to collect information on a broad set of indicators. The process through which CMS calculates the star rating in-

---

[12]Medicare Part A includes inpatient hospital, skilled nursing, and some home health services. Medicare Part B includes physicians' services, outpatient care, and durable medical equipment.

[13]The subset of plans offering both Pat C and D coverage are usually indicated as MA-PD plans. With a slight abuse of notation we will typically refer to all Part C plans as MA plans.

[14]Newhouse and McGuire (2014) and Duggan, Healy and Scott Morton (2008) are recent studies discussing more broadly the institutional aspects of Part C and D respectively.

[15]In Part C, a contract is a particular product type (HMO, PPO or Private FFS) covering a specific service area (i.e county or group of counties), while a plan is finer specification of benefit package that include type of coverage, premium, copayment, etc. In Part D, a contract typically indicates a drug formulary and, then, each plan within the contract applies different conditions (for instance copays) to the same formulary.

volves several steps. At the most disaggregated level there is a large number of "individual measures," which are aggregated into a smaller number of "domain measures" and finally into the "summary rating" through a complex weighting system.[16] Table 2 reports the domain measures: for Part C, they cover features such as clinical quality, patient experience, and contractor performance; for Part D, they cover aspects such as call center hold time, members' ability to get prescriptions filled easily when using the drug plan, and plan fairness in denials to members' appeals. The overall rating, expressed in a 5-Star scale with increments of half a star, is released every year in October on the CMS Plan Finder web site.

A notable feature of the rating system is that it is hard to manipulate for insurers, especially in the short run. There are at least three reasons for this: first, CMS changes the system from year to year in terms of both which parameters are evaluated and how they are aggregated into the overall rating. This aspect is particularly salient given the large number of different measures that are evaluated, as shown in Table 2. Second, ratings on individual measures are assigned by comparing the *relative performance* of each contract to the entire population of contracts so that manipulations would require detailed information on all competing contracts. Third, and most crucially, the rating is based on lagged data: year t ratings (released on October of year $t-1$) use data for the period between January of year $t-2$ and June of year $t-1$. To exploit these features limiting potential rating manipulations in the short run, we focus on the first two years after the enrollment reform.

Very few contracts obtain the 5-star maximum. In 2012 and 2013, for instance, only two firms offer 5-star PDP, while for Part C seven firms offer 5 star plans, as shown in Table 3. Regarding the geographical distribution of plans, out of the 34 regions into which Part D divides the United States, only 2 regions (region 3, New York, and region 25, formed by 7 midwest states) had a 5-star PDP. 5-star plans are more frequent among MA. However, while PDP must be offered to all counties within a region, Part C plans are offered at the county level. Figure 1 presents a heat map showing the offerings of MA plans. In 2012-2013 period, 5-star plans are offered in 160 counties belonging to 10 different states and spanning

---

[16]More precisely, for PDP and MA plans not offering Part D, the summary rating is also the overall rating. For MA plans, the Part C and D summary ratings are combined to obtain an overall rating. A more complete description of the process through which CMS calculates the star rating is detailed in the web appendix.

almost all the U.S. geographical areas, with the relevant exception of the center-south area. This geographical dispersion of 5-star MA plans plays a fundamental role in our empirical strategy and we return to it in the next section.

B. Demand Side Reform: Plan Rating and Enrollment Periods

Generally, beneficiaries enroll in a plan between mid October and early December of the year before the coverage period (Open Enrollment Period, OEP) and must keep the same plan for the entire coverage year (i.e., from January though the end of December). Exceptions to the OEP, known as Special Enrollment Periods (SEPs), permit enrollees to change plans, but are typically confined to special circumstances.[17]

Starting with the 2012 coverage period, CMS introduced a new type of SEP linked to the star rating system. This reform allows all beneficiaries to enroll in a 5-star Part C or D plan at any point in time.[18] This SEP rule can only be used once per year and is available even to enrollees already in a 5-star plan, but who want to switch to another 5-star plan. Coverage with the new 5-star plan takes effect the first day of the month following the enrollment. Similar to any other enrollment request, 5-star plans must accept all applicants. The SEP cannot be used to enroll in a plan that does not have an overall 5-star rating, even if the plan receives 5-stars in some rating categories, or if the plan is in the same parent organization.[19] CMS has extensively advertised this new SEP rule in its communications to consumers. Insurers were publicly informed of the introduction of the 5-star SEP on November 2010. Since the next round of plan bids was in June 2011 for the menu of plans to be offered in 2012, then we can consider 2012 as the first year from which we can expect to see reactions in plan features driven by the policy change.

C. Supply Side Reform: Plan Rating and Insurers' Payments

Payments to insurers come mostly from various types of Medicare payments and, only in

---

[17]The most relevant SEPs are: (i) for change of residency, including moving to a nursing home; (ii) for low income people (dual eligible or qualifying for the LIS or for SPAPs); (iii) for people who enroll in a MA plan when they are first eligible at age 65 get a "trial period" (up to 12 months) to try out MA. This SEP allows them to disenroll from their first MA plan to go to TM.

[18]See the *2012 Newsletter* at `http://www.cms.gov/Medicare/Prescription-Drug-Coverage/PrescriptionDrugCovContra/downloads/Announcement2012final2.pdf`.

[19]There is also a special provision for which, if the enrollee uses the 5-Star SEP to enroll in either a 5-star PFFS plan or a 5-star Cost Plan, then he gets a "coordinating Part D SEP" allowing him to enroll in a stand-alone PDP, or in the Cost Plan's Part D optional benefit, if applicable.

small part, from enrollees premiums, see Newhouse and McGuire (2014) and Decarolis (2014). The ACA of 2010 reformed various aspects of the system and, crucially, introduced a link between the star rating system and payments.

This supply side reform exclusively affects Part C and, like the enrollment reform, became effective in 2012. Essentially, the reform sought to reduce overall plan payments, but also to make payments relatively more generous for higher quality plans than for lower quality plans. For the purposes of our study, this reform implies that after 2012 per enrollee payments of 5-star plans are more comparable to those of 4 and 4.5 then to those of plans with lower ratings. In essence this is due to how this reform affects the following two features of the payment system.

The first is the benchmark. The benchmark is a function of what TM spends in the plan's service area. CMS determines the payment to an MA plan by comparing its "bid" (the amount the insurer requests to enroll a beneficiary in the plan) to the service area benchmark. Plans with a bid below benchmark (the typical case) receive their bid plus a rebate based on the difference between benchmark and the bid. The ACA reform aligned benchmarks more closely with TM spending[20] and, instead of the flat 75% rebate used before 2012, introduced a variable rebate, ranging from 50% to 70%, linked to the plan star rating.[21]

The second is the bonus. Bonuses were introduced in 2012 to bolster payments for high-quality plans by proportionally increasing their benchmarks. For instance, in 2012 the bonus for 5-star plans is 5% of the benchmark. Thus, a 5-star plan with a bid below the benchmark receives a rebate equal to 73% of 1.05 times its service area benchmark. While under the ACA bonuses were reserved for plans with 4 or more stars, CMS used its demonstration authority to extend bonuses to plans with 3 or more stars. In the period that we study, benchmarks are increased by 4% for 4.5-4 star plans, by 3.5% for 3.5 star plans, by 3% for 3 star plans and plans that are too new or with too few enrollees to be rated.[22]

---

[20]It ties the benchmarks to a percentage of mean TM cost in each county and caps them at the pre-ACA level. These benchmarks are phased in from 2012 to 2017 by blending them with the old benchmarks.

[21]The new rebates are phased in from 2012 to 2014. In 2012, the rebate equals the sum of two-thirds of the old rebate amount and one-third of the new rebate amount. In 2013, the rebate equals the sum of one-third of the old rebate amount and two-thirds of the new rebate amount. From 2014 onward, the rebate is 70% for 5-4.5 star contracts, 65% for 4-3.5 contracts and 50% for the rest of the contracts.

[22]The demonstration is expected to cost more than $8 billion, making it more costly than the combined

III   Conceptual Framework

It is useful, before presenting the data, to illustrate through a simple conceptual framework the incentives produced by the regulations described above for the insurers. Consider the typical RS setting with a pool of consumers having homogenous preferences over wealth and heterogeneous expected health care costs, who make up the demand side of a health insurance market. Suppose that on the supply side there are two insurers setting premiums and benefits and that only one of them is labeled 5-star.[23] Now, extending RS, suppose that enrollees exhibit inertia in the form of a switching cost and that the 5-star SEP reform acts as a reduction in this switching cost, but only if the switch is toward the 5-star plan.[24] Notice that, by comparison with RS, given an exogenous initial allocation of consumers, a plan seeking to expand its enrollment will have to alter price and/or benefits by a larger margin to overcome the inertia. Hence, the asymmetric reduction in switching costs driven by the 5-star SEP increases the incentive for the 5-star plan to lower its premium and benefits: the 5-star SEP increases the sensitivity to contract changes for consumers who are not the 5-star plan's enrollees, without altering the sensitivity to contract changes of its current enrollees. Under most initial allocations, since enrollees value premiums equally, but benefits are valued more by the sickest, who are also more costly to serve, the 5-star plan has an incentive to lower premiums to expand its pool of enrollees, capitalizing on the reduced switching cost, while limiting its appeal for the bad risk by lowering generosity.[25] The main hypothesis that our analysis will seek to test is whether the data indicates declines of premiums and benefits compatible with a RS framework, that are also present when consumers exhibit inertia.

In a market with multiple, competing 5-star plans, the response by insurers to the policy might reflect the degree of market power of the different plans. Although the presence of mar-

---

cost of all 85 other Medicare demonstrations that have taken place since 1995. See Layton and Ryan (2014) for a first assessment of its effects.

[23]Despite the presence of risk adjustment in Part C, recent empirical evidence from Brown et al. (2014) shows that under the current arrangement incentives for risk selection are still present.

[24]Inertia in health insurance demand is a well documented issue, see Nosal (2012) for a study on Part C.

[25]The supply-side reform boosting payments to 5-star plans described above might be partially passed-through to consumers in the form of lower premiums or higher benefits. Furthermore, depending on the nature of the inertia, it might be that the enrollees that switch during the year to a 5-star plan are relatively healthier than those already forming the risk pool of the 5-star plans, thus lowering the plans average risk score. These forces might exacerbate premiums decline, but also limit benefit reductions.

ket power makes an exact characterization of the equilibrium hard to formulate (Einav and Finkelstein (2011)), recent results by Lester et al. (2015) show that greater cream-skimming occurs when insurers face more competition. In our context, competition between 5-star insurers is likely to put relatively more downward pressure on benefits than on premiums.[26]

For a managed care health insurance markets like Part C and D, insurers are limited in their design choices. Thus, together with premiums and expected out of pocket costs, which we will use to proxy for benefit generosity, the next section will also look at a few other plan features. The first measure is the maximum out of pocket cost which, clearly, is going to be salient for the enrollees in poorer health conditions as these individuals are likely to have larger out of pocket expenditures. The second is the Part D deductible. Under the 5-star SEP, enrollees switching to a 5-star plan during the year face a deductible equal to the maximum between zero and the difference between the new and old plan deductibles. Thus, if 5-star plans were to ask for high deductibles, this would reduce their appeal for every consumer considering a within-year switch and especially so for the healthiest ones who are unlikely to have made large deductible payments. On the other hand, for non 5-star plans increasing the deductible might not trigger a major loss of enrollees since these enrollees are aware of the possibility of switching to 5-star plans.[27] Since the deductible is a transparent and understandable plan feature for enrollees, the logic in Frank, Glazer and McGuire (2000) suggests that insurers are likely to use it as a tool to influence consumer choice.

Finally, a last group of plan features involves "soft" quality measures. In a more realistic setup where enrollees have multiple dimensions along which they value plan features, the plan's quality in terms of "customer care" can be an additional screening tool. In particular, to the extent that enrollees' switching behavior is driven not only by health shocks, but also by negative experiences with service quality, and that consumers attentive to quality tend to be good risk (in the spirit of Fang, Keane and Silverman (2008)'s advantageous selection), then 5-star plans might respond to the 5-star SEP by enhancing their perceived quality.

---

[26]Lowering benefits can have smaller effects in terms of discouraging initial enrollment choices, when plan choices are made based on an expectation of health status, and large effects on within-year plan switches, when choices reflect new information. Higher premiums, instead, discourage equally both enrollment choices.

[27]The opposite incentives for 5 and non-5-star plans might thus move their deductibles further apart, with the one of 5-star plans becoming relatively more generous, despite the overall generosity of the plan declining.

IV Data and Descriptive Evidence

Our analysis is based on publicly available data released by CMS describing plan/contract characteristics. For all the MA-PD plans offered, we observed their monthly enrollment, Part C and D premiums, star rating (with all the associated individual components), several out of pocket cost measures and several Part D features (deductible, extra coverage in the gap and multiple measures of drug generosity). We also use the Area Health Resource File released by the Health Resource Service Administration to assess a number of county-level demographic, economic and heath indicators. We focus on the period from 2009 to 2013.

Although data on the generosity of Part C benefits is available to researchers, we decided to capture this crucial element of our analysis focusing on three synthetic measures related to out of pocket costs that CMS releases at plan level. The first is the maximum out of pocket an enrollee can be responsible for within a year from in-network utilization.[28] The other two measures are representative expected cost measures. Specifically CMS applies plan designs to representative consumers to simulate an expected out of pocket cost (OOPC). This expected cost measure is calculated separately for representative enrollees in different health conditions and, in our analysis, we focus on the two polar cases of "poor health" enrollees and "excellent health" enrollees. These measures are commonly used by researchers and practitioners to approximate plans' generosity (see Stockley et al. (2014); Guglielmo (2016)) and are shown to enrollees when they choose their plans.

Table 1 reports summary statistics separately for contracts with different star ratings (5 star; 4 and 4.5 star; less than 4 star), for different set of counties (160 treated counties and 1,084 control counties) and for different time periods (2009-11 and 2012-13). We present the data at contract and not at plan level both because the rating does not vary among plans under the same contract and because missing enrollment data are more common at plan than at contract level.[29] Our main dependent variables are premiums (for both Part C and D) and the three out of pocket cost measures described above (MOOP and OOPC for

---

[28] Any in-network covered utilization will be paid in full by the insurer once out of pocket expenditure reaches this threshold.

[29] A subset of our measures are available only at plan level. We aggregate them at contract level by weighting the plan characteristics by the enrollment of the plan. As discussed below, we tested the robustness of our results to aggregation (i.e. simple average), the results are reported in appendix.

enrollees in "poor health" and "excellent health"). Additionally, among the various Part D features that we observe, we report in the analysis below the Part D deductible.

For these six outcome variables, Figure 2 reports their evolution over time among different sets of contracts. The visual inspection of the plots immediately suggests a few tendencies that will be confirmed by the subsequent regression analysis. In particular, regarding the Part C premium (plot (a)), the average value among 5-star plans declines sharply in 2013, relative to the previous years, passing from about $800 per year to slightly less than $600 per year. In comparison, the premiums of all other star rating groups, for both treatment and control counties, do not express such a pronounced decline post 2011. A similar description applies for the Part D premium (plot (b)) where, however, the premium decline of the 5-star plans post 2011 is also accompanied by an increase for almost all the other plan groups. Regarding the OOPC measures, plots (d) and (e) are suggestive of an increase in both the MOOP and the OOPC for enrollees in poor health that after 2011 is steeper for 5-star plans relative to all other plan groups. No apparent change is instead visible for the OOPC of enrollees in excellent health.

Although most of the analysis below will focus on these six main outcome variables, auxiliary results will also explore the effects of the 5-star SEP on additional features. These involve soft quality measures, such as *health care quality*, *customer service* and *drug access*. The latter set of measures are all components of the star rating system illustrated in the previous section for which we take the appropriate time window.[30] Finally, an analysis involving additional plan observable characteristics is reported in the web appendix.[31]

Regarding the insurers offering 5-star MA plans, Table 3 shows that there are seven of these insurers offering plans in 2012-2013.[32] A first interesting feature revealed by the table is the fact that the 5-star SEP did not trigger any major entry/exit of plans. Table 3 illustrates

---

[30]As mentioned earlier, certain components of the rating enter its calculation with a time lag and, hence, their usage requires attention to their period of reference.

[31]The outcomes are: the Part D OOPC for enrollees in poor and excellent health, the number of most frequently purchased drugs included in the formulary, the number of drugs without utilization restrictions and the Part C and D risk scores.

[32]The overall set of firms active on the supply side of Part C and D are many and heterogeneous. They range from large scale, nation-wide insurers like United Healthcare and Humana, to a plethora of small local companies. Almost all insurers offering Part C also offer Part D, but some major Part D insurers, like CVS Caremark, are not present in Part C.

this point by reporting, for each of the seven insurers offering at least one 5-star contract in 2012-2013 (i.e., treated contracts), the total number of counties where these treated contracts where offered in all the years from 2009 to 2013. This table indicates that there was no exit of these plans and, instead, they expanded their presence to more counties in both 2012 and 2013 relative to the previous years. This is clearly relevant to alleviate any concern that a selective exit from certain counties might drive any of the results described below.[33]

A second feature related to these seven insurers is that only three of them, Group Health, Humana and Kaiser Foundation, are major national players. However, while the 5-star plans of Group Health and Humana are offered only in a limited geographical area (Wisconsin for Humana and Oregon-Washington for Group Health), Kaiser has 5-star plans in various states: California, Colorado, Hawaii, Oregon and Washington. Moreover, Kaiser's 5-star contracts have large market shares in all of these states, ranging from 12 to 48 percent of the relative markets. For Group Health and Humana, the market shares of their 5-star plans are smaller, but in both cases greater than 5 percent.[34]

The small number of insurers with relatively large market shares, makes it both feasible and interesting to look at the possible strategies with which these insurers responded to the 5-star SEP. In fact, both Humana and Kaiser offer non-5 star plans in control counties, where no 5-star plan is offered by any company, which allows for some additional descriptive comparisons. The most surprising aspect that we find is that Humana and Kaiser seem to follow different strategies. Humana reduced generosity without much change in premiums, while Kaiser decreased premiums without a measurable change in generosity. Specifically, comparing the periods before and after the 5-star SEP, Humana's 5-star plans offered in Wisconsin lower their generosity (the average MOOP grows from $3,400 to $6,260), substantially more than what is done by both its 4.5 star plans also offered in Wisconsin (the average MOOP grows from $4,500 to $6,331) and its 4.5 star plans offered in other Midwest counties (the average MOOP grows from $3,400 to $4,358). In the same period, the average premium

---

[33]It is also important to point out that CMS poses limits to the exit of plans as it can impose a two year ban to a firms that retires all its contracts from MA.

[34]For expositional purposes, the market share is calculated considering as the geographical markets the 34 regions of Medicare Part D, and not individual counties. The goal here is to identify large insurers with a national, or at least regional, footprint.

of 5-star plans registers an increase, but in line with that of the 4.5 plans in Wisconsin. For Kaiser, instead, we can compare its 5-star plans with the 4.5 star plans it offers in Georgia. We observe that generosity remains nearly identical for both the 5-star plans (the average MOOP goes from $3,200 to $3,233) and 4.5 star plans (the average MOOP remains identical at $3,400). Average premiums, however, decline slightly more for 5-star plans than for 4.5 star plans (Part D premiums decline from $178 to $124 for 5-star plans, while they increase from $18 to $25 for 4.5-star plans; Part C premiums, instead, decreases $807 to $584 for 5-star plans and from $499 to $421 for 4.5 star plans) .

This descriptive evidence is suggestive that insurers' response to the policy involves changes in both premium and generosity dimensions. To draw more consistent conclusions about such responses, however, it is necessary to take into account how not only 5-star plans response, but also how their competitors reacted to the policy change. Insurers offering non 5-star plans in markets with 5-star plans face the possibility of losing enrollees during the year and, accordingly, of experiencing changes to their risk pools. Indeed, they might face a worsening of selection if 5-star plans increase their cream skimming activity. We describe below an empirical strategy that aims to detect this type of insurers' responses.

V Empirical Analysis

A. Empirical Strategy

The empirical strategy that we pursue is a form of difference-in-differences (DID) strategy. The two key features of our approach are as follows. First, our unit of analysis is the county, and not the contract (or plan). As discussed above, since all contracts offered in a county with at least one 5-star contract can respond to the 5-star SEP reform, our interest is in understanding how the market (i.e., the county) responded to the policy change. Hence, we label *treated counties* those with at least one 5-star contract in either 2012 or 2013, and as *control counties* those having either 4 or 4.5 star contracts as their highest rated contracts.

The second feature is that, to capture the changes in how the overall market readjusts, we pursue a quantile-based DID analysis. This allows us to evaluate changes along the

whole distribution of each one of the dependent variables that we will consider (premium, deductible, etc.). The goal is to understand how the 5-star SEP affects the nature of competition within a market. For example, in the case of the premium, a 3 star contract with a low premium and a 5-star contract with an high premium would probably have a different reaction to the 5-star SEP, and analyzing different quantiles of the premium distribution within a market can be more informative than just focusing on the mere average effect.

We model the τth quantile of the distribution of characteristic Y in county c at time t as:

$$Y_{ct}(\tau) = a_c(\tau) + b_t(\tau) + \beta(\tau) \times 5StarCounty_{ct} + \varepsilon_{ct}(\tau) \qquad (1)$$

where the coefficient of interest is $\beta(\tau)$, the effect of the 5-star SEP on the dependent variable Y for the τth quantile. For instance, when analyzing the Part C premium, estimating $\beta(0.2)=2$ implies that the 5-star SEP induced an increase in the 20th percentile of the Part C premium distribution by \$2. $a_c(\tau)$ and $b_t(\tau)$ represent the county and year fixed effects. The error term $\varepsilon_{ct}(\tau)$ includes all the unobserved factors that may affect the τth quantile at the county-year level.

The assumptions required for the validity of this strategy are the same as those of the standard DID framework, in particular the presence of a five star contract in a county after 2011 must be uncorrelated with other unobserved county-year specific shocks ($\varepsilon_{ct}(\tau)$). Our model is a special case of the grouped instrumental variables quantile model of Chetverikov, Larsen and Palmer (2015) and can be estimated using OLS. As explained in Larsen (2015), we can easily estimate this model in two steps: first, we compute the quantile for the contracts characteristic of interest (i.e. Part C premium) for each group (county-year); second, we estimate equation (1) using the computed quantile as a dependent variable in an OLS regression where the units of observation are the groups.[35]

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 a valid control group. In our setting,

---

[35]Compared to standard quantile method, the simplicity of this approach allows us to include a rich control structure, such as county and year fixed effects, while limiting the computational time given the use of OLS. Moreover, standard quantile methods would retrieve a biased $\beta(\tau)$ in presence of a county-year specific shock $\varepsilon_{ct}(\tau)$ (see Chetverikov, Larsen and Palmer (2015)).

counties with highest rated contracts having a rating of 4 or 4.5 stars are an appropriate control group. Clearly, both treatment and control counties have similar quality plans at the top of their respective menu of offerings. As discussed above, this is relevant to ensure that insurers in both sets of counties face similar financial incentives, thus allowing us to identify the effect of the 5-star SEP policy reform separately from any other effect produced by the simultaneous payment reform. The geographical location of the two sets of counties is also similar: Figure 1, shows the geographic distribution of treated (dark red) and control (light red) counties. Nevertheless, treatment and control groups differ along several observable characteristics, like size of the enrollment base and features of the enrollment pool. Indeed, the fact that the 5-star plans are scattered across many different counties does not ensure that their assignment to counties is random. We have two arguments to address this concern, the first is that, for the three reasons explained earlier, 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 analyzed. 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 in terms of observable characteristics. 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 light red in Figure 1.

B. Baseline Results

The plots of Figure 3 summarize our findings for each of the plan characteristics analyzed. Plot (a), for instance, reports the effect of the policy change on the Part C premium. The plot contains a great deal of information: the solid, dark line is drawn using the regression coefficients, $\beta(\tau)$, estimated separately for each one of the quantiles ($\tau = 0.05, 0.1, ..., 0.9, 0.95$) of the Part C premium distribution. The two solid lines around it show the 95 percent confidence interval. This plot reveals that the policy change is associated with a premium increase at the lower end of premiums (up until the third decile) and with a premium decrease in the top end of the premiums (starting from the seventh decile). The decline is about \$250

19

for plans at the 90th percentile of the distribution and it his highly statistically significant. To illustrate the usefulness of a distributional analysis, the plots also report the average effect. The dark, horizontal, dashed line shows the mean effect (with the associated surrounding lines denoting the 95 percent confidence interval) that is estimated by applying a conventional DID method, like the one used for the enrollment analysis. For Part C premium, this mean effect is negative but not statistically significant. The mean effect is clearly unable to reveal the nature of the market readjustment uncovered by the distributional analysis.

Plot (a) also describes where 5-star plans are located within the Part C premium distribution. Small squares and circles are used to mark the share of 5-star plans present at each decile of the distribution (relative to total number of 5-star plans): squares measure the share of 5-star plans in the pre-policy period, while circles measure them in the post-policy period. In terms of the Part C premium distribution, 5-star plans are mostly concentrated in the top 50 percent of the distribution, both pre and post policy. In plot (a), we can thus observe that prior to 2012 about 10 percent of all 5-star plans are located in the top 10 percent of the premium distribution, while after the 5-star SEP none of the 5-star plans is in the top decile. A similar pattern is also observed for the next two deciles. The mass of 5-star plans that leaves the upper portion of the premium distribution reappears in its lower portion. Across all the bottom half of the distribution their increased presence is reveled by the positive gap between the hollow circles and the squares. Combined with the descriptive evidence presented earlier about the sharp decline of 5-star premiums relative to those of plans in different star rating groups of both treatment and control counties (see Figure 2), these results indicate that the distributional effects estimated are produced by a change in 5-star plan premiums, and not merely by changes in the other plans.

Using the same logic to interpret the evidence in the remaining plots, we document a number of interesting results. First, consistent with the behavior of Part C premiums, also for the Part D premium (plot (b)) we observe a slight tendency of premium increases for plans in the medium-low end of the distribution and decreases for plans in the medium-high end of the distribution (where 5-star plans are mostly located). Second, for the Part D deductible, first notice that about 80 percent of all 5-star plans are located in the bottom

decile of the deductible distribution both before and after the policy (the squares and circles in plot (c), respectively). The estimates in plot (c) indicate that low deductible plans (like 5-star plans) reduce their deductible even further, while the deductible increases further for high deductible plans. This evidence, is likely explained by the very peculiar role played by the deductible under the 5-star SEP explained through our simple conceptual framework.[36]

Third and most crucially, plan generosity - as summarized by the MOOP - tends to worsen for plans at the low and medium end of the MOOP distribution, while it remains unchanged for plans in the high end of the MOOP. 5-star plans, that are disproportionately concentrated in the lowest end of the MOOP distribution, seem to respond by reducing their generosity and so do the plans closest to them in terms of MOOP. The following plots, (e)-(f), report additional results in terms of the OOPC. It is particularly interesting to compare the estimates for the Part C OOPC of beneficiaries in poor health and excellent health. For enrollees in poor health, the evidence in Plot (e) is once again of an increase in costs for the plans at the low end of the OOPC distribution and a decline in costs for the high OOPC plans. This is not surprising given the close connection between this OOPC measure and the MOOP. For enrollees in excellent health, however, Plot (f) shows that for all deciles there is no effect. This is reasonable since the representative healthy enrollees' expected out of pocket is not sensitive to changes in benefits that the healthy rarely utilize which are also the most likely candidates for plan changes, if insurers are looking to improve their risk pool.

In interpreting the above results, an important caveat regards the potentially confounding effect created by the pass-through of the 5-star "subsidy." As discussed earlier, starting in 2012 plan payments begun to be linked with their star rating, both for the calculation of the rebate they receive (in case their bid is below the benchmark) and for the bonus applied in the calculation of their benchmark. It is thus possible that part of the changes observed for 5-star plans might be driven by the payment reform. Indeed, the question of whether an higher Part C benchmark is passed through to enrollees in the form of lower premiums

---

[36]Additional evidence, consistent with this interpretation is presented in the web appendix. For the Part D OOPC, the results indicate an improvement of generosity for the plans that, like the 5-star ones, were already low in terms of their Part D OOPC and a worsening of generosity for high Part D OOPC plans. These features involve both the case of poor health beneficiaries and of excellent health beneficiaries. A likely explanation for the different behavior of the Part C and D OOPC measures is based on what we report regarding the Part D deductible.

has been extensively discussed. As series of recent papers exploring this question, Song and Chernew (2013), Curto et al. (2014) and Cabral, Geruso and Mahoney (2014), agree on an estimate of the pass-through of around 50%, but Duggan, Starc and Vabson (forthcoming) report a substantially higher value closer to 100%. Stockley et al. (2014), instead, finds a nearly zero pass-through for premiums, but a significant pass-through in terms of benefits that is rationalized by fact that consumers do not observe premium reductions below the FFS Medicare (i.e., Part B) premium. Although the evidence from these studies is mixed, our results on the premium might in part be due to the benchmark change. Nevertheless, it is interesting to point out that the evidence in Figure 2, panel (a), clearly shows that the Part C premium reduction is substantially more marked for 5-star plans than for 4/4.5-star plans. Since for the years that we study the subsidy created by this reform was nearly identical for 5-star and 4.5-star plans,[37] the sharper decline for 5-star premiums is suggestive of a response to the 5-star SEP. Furthermore, in terms of benefits, observing a decline in generosity for 5-star plans is clearly not the result of the payment reform which, following Stockley et al. (2014), could have instead caused an expansion. The pass-through would also not explain why the OOPC differs for enrollees in different health conditions, nor would it explain the differences between monopoly and duopoly markets that we discuss below.

Finally, the presence of such heterogenous responses between the deductible and the OOPC measures is particularly interesting as it indicates the need, stressed by Glazer and McGuire (2000), to broaden the view of the margins along which insurers compete. Therefore, we conclude by exploring the effect of the 5-star SEP on further margins that insurers could modify. In Figures 4 (a)-(c), we thus report additional estimates of the quantile DID for three soft quality measures: *health care quality*, *customer service* and *drug access*. These are three of the individual measures composing the summary star rating. An interesting result revealed by these estimates is that, while the distribution of premiums and MOOP tend to converge toward the middle, the distribution of various quality measures like health care quality and customer service widens: plans at the higher end of the distribution experience an increase relative to plans at the lower end of the distribution. There is an apparent

---

[37]The new rebate is identical for 5 and 4.5 star contracts, while the bonus is set to 5% for 5-star and 4% for 4.5-4 star plans. See further details in section II.

heterogeneity, however, across the various measures: while for health care quality plans at the high end of the distribution experience a positive and statistically significant effect, for customer service the effect is negative essentially throughout the entire distribution. We should also notice that while we observe a statistically significant increase in heterogeneity across quantiles, the effect is rather small in absolute terms.[38] To help understand the quantile DID results in the plots (a)-(c) of Figure 4, we can look at the evolution overtime of the three soft measures in plots (d)-(f) of Figure 4. Clearly, the individual measures and the overall star rating are likely to be correlated and, therefore, it is fairly straightforward to identify which of the 5 groups is driving our quantile results. For example, in the case *health care quality*, plans with less than 4-star in the treated counties performing worst than similar plans in control counties are driving the results in the lower quantile. Instead, for higher quantiles what is mostly driving the effect are the 5-star plans that in 2012 perform better than their counterparts in the control counties. The case of *health care quality* also confirms how our quantile approach allows us to uncover market-wide effects that a traditional DID analysis focusing only of 5-star plans would not be able to highlight.[39]

C. Markets with 5-Star Contracts Monopoly or Duopoly

As discussed at the beginning of this section, counties where 5-star plans are present have either one or two insurers offering these plans.[40] The distinction between markets with 5-star plan monopoly and duopoly is potentially informative of the interactions between competition and the 5-star SEP reform. Indeed, the reform is such that even enrollees of a 5-star plan can switch plan within the year, provided they move to another 5-star plan. As argued through our discussion of the insurers' incentives, while irrelevant in monopoly markets, this provision can exacerbate the downward pressure on plan generosity in duopoly

---

[38]The small magnitude of these effects could be due to two factors. First, the rating can assume only 5 values, from 1 to 5 stars. Second, overtime the majority of plans achieved 4 stars or more in these three individual measures. This latter feature can also explain the large change in the quantile position of the 5-star plans post 2012 for *health care quality*: given that several plans have 4 or more stars, a drop of one star can generate a large changes in the quantile ranking.

[39]In the web appendix we explore the effect of the 5-star SEP on further margins that insurers could modify and that are related to Part D plan features that we observe. The decline in generosity of 5-star plans is confirmed by two Part D plan characteristics: the share of most frequently used drugs that the plan covers and the number of drugs that the plan covers without placing any utilization restrictions. For both variables, generosity improves for plans in the low end of the distribution, while it declines for plans in the medium-high end (where 5-star plans are located).

[40]We observe 7 counties for which there were more than one 5-star plan in either 2012 or 2013.

markets. Moreover, since the typical 5-star plan in the data enrolls high risk beneficiaries, a 5-star plan must internalize the risk of attracting the high cost enrollees of other 5-star plans, which can alter the incentives and ability to engage in selection.

To evaluate differences in market responses to the policy between monopoly and duopoly markets, we repeat the previous analysis on two subsamples. The six top panels of Figure 5 report the distributional effects for the monopoly case, while the latter six report the effect for the duopoly cases. The comparison of the two environments reveals that, while the decline in premiums is roughly similar, the worsening in generosity for enrollees in poor health is stronger for duopoly than for monopoly markets. Interestingly, for the duopoly case we observe a slight worsening of the OOPC also for individuals in excellent health, suggesting that insurers are constrained in their ability to select based on expected health cost, and in the presence of competition may even dissuade good risk beneficiaries to mitigate potential entry by bad risk beneficiaries.[41]

Altogether, this evidence is suggestive that 5-star plans in duopoly markets decreased their generosity and quality more than 5-star plans in monopoly markets. On the other hand, these reductions are not accompanied by a more pronounced premium decline. Thus, relative to the pre-policy period, the effect of the 5-star SEP appears to have been more beneficial for consumers located in counties with a single firm offering 5-star plans than in areas with competition between 5-star plans. This potentially problematic effect of competition is an interesting manifestation of the complexity of making competition work in healthcare markets. This result complements similar findings by Kuziemko, Meckel and Rossin-Slater (2014) for the related, but different setting of Medicaid managed care where no outside option is present.

D. Robustness Checks

Finally, we discuss two sets of robustness checks. The first one deals with the non-random presence of 5-star plans across counties and it entails using a control group that matches the treatment group on observable characteristics. By comparing demographic characteristics

---

[41]This evidence is further supported by the results involving the risk score reported in the web appendix. Both Part C and D risk scores experience a clear decline for 5-star plans in duopoly markets, but there is no statistically significant decline for the case of monopoly markets.

of treated and control counties collected from the AHRF files of the Health Resources and Services Administration, we find that treated counties tend to have a larger population of Medicare enrollees (and eligibles) and slightly less of both female Medicare enrollees and hospitals accepting Medicare patients (see Table A.2 in the web appendix). Thus, we repeat the analysis a matched DID strategy: we perform the DID analysis on a sample that matches the control counties to the treated ones by using a propensity score method.[42] The results obtained are reported in the web appendix and they show patterns nearly identical to what is reported as our baseline results.

The second set of robustness checks involves the way plan features are aggregated at contract level. Indeed, while we perform our analysis at contract level, certain features, like the Part D deductible are plan-specific and will differ among plans within the same contract. For our baseline estimates presented above, the aggregation method used is an enrollment-weighted average of the plans. As an alternative, we report in the web appendix estimates obtained from using equally-weighted plans which reveal patterns that are broadly in line with the baseline estimates discussed above.[43]

VI   Conclusions

The reform that, starting in 2012, allowed Medicare enrollees to switch at any point in time to the highest quality, 5-star plans could have backfired. By undermining the use of rigid open enrollment periods, a pillar of most insurance markets, this policy could have exacerbated the adverse selection faced by 5-star plans, potentially triggering premium spikes or even plan exit. The fact that this did not happen and that, despite the substantial growth in within-year enrollment in 5-star contracts, their risk pool did not worsen is consistent with theoretical literature starting from Rothschild and Stiglitz (1976) and Glazer and McGuire (2000) that suggests health care plans alter their product seeking to attract good risks.

---

[42]For the 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 counties on the common support of the propensity score between the treatment and the control groups are included. The matched DID estimates reported in Figure A.2 are based on the probit estimates in column 6 of Table A.2 in the web appendix.

[43]For this analysis, we consider only the subset of characteristics varying at plan level.

This paper shows that a relevant force behind these facts is the sophisticated response adopted by suppliers. Both 5-star insurers and their competitors responded to the new policy. The 5-star insurers lowered their premiums, while, at the same time, worsening the amount of coverage offered by their plans. This contributed to expand their enrollment base, without worsening their risk pool. The overall market adjustment entails a compression in the characteristics of the available plans, with greater convergence in terms of both premiums and financial characteristics of the plans.

These results, based on a clean identification strategy, empirically document key features of insurance markets. There are various implications for both research and policy. In terms of research, our findings suggest the relevance of three main avenues for future research. First, when modeling insures behavior it is necessary to consider that competition extends well beyond premium competition and entails subtle aspects of plan design. Second, enrollees inertia in plan choices makes prominent the need to better understand the drivers of plan switching behavior. Third, effective risk adjustment systems need to take into account plan switching behavior associated with the presence of special enrollment periods. The potential enlargement of the set of "life qualifying events" in the ACA exchanges referenced in the introduction might be a fruitful area to further analyze this issue.

Finally, in terms of policy, our results are both encouraging and problematic. On the one hand, the flexibility in product design that insurers retain in Medicare Pact C and D has allowed the 5-star SEP to achieve the goal of bolstering enrollment into 5-star plans. More generally, such flexibility is likely to help making the market sustainable for insurers. On the other hand, however, the very presence of such flexibility implies difficulties in designing rules capable of steering the market toward any public goal. In the context of the 5-star SEP, the reduced generosity of 5-star plans could negatively affect the well being of the weakest beneficiaries and could also represent a diminished allocative efficiency in the market.

References

Abaluck, Jason, and Jonathan Gruber. 2015. "Evolving choice inconsistencies in choice

of prescription drug insurance." *American Economic Review*, forthcoming.

Akerlof, George. 1970. "The Market for Lemons: Quality Uncertainty and the Market Mechanism." *Quarterly Journal of Economics*, 488Ð500.

Azevedo, Eduardo, and Daniel Gottlieb. 2015. "Perfect Competition in Markets with Adverse Selection." mimeo.

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.

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.

Bundorf, M. Kate, Jonathan Levin, and Neale Mahoney. 2012. "Pricing and Welfare in Health Plan Choice." *American Economic Review*, 102(7): 3214–48.

Cabral, M., M. Geruso, and N. Mahoney. 2014. "Does Privatized Medicare Benefit Patients or Producers? Evidence from Medicare Advantage." NBER Working Paper 20470.

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).

Chetverikov, Denis, Bradley Larsen, and Christopher Palmer. 2015. "IV Quantile Regression for Group-level Treatments, with an Application to the Effects of Trade on the Distribution of Wages." *Econometrica*, forthcoming.

Crawford, Gregory S. 2012. "Endogenous product choice: A progress report." *International Journal of Industrial Organization*, 30(3): 315–320.

Crocker, Keith J., and John R. Moran. 2003. "Contracting with Limited Commitment: Evidence from Employment-Based Health Insurance Contract." *RAND Journal of Economics*, 34(4).

Curto, Vilsa, Liran Einav, Jonathan Levin, and Jay Bhattacharya. 2014. "Can Health Insurance Competition Work? Evidence from Medicare Advantage." National Bureau of Economic Research Working Paper 20818.

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.

Darden, Michael, and Ian M McCarthy. 2014. "The Star Treatment: Estimating the Impact of Star Ratings on Medicare Advantage Enrollments." *Journal of Human Resources*, forthcoming.

David Dranove, Daniel Kessler, Mark McClellan, and Mark Satterthwaite. 2003. "Is more information better? The effects of Òreport cardsÓ on health care providers." *Journal of Political Economy*, 111(3): 555–588.

Decarolis, Francesco. 2015. "Medicare Part D: Are Insurers Gaming the Low Income Subsidy Design?" *American Economic Review*, 105(4): 1547–80.

Decarolis, Francesco, Andrea Guglielmo, and Calvin Luscombe. 2016. "Flight to Quality in Medicare Plans: Who Switches to Top Quality Plans and Why." mimeo.

Duggan, Mark, Amanda Starc, and Boris Vabson. forthcoming. "Who Benefits When the Government Pays More? Pass-Through in the Medicare Advantage Program." *Journal of Public Economics*.

Duggan, Mark, Patrick Healy, and Fiona Scott Morton. 2008. "Providing prescription drug coverage to the elderly: America's experiment with Medicare Part D." *The Journal of Economic Perspectives*, 69–92.

Einav, Liran, Amy Finkelstein, and Jonathan Levin. 2010. "Beyond Testing: Empirical Models of Insurance Markets." *Annual Review of Economics*, 2: 311–336.

Einav, Liran, Amy Finkelstein, and Mark R. Cullen. 2010. "Estimating Welfare in Insurance Markets Using Variation in Prices." *The Quarterly Journal of Economics*, 125(3): 877–921.

Einav, Liran, and Amy Finkelstein. 2011. "Selection in Insurance Markets: Theory and Empirics in Pictures." *Journal of Economic Perspectives*, 25(1): 115–138.

Ellis, Randall, and Thomas G. McGuire. 2007. "Predictability and predictiveness in health care spending." *Journal of Health Economics*, 26(1): 25–48.

Fang, Hanming, Michael Keane, and Dan Silverman. 2008. "Sources of Advantageous Selection: Evidence from the Medigap Insurance Market." *Journal of Political Economy*, 116(2): 303–350.

Fan, Ying. 2013. "Ownership consolidation and product characteristics: A study of the US daily newspaper market." *American Economic Review*, 103(5): 1598–1628.

Farinha Luz, Vitor. 2015. "Dynamic Competitive Insurance." mimeo.

Frank, Richard G., Jacob Glazer, and Thomas G. McGuire. 2000. "Measuring adverse selection in managed health care." *Journal of Health Economics*, 19(6): 829–854.

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

Glazer, Jacob, and Thomas G. McGuire. 2006. "Optimal quality reporting in markets for health plans." *Journal of Health Economics*, 25(2): 295–310.

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.

Hendel, Igal, and Alessandro Lizzeri. 2003. "The Role of Commitment in Dynamic Contracts: Evidence from Life Insurance." *The Quarterly Journal of Economics*, 118(1): 299–328.

Ho, Kate, Joseph Hogan, and Fiona Scott Morton. 2014. "The Impact of Consumer Inattention on Insurer Pricing in the Medicare Part D Program." mimeo.

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.

Larsen, Bradley. 2015. "Occupational Licensing and Quality: Distributional and Heterogeneous Effects in the Teaching Profession."

Layton, Timothy J., and Andrew M. Ryan. 2014. "The Effect of the Medicare Advantage Quality Based Payment Demonstration on Quality of Care in Medicare."

Layton, Timothy J., Rndall P. Ellis, and Thomas G. McGuire. 2015. "Assessing Incentives for Adverse Selection in Health Plans Payment Systems." NBER Working Paper 21531.

Lester, Ben, Ali Shourideh, Venky Venkateswaran, and Ariel Zetlin-Jones. 2015. "Screening and Adverse Selection in a Frictional Market." mimeo.

Lu, SF. 2012. "Information disclosure, multitasking and product quality: Evidence from nursing homes." *J. Econom. Management Strategy, v21*, 673–705.

Lustig, Joshua. 2012. "The Welfare Effects of Adverse Selection in Privatized Medicare." mimeo.

Madeira, Tarso. 2015. "The Cost of Removing Deadlines: Evidence from Medicare Part D." mimeo.

Mahoney, Neale, and Glen Weyl. 2014. "Imperfect Competition in Selection Markets." mimeo.

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.

Newhouse, Joseph P, and Thomas G McGuire. 2014. "How Successful Is Medicare Advantage?" *Milbank Quarterly*, 92(2): 351–394.

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.

Reid, Rachel O, Partha Deb, Benjamin L Howell, and William H Shrank. 2013. "Association Between Medicare Advantage Plan Star Ratings and EnrollmentStar Ratings for Medicare Advantage Plan." *Journal of American Medical Association*, 309(3): 267–274.

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.

Song, Z., Landrum M.B., and M.E. Chernew. 2013. "Competitive Bidding in Medicare Advantage: Effect of Benchmark Changes on Plan Bids." *Journal of Health Economics*, 32(6): 1301–1312.

Starc, Amanda. 2014. "Insurer pricing and consumer welfare: evidence from Medigap." *The RAND Journal of Economics*, 45(1): 198–220.

Stockley, K., T.G. McGuire, C. Afendulis, and M.E. Chernew. 2014. "Premium Transparency in the Medicare Advantage Market: Implications for Premiums, Benefits and Efficiency." *NBER Working Paper No. 20208.*

Veiga, Andre, and E. Glen Weyl. 2016. "Product Design in Selection Markets." *Quarterly Journal of Economics*, 126(4).

Werner, Rachel M, R Tamara Konetzka, Elizabeth A Stuart, Edward C Norton, Daniel Polsky, and Jeongyoung Park. 2009. "Impact of public reporting on quality of postacute care." *Health services research*, 44(4): 1169–1187.

Wollman, Thomas. 2014. "Trucks without bailouts: Equilibrium product characteristics for commercial vehicles." Mimeo, Harvard Unicersity.

Table 1: Descriptive Statistics for Part C

	2009-2011									
	Control Group				Treatment Group					
	Less 4 Star		4&4.5 Star		Less 4 Star		4&4.5 Star		5 Star	
	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N
Part C Premium	356.0	25,569	471.3	5,672	373.5	3,373	450.8	665	757.1	421
Part D Total Premium	235.7	25,588	329.4	5,672	223.2	3,375	206.7	665	235.3	421
Maximum OOP Part C	4,525	4,021	3,863	1,968	4,253	402	3,820	227	2,771	150
Part D Deductible	69.95	25,588	47.46	5,672	71.73	3,375	30.73	665	20.74	421
Part C OOPC - Excellent Health	937.3	24,664	835.4	5,260	926.6	3,255	788.5	638	797.4	421
Part C OOPC - Poor Health	2,234	24,664	1,793	5,260	2,168	3,255	1,775	638	1,623	421
Drug Access	3.199	17,461	4.141	5,341	3.159	2,251	4.157	644	4.953	407
Customer Service	2.526	14,556	3.767	4,150	2.365	2,007	3.470	498	4.700	407
Health Care Quality	3.102	18,411	4.038	5,345	3.057	2,318	4.019	644	4.749	407

	2012-2013									
	Control Group				Treatment Group					
	Less 4 Star		4&4.5 Star		Less 4 Star		4&4.5 Star		5 Star	
	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N
Part C Premium	358.7	10,325	424.0	4,811	296.6	1,022	425.3	464	631.4	270
Part D Total Premium	254.6	10,344	305.0	4,811	249.8	1,026	232.1	464	215.4	270
Maximum OOP Part C	4,864	8,318	3,739	4,521	4,494	851	3,798	437	3,377	270
Part D Deductible	88.52	10,344	45.58	4,811	113.5	1,026	35.35	464	30.61	270
Part C OOPC - Excellent Health	1,038	9,656	982.9	4,531	979.1	933	949.8	447	988.8	270
Part C OOPC - Poor Health	2,649	9,656	2,234	4,531	2,443	933	2,186	447	2,174	270
Drug Access	3.257	7,967	3.912	4,738	2.957	792	3.848	460	4.656	270
Customer Service	3.322	8,339	3.926	4,685	2.925	832	3.445	456	4.296	270
Health Care Quality	3.729	8,330	4.250	4,734	3.566	837	4.222	459	4.815	270

Notes: the unit of observation is Contract/County/Year. There are 160 treated counties (i.e., counties with at leas one 5-star plan in 2012 or 2013) and 1,084 control counties (i.e., counties with no 5-star plans, but at least one 4 or 4.5-star plan in 2012 or 2013). The top panel includes observations from 2009 to 2011. The bottom panel includes observations from 2012 to 2013. The "Treatment" sample includes observation from contract with 5 Star rating in either 2012 or 2013. The "Control" sample include contracts with either 4 star in either 2012 or 2013 in counties without 5-star contracts. "Premium Part C" is the annual Premium for Part C. "Premium Part D" is the annual Premium for Part D. "Maximum OOP Part C" or MOOP is the maximum outside of pocket expenditure for in network service, excluding Part D drugs (we observe it starting from 2011). "Part D Deductible" is the maximum annual amount of initial out of pocket expenses for Part D drugs."Part C OOPC Excellent (Poor)" is the average yearly out-of-pocket for individuals with Excellent (Poor) heath status for Part C coverage. "Health Care Quality" is a star rating (1-5), over member's evaluation of health care quality (CAHPS Survey). "Customer Service" is a star rating (1-5), over ability of the health plan to provide information or help when members need it (CAHPS Survey). "Drug Access" is a star rating (1-5) over the ease of getting prescriptions filled when using the plan (CAHPS Survey). "Health Care Quality", "Customer Service" and "Drug Access" are measured at contract level. "Premium Part C", "Premium Part D", "Maximum OOP Part C"', "Deductible Part D", "Part C OOPC Excellent (Poor)" and "Drug OOPC Excellent (Poor)" are measured at plan level and aggregated at contract level as weighted average, with enrollment as weights. Plan with less than 10 enrollees are imputed 5 enrollees.

'

Table 2: Domain Measures for Part C and D - 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 Part C and D summary ratings in 2012. There are 5 domain measures for part C and 4 for Part D. The numbers in the table that follow the description of each domain measure indicate the number of underlying individual measures.

Table 3: Number of 5-star Contracts by Insurer and Year

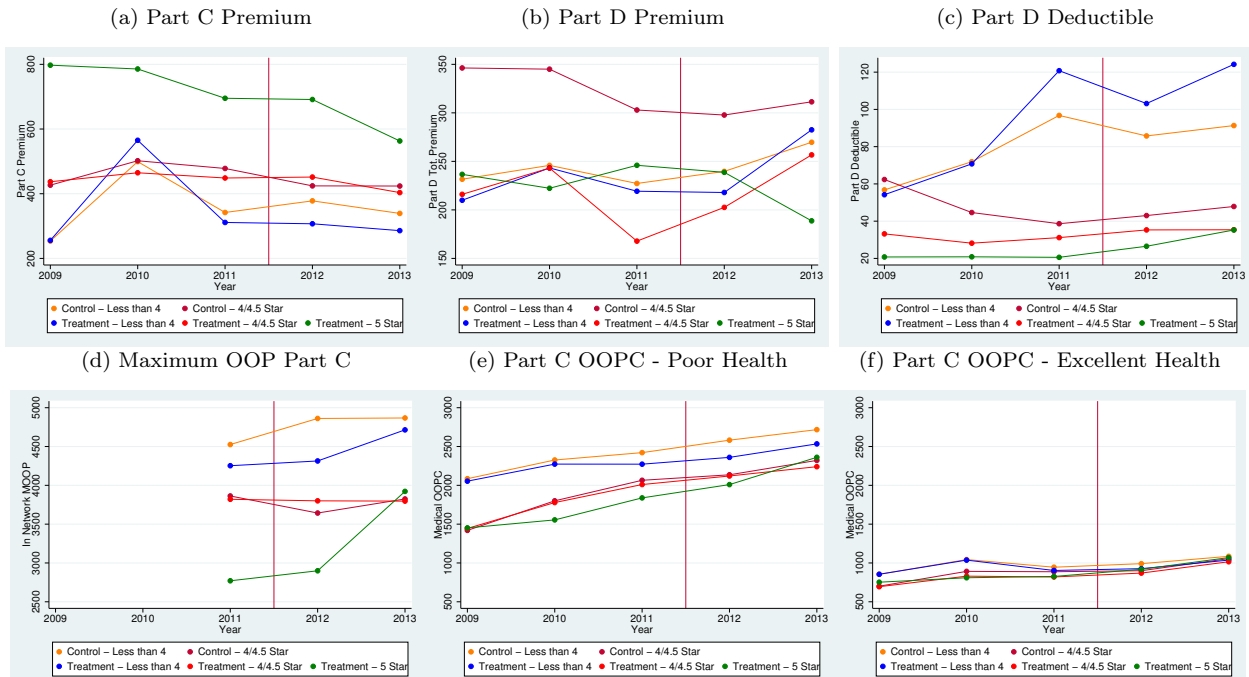| Year	2009	2010	2011	2012	2013
Baystate Health, Inc.	3	3	3	3	3
Group Health Cooperative	13	13	13	13	13
Gundersen Lutheran Health System Inc.	11	11	11	16	16
Humana Inc.	0	0	11	30	30
Kaiser Foundation Health Plan, Inc.	63	63	64	64	64
Marshfield Clinic.	32	32	32	32	36
Martin's Point Health Care, Inc.	12	15	16	16	18
Total	134	137	150	174	180

Notes: For each of the seven insurers offering at least one 5-star contract in 2012-2013 (i.e., treated contracts), the table reports the total number of counties where these treated contracts where offered in all the years from 2009 to 2013.

Figure 1: Maps of 5-Star Counties
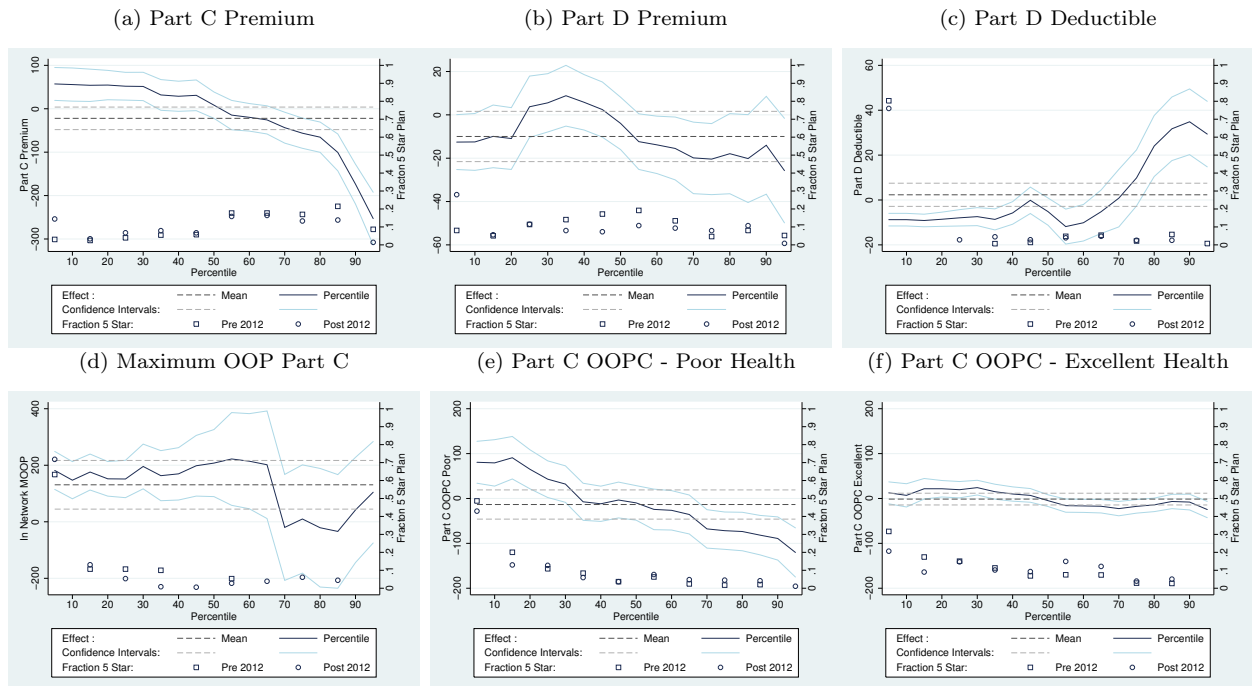

- 5 Star
- 4–4.5 Star
- < 4 Star
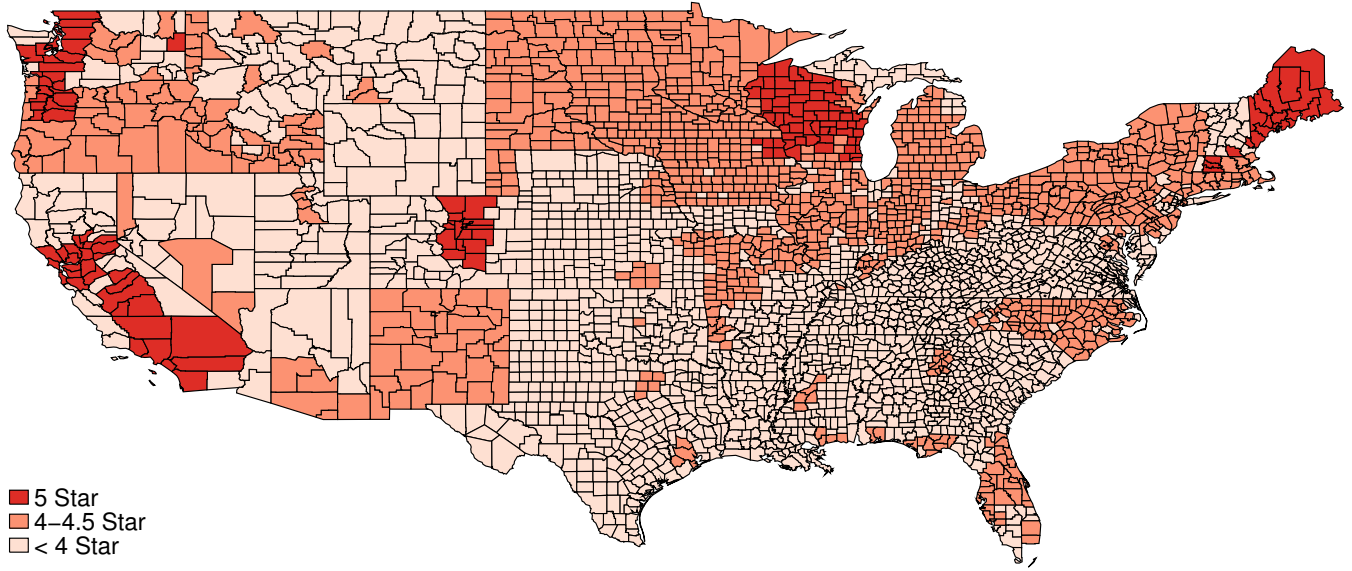
Notes: The heat map reports with the darkest color the set of counties where at least one 5-star plan was offered in 2012 or 2013. The lightest color counties are those where in the same period no plan got a score of 4 or higher. The remaining counties have at least one plan with a score of al least 4, but no plan with a score of 5.

Figure 2: Plan Characteristics Overtime

(a) Part C Premium

(b) Part D Premium

(c) Part D Deductible

(d) Maximum OOP Part C

(e) Part C OOPC - Poor Health

(f) Part C OOPC - Excellent Health

Notes: The solid lines of different colors represent the average value of the outcome variable across contracts with different star ratings (5 star; 4 and 4.5 star; less than 4 star) and for different set of counties (treated and control).

35

Figure 3: Quantile Regression Estimates for Plan Characteristics

(a) Part C Premium

(b) Part D Premium

(c) Part D Deductible

(d) Maximum OOP Part C

(e) Part C OOPC - Poor Health

(f) Part C OOPC - Excellent Health

Notes: The solid, dark line is drawn using the coefficient estimated separately for each one of the quantiles ($\tau = 0.05, 0.1, ..., 0.9$, 0.95). The two slid lines around it show the 95% confidence interval. The dark, horizontal, dashed line shows the mean effect, the lighter lines denotes the 95% confidence interval. Squares measure the share of 5-star plans in the pre-policy period, while circles measure them in the post-policy period.

Figure 4: Quantile Regression Estimates for Plan Characteristics - Other Characteristics

(a) Health Care Quality

(b) Customer Service

(c) Drug Access

(d) Health Care Quality

(e) Customer Service

(f) Drug Access

Notes: For plot (a)-(c) the structure of the figure is analogous to what described for Figure 3, only the outcomes differ as here we consider thee components of the star rating system: Health Care Quality, Customer Service and Drug Access. Similarly, for plot (d)-(f) the structure of the figure is analogous to what described for Figure 2.
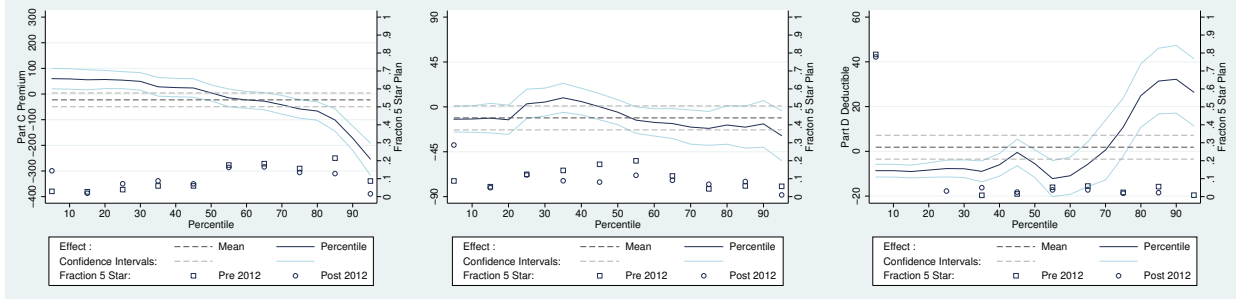
Figure 5: Quantile Regression Estimates - Monopoly and Duopoly Counties

Part I: 5-Star Monopoly Counties

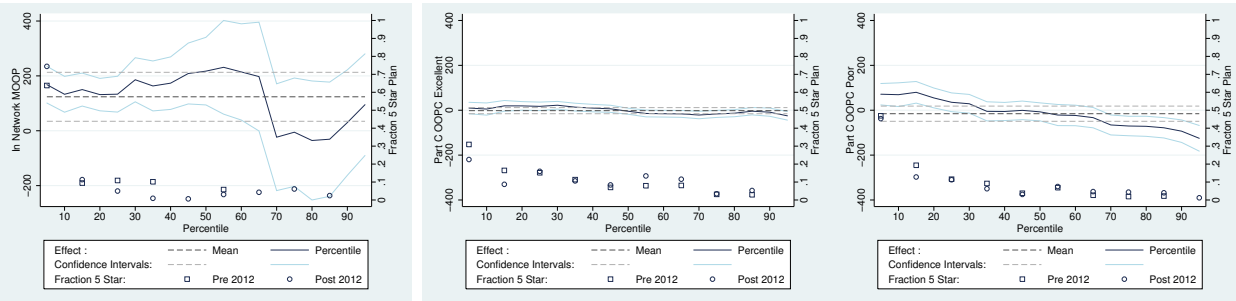(a) Part C Premium   (b) Part D Premium   (c) Part D Deductible

(d) Maximum OOP Part C   (e) Part C OOPC - Excellent Health   (f) Part C OOPC - Poor Health
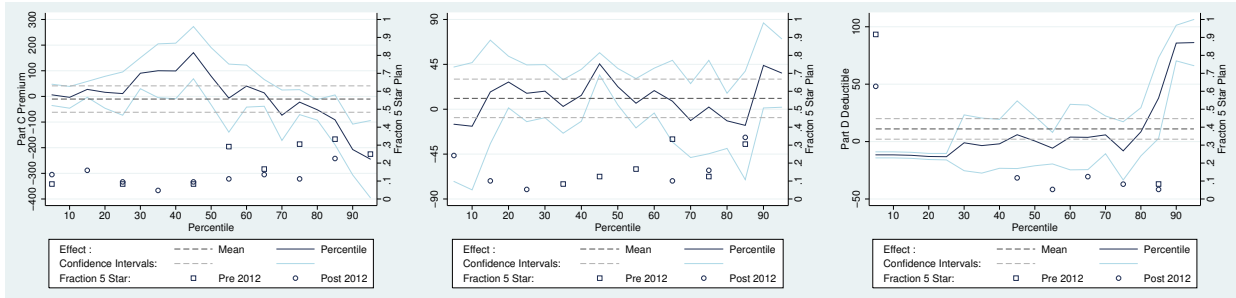
Part II: 5-Star Duopoly Counties

(g) Part C Premium   (h) Part D Premium   (i) Part D Deductible
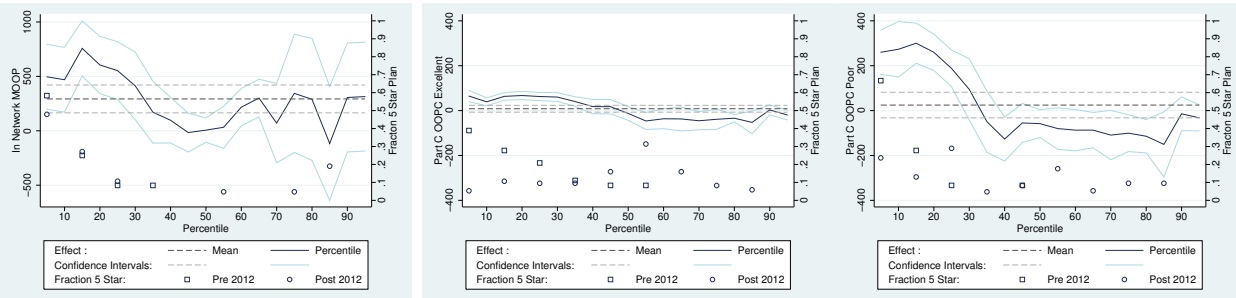
(j) Maximum OOP Part C   (k) Part C OOPC - Excellent Health   (l) Part C OOPC - Poor Health

Notes: The structure of the table is analogous to that of Table 3. The top panel includes as treated counties only those with one insurer offering all 5-star plans. Bottom panel includes as treated counties only those with two insurers offering all 5-star plans.

For Publication on the Authors' Web Page

Insurers' Response to Selection Risk:

Evidence from Medicare Enrollment Reforms

Web Appendix

A. Data and Institutions

The dataset was assembled from data made publicly available by CMS (Center for Medicare and Medicaid Services). 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`.

The calculation of the star rating described in the main text is illustrated below in greater details for the case of the Part D rating for year 2012. A weighted average of the scores earned on each of the individual measures determines the final score.

Table A.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 to each of 17 the individual measures in the calculation of the corresponding domain measures. The 4 domain measures are equally weighted in the calculation of the summary rating.

C. Matched Sample: Probit Estimates

We report in Table A.2 the probit estimates used for the construction of the matched DID estimates. 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 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.

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

D. Evidence on Additional Outcomes: Baseline Estimates and Robustness Checks

This section presents three sets of additional estimates that complement those reported in the quantile estimates in the main text.

- In Figure A.1, we complement the estimates in Figure 3 in the main text by showing the baseline quantile estimates for four additional outcome variables: the OOPC for drugs for enrollees in poor health, panel (a); the OOPC for drugs for enrollees in excellent health, panel (b); the risk score relative to Part C, panel (c); and, finally, the risk score relative to Part D, panel (d).

- In Figure A.2, we report the estimates the robustness check analysis using a control group that matches the treatment group on observable characteristics.

- In Figure A.3, we report the estimates the robustness check analysis using using equally-weighted plans, instead of enrollment-weighted plans as in the analysis in the main text.

Figure A.1: Baseline Estimates: Additional Contract Characteristics

(a) Drug OOPC - Poor Health

(b) Drug OOPC - Excellent Health

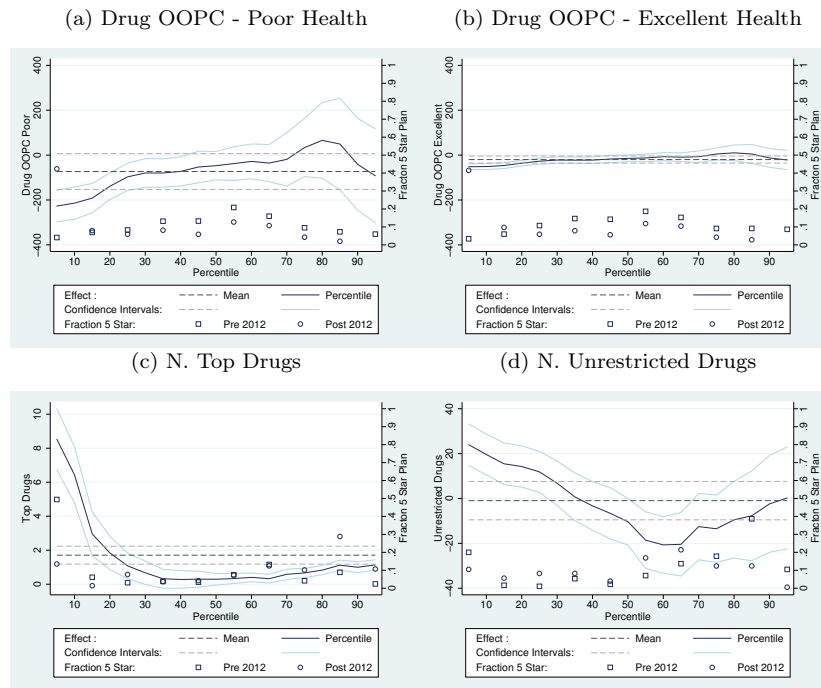(c) N. Top Drugs

(d) N. Unrestricted Drugs

Figure A.2: Quantile Regression Estimates for Plan Characteristics - Matched Samples
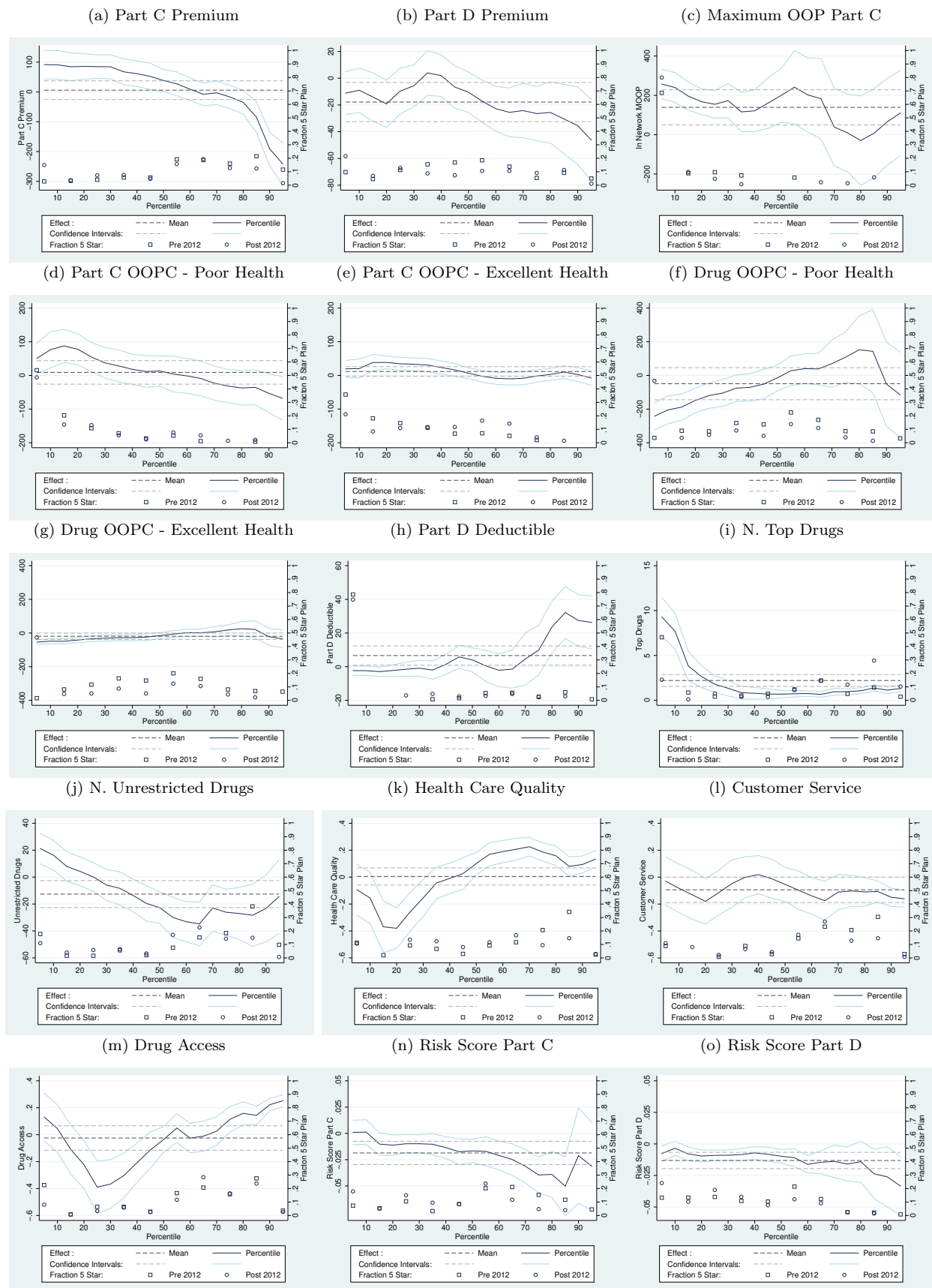
Figure A.3: Quantile Regression Estimates for Plan Characteristics - Mean Characteristics

(a) Part C Premium   (b) Part D Premium   (c) Maximum OOP Part C

(d) Part C OOPC - Poor Health   (e) Part C OOPC - Excellent Health   (f) Drug OOPC - Poor Health

(g) Drug OOPC - Excellent Health   (h) Part D Deductible   (i) N. Top Drugs

(j) N. Unrestricted Drugs   (k) Risk Score Part C   (l) Risk Score Part D