THE WELFARE EFFECTS OF SUPPLY-SIDE REGULATIONS IN MEDICARE
PART D

Francesco Decarolis
Maria Polyakova
Stephen P. Ryan

The Welfare Effects of Supply-Side Regulations in Medicare Part D
Francesco Decarolis, Maria Polyakova, and Stephen P. Ryan
NBER Working Paper No. 21298
June 2015
JEL No. H2,H4,I11,I18,L1,L2

ABSTRACT

The efficiency of publicly-subsidized, privately-provisioned social insurance programs depends on the interaction between insurer behavior and public subsidies. We study this interaction within Medicare Part D Prescription Drug Plan (PDP) markets. Using a structural model of supply and demand, we find: consumers purchase too few and too socially-costly PDP plans; insurers price near marginal cost; the primary driver of welfare is the opportunity cost of government spending on other Medicare programs; and the current subsidization policy achieves a level of total welfare close to that obtained under an optimal in-kind subsidy, but is far from the social planner's first-best solution.

Francesco Decarolis
Department of Economics
Boston University
270 Bay State Rd.
Boston, MA 02215
fdc@bu.edu

Maria Polyakova
Stanford School of Medicine
Redwood Building T111
150 Governor's Lane
Stanford, CA 94305
and NBER
maria.polyakova@stanford.edu

Stephen P. Ryan
Department of Economics
University of Texas at Austin
BRB 3.134D
2225 Speedway Stop C3100
Austin, TX 78712-1690
and NBER
sryan@utexas.edu

1   Introduction

Subsidizing the private provision of health insurance constitutes a large and growing fraction of government's expenditures in health care. Recently, the design of complex subsidization policies within the Affordable Care Act has spurred extensive political debate and become the centerpiece of disagreement around the Act. Potentially even more important, however, is the ongoing shift from the traditional fee-for-service model to subsidized private provision of Medicare coverage, as the budgetary outlays on the Medicare program are roughly five times the projected annual spending on the Exchanges, and amounted to more than $500 billion in 2014.[1] This development substantially changes how the government spends money in health care, pivoting away from direct reimbursement of physician and hospital services toward the subsidization of private insurance plans. Despite the importance of subsidy spending in budgetary outlays, very little is known about the efficiency and the distributional effects of the existing subsidization policies.

In this paper we use the institutional environment of Medicare Part D, a privately provided, publicly-subsidized insurance program for prescription drugs, to derive lessons about the efficiency implications of different subsidy designs. Part D is an important, controversial, and expensive program, with federal spending totaling more than $76 billion annually. We focus our study on the mechanism determining subsidies for Prescription Drug Plans (PDP) in Medicare Part D, particularly on strategic behavior by insurers. Despite the importance of supply-side incentives in this setting, the academic and policy debate has so far mostly focused on individuals' choices in the Part D program, leaving aside insurers' pricing incentives. To start closing this gap, we explore the allocative efficiency implications of the mechanism that Medicare uses to determine subsidies in Medicare Part D. This mechanism relies heavily on insurers' behavior, deriving government subsidies from prices set by insurers. Since subsidies account for the majority of insurers' revenues, the endogeneity of subsidies to insurers' pricing raises concerns about potential efficiency distortions on the supply-side.

Our research strategy starts with the estimation of demand for prescription drug plans. In each market, firms offer a list of insurance plans which vary across several dimensions such as the size of the deductible, the set of drugs that are covered, whether the plan has a "donut-hole," which is a region of expenditures for which the plan reverts to 100 percent co-insurance, and the plan's premium. Demand in Part D is slightly more complicated than

---

[1]Source: Congressional Budget Office, 2014 Medicare Baseline. In terms of federal spending, out of total federal outlays of $3.5 trillion in 2013, the net federal outlays for Medicare amount to 14 percent or $492 billion.

the typical setting due to the presence of two groups of consumers: so-called regular enrollees and low-income (LIS) enrollees.[2] Regular enrollees make unrestricted choices from all plans offered in their region and pay a partially-subsidized premium. In contrast, low-income enrollees, who constitute 35 percent of all enrollees, are randomly assigned to eligible plans by the Centers for Medicare and Medicaid Services (CMS) and pay nothing. These enrollees, however, can and do opt out of the random assignment process and freely choose any plan at additional cost.[3] Using four years of data on the characteristics and enrollments of all Part D PDP plans across all 34 Medicare Part D in the US, we estimate demand for both regular and LIS enrollees using the random coefficients discrete choice framework pioneered by Berry, Levinsohn, and Pakes (1995). Our use of a random coefficients framework allows us to address heterogeneity across consumers, which is a motivation for introducing choice into these types of programs.

Given demand estimates for insurance plans, we then turn our attention to modeling the behavior of firms. A critical piece of this puzzle is the rule for how a firm's pricing decision, hereafter referred to as its bid, are turned into premiums that enrollees face. Medicare beneficiaries do not face full prices or bids set by insurers; instead, there is an intermediate process by which CMS decides on how much of the insurer's bid will be paid by the government in subsidies, and how much will be paid by enrollees in premiums. In this process, CMS takes the sum of all bids for all participating insurers in the US, combines them with prices from the Medicare Advantage market, averages them using enrollment weights from the previous year, and takes a fraction of the resulting number to obtain the base subsidy. The premium of a given plan is then determined by taking the maximum of zero and the firm's bid minus this base subsidy. This pricing mechanism has three effects on market outcomes. First, consumers face premiums that are strictly lower than firm bids, which increases demand. Second, the relative premiums of plans are distorted by this mechanism; this is important since it distorts the choices behavior of consumers across plans. Third, the same bids determine the plans' eligibility to enroll the randomly-assigned LIS enrollees. Only plans with a premium below the average premium in their market are eligible for random assignment of LIS enrollees. Consequently, there is a key linkage between the two groups: the bidding process by which plans qualify to be eligible for low-income assignments also influences premiums for regular enrollees. Thus, these incentives distort both the public payments for

---

[2]Center for Medicare and Medicaid Services (2009) explains how LIS eligibility is determined. In general, consumers below some multiple of the federal poverty level, depending on family size, are eligible for LIS classification.

[3]As of 2011, about one-third of LIS enrollees had opted out of the random assignment system.

low-income enrollees and the prices and choices of regular enrollees. We model firm behavior in light of these incentives and recover estimates of plan-level marginal costs.

With demand and supply cost estimates in hand, we then characterize the welfare effects of the current subsidy mechanism. Our welfare estimates depend on the estimated consumer surplus, producer profits, and the social cost of government spending. We assume that the deadweight loss of taxation is given by 30 cents per dollar of revenue raised (Hausman and Poterba, 1987). We also make two critical assumptions in computing welfare. First, we assume that the rest of the world does not change as we modify the subsidy mechanism in Part D PDP. As such, all of our counterfactual results are subject to the usual partial-equilibrium critiques. Second, all of our estimates, demand, marginal cost, and government spending, are measured relative to their opportunity cost. Consumers in this market are not left without coverage if the Part D PDP market were to shut down; one can readily see this as the inside share of consumers in Part D PDP is only 37.5 percent in 2012. The remaining 62.5 percent are primarily covered by private insurance or a similar insurance program offered under Medicare Advantage (MA-PD). Evidence from consumer level data indicates that out of the few enrollees that switch out of the PDP plans, two-thirds move to an MA-PD plan. Producers face a direct marginal cost of providing the good here, but also the opportunity cost of potentially serving the same consumer in the MA-PD market. Indeed, about 90 percent of PDP plans are offered by insurers that also offer an MA-PD plan. The government spending opportunity cost is particularly salient, as we conservatively assume that consumers would substitute from Part D PDP plans exclusively to MA-PD plans rather than dropping a publicly subsidized program altogether. This implies that all of our estimates—demand, marginal cost, government spending, and, thus, social welfare— are relative to the outside option.

We first calculate welfare estimates for the observed prices and allocations. Our findings suggest that relative to the existing outside option, the current levels of subsidies in the stand-alone Prescription Drug Plans are generating negative nominal welfare with a return of only 17 cents of surplus for every dollar of government spending. However, once the foregone costs of providing similar services in MA-PD are considered, the program generates substantial surplus, with a return of $1.12 per dollar of opportunity cost. This is one of our primary findings; the positive welfare effect of Part D PDP is driven almost exclusively by the opportunity cost of government spending. On its own merits, the total cost of providing subsidized goods exceeds their benefits; expenditures of $9.4 billion generated $2.5 billion of consumer surplus and $459 million of producer profit. However, we estimate that foregone

3

costs of providing similar coverage in MA-PD is $8.3 billion. Considering the opportunity cost and the deadweight loss of taxation to raise government funds, we estimate that the program in its current form generates $1.5 billion in surplus.

To understand the role of the various pieces of the subsidy mechanism, and to assess its efficiency relative to several alternative approaches, we perform several counterfactual experiments. We start with two counterfactuals that examine regulatory links between parts of the PDP subsidy mechanism. Recognizing potential problems arising from mixing together the regular enrollees and the LIS enrollees, several policy initiatives have proposed removing the LIS enrollees to their own market. We simulate this policy and we find that consumer surplus and producer profit increase relative to the observed mechanism, but overall surplus declines as the net surplus generated by the marginal consumers is exceeded by the social cost of subsidizing the program. This finding is further exacerbated if we were to remove the MA-PD plans from the weighted average used to compute the baseline PDP subsidy. The message from these two counterfactuals is that these links help reduce the generosity of the subsidy, which in turn leads to improvements in welfare. We take the latter counterfactual as our baseline for further investigation of the properties of the mechanism, as it removes strategic incentives for competing for LIS enrollees and possible feedback effects in the MA market.

One of the primary motivations for managed competition is that firms will compete for consumers, driving down prices. To assess the competitiveness of the market, we perform two counterfactuals where we change the ownership structure. In the first, we assume that each plan is its own firm; in the second, we assume that every plan in each market belongs to one firm. Compared to the baseline counterfactual, we find the expected pattern that profits increase greatly and consumer surplus declines under the monopolistic regime, with the opposite pattern under atomistic competition. Interestingly, total surplus declines slightly under atomistic competition but actually increases under the monopolist. The monopolist increases prices, but this generates welfare increases as the marginal consumers who exit the market had valuations for PDP plans that were below their social marginal cost.

This highlights a general tension in this setting: the social planner must balance the benefits of additional consumer surplus and producer profits against the social cost of subsidizing the provision of those goods. To formalize this, we perform several counterfactuals where the government sets prices directly. In the first, prices are set at private marginal cost. In the second, the government acts as the social planner, maximizing total welfare.

Under marginal cost pricing, consumer surplus is half of the current mechanism, driven

by a more than doubling of consumer premiums and a corresponding precipitous decline in the amount of consumers choosing to buy a Part D PDP plan. This is not a completely unexpected result; on the one hand, prescription drug coverage in general is certainly a valuable product for seniors. For example, Town and Liu (2003) conclude in their estimates of welfare effects from the introduction of Medicare Advantage program that the prescription drug insurance part of the program was extremely valuable for the Medicare population. At the same time, Engelhardt and Gruber (2011) find evidence of substantial crowd-out, where Part D insurance was used merely as a substitute for other prescription drug coverage sources. Given the outside option, we may have expected to see a large substitution to the outside good if consumers faced the marginal cost.

Marginal cost pricing mechanism, which is usually the benchmark for welfare maximization, produces terrible outcomes by ignoring an important component of welfare: the opportunity cost of government spending. To assess that situation, we compute the social planner's problem. As expected, the social planner has high total surplus of $3.6 billion, or approximately twice the current mechanism. Enrollment in Part D PDP under the social planner is nearly 50 percent of the market. Consumer surplus is slightly higher than the observed mechanism, but the distribution of equilibrium prices is completely different. The social planner highlights and solves two inefficiencies of the current mechanism: average prices are lowered to induce more people to consume in PDP, but the relative prices are also changed to induce more consumers to choose less-costly plans.

With these benchmarks in mind, we then proceed to investigate a menu of counterfactual subsidy-setting policies that CMS could implement in lieu of the current bid averaging process. The simplest scenario would be to provide fixed vouchers that could be used to buy a plan in the Part D market. We find that the current system operates like a voucher, in that the average bid mechanism is set by bids of all plans, and any individual firm has little influence on that average. Unsurprisingly, we can replicate the observed surplus very closely using a fixed voucher. Interestingly, the current system gives welfare very close to the optimal uniform voucher. Bridging the gap between a uniform voucher at the national level and the social planner's plan-specific prices, we also evaluate the welfare gains of instituting vouchers that vary at the regional level, but find that the welfare increase is very minor.

A second option would be to use a uniform proportional discount on all plans' bids. We find that proportional subsidies are, in general, a disastrous idea as firms simply scale their bids in proportion to the subsidy. Consumers face increasingly low premiums, firms are paid increasingly large bids, and government expenditures explode. That combination results in

large negative welfare losses.

Our paper is related to a large theoretical literature that has examined the role and motivation for in-kind subsidies in different sectors of the economy; surprisingly, however, the empirical analysis of the motivation and effects of such government policies is much less explored (Currie and Gahvari, 2008). In health insurance, the literature has focused on the effects of tax subsidies to employer-provided health insurance (Gruber and Washington, 2005). At the same time, the recent expansion of federal health insurance programs into private markets has brought a large public policy interest to how the federal budget subsidizes these programs—from privatized Medicare and Medicaid plans to the ACA health insurance exchanges. For example, Enthoven (2011) and Frakt (2011) discuss some of the key conceptual points and the policy debate. Conceptually and methodologically, our paper is closest to Curto et al. (2015) that explores the questions about subsidies, competition, and market design in the context of Medicare Advantage.

The current paper is also related to the growing literature that analyzes the Medicare Part D program as a prominent example of a health insurance program with consumer choice. This literature has so far mostly focused on demand questions. Several papers have explored the rationality of individual choices (Heiss et al., 2010, 2013; Abaluck and Gruber, 2011, 2013; Ketcham et al., 2012; Kesternich et al., 2013; Kling et al., 2012; Vetter et al., 2013; Winter et al., 2006; Ketcham et al., 2015). Relatedly, Ericson (2014); Miller and Yeo (2014b); Abaluck and Gruber (2013); Ho et al. (2015); Polyakova (2015) explore the presence and role of inertia in the individual choices of Part D contracts. Einav et al. (2015) study the effect of non-linear contract structure on the drug consumption decisions in Part D. Related work has explicitly considered the dynamic incentives within Part D contracts (Abaluck et al., 2015; Dalton et al., 2015). A number of papers, in economics and health services research, have examined the effect of Part D on drug utilization, adherence, and health outcomes for the elderly, for example Ketcham and Simon (2008).

Further, this paper is related to a substantial theoretical and empirical literature on the supply-side effects of government regulation. Laffont and Tirole (1993) gives a classic reference on the multitude of theoretical issues. Our research question is related to the issues of government procurement in health care, such as Duggan (2004) and Duggan and Scott Morton (2006). The literature on the supply side of Part D is still rather small. Ericson (2014) raises the questions of insurer strategies in Part D, arguing that insurers are exploiting individual inertia in their pricing decisions. Ho et al. (2015) expand on this theme, exploring the extent of strategic supply-side pricing in response to consumer inertia.

Duggan et al. (2008); Duggan and Scott Morton (2010) estimate the effect of Part D on drug prices, and Yin and Lakdawalla (2015) analyzes how Part D enrollment affects private insurance markets. Decarolis (2015) focuses entirely on the supply-side, documenting that insurers are pricing strategically to take advantage of low-income-subsidy policies in Part D. Chorniy et al. (2014) explore the issues around Medicare Part D mergers. Miller (2015) and Miller and Yeo (2014a) consider questions of risk adjustment, low-income subsidies, and the effect of providing a public option in Part D, respectively—both of these issues of market design are close in spirit to the questions we explore in the current paper. Also close to the current paper is Lucarelli et al. (2012), which considers the welfare effects of imposing an upper bound on how many plans each insurer can offer in each Part D market, or removing plans that provide coverage in the gap. Their paper also discusses the potential supply-side effects of introducing ex ante competition for entry at market-level rather than price competition among many insurers in the same market. The entry dimension is largely outside the scope of the current paper, as empirically few large plans appear to exercise the entry margin in response to changes in subsidies, but combining the entry margin analysis with the subsidization problem may provide a productive avenue for future research.

The remainder of the paper is organized as follows: Section 2 discusses the key economic concepts. Section 3 describes the institutional details of the Medicare Part D market and our sources of data. Section 4 introduces the theoretical model underpinning our analysis, while Section 5 describes our empirical application of that model to the data and our results. Section 6 discusses our counterfactual pricing mechanisms and presents our results. Section 7 concludes.

2 Conceptual Framework

In imperfectly-competitive markets with differentiated products, such as Medicare Part D, subsidy policies create incentives that affect both consumer and producer behavior. To illustrate these effects, consider the following simple model of subsidized competition with differentiated products. Suppose there are two firms, each selling one product to a unit mass of consumers. The utility to consumer i from product j is given by $u_{ij} = \delta_j - \alpha p_j + \epsilon_{ij}$, where δ_j is a measures of desirability of the products. The outside option gives zero utility. The idiosyncratic taste shock, ϵ_{ij}, is distributed type I extreme value with the usual scale normalization. Consumers purchase the good that gives them the highest utility. Firms maximize profits by setting prices; under the assumption that marginal costs are zero, the

pricing first-order condition for the j-th product is $p_j = 1/(\alpha(1-s_j))$, where s_j is the share of product j: $s_j = \exp(\delta_j - \alpha p_j)/(1 + \sum_{k=1}^{2} \exp(\delta_k - \alpha p_k))$. Let us first consider the case where the government introduces a proportional subsidy that reduces the effective price by $1 - z$; that is, the effective price is $\hat{\alpha} = z \cdot \alpha$. The government raises taxes through distortionary taxes that create a deadweight loss, λ, for each dollar of subsidy provided. For our illustrative examples below, we set $\alpha = 1$, $\lambda = 0.3$, and $\delta = \{1.2, 1\}$, so that the first product is more desirable to consumers; in an insurance context, this would be a plan with more generous coverage.

Figure 1 illustrates the effects of increasing z from zero to one. For consumers, government subsidies distort choices along two margins. First, subsidizing a market makes it more attractive to marginal consumers who will purchase a product only when prices are subsidized; this is shown in the figure as a decline in the outside option as the subsidy becomes more generous. This distorts consumption choices *across* markets, and has important implications for welfare, as it can lead to lower-value consumers purchasing goods with social costs above their valuations. A second, more subtle, margin influences both those marginal and inframarginal consumers if the subsidy changes the relative demand for products within the market. In general, consumers substitute from the now relatively more expensive plans to those that are now relatively cheaper as the subsidy mechanism changes their relative prices. This is shown in the figure as the relative inside shares, computed as the share of product one over the share of product two. Here, consumers substitute to the more generous plan in equilibrium. Hence, subsidies may distort the allocative efficiency *within* the market.

The subsidy also affects the supply side of the market. Firms know that consumers have less elastic demand curves, and will charge higher markups as a result. To measure this effect, we define "excess price" as the percentage of the proportional subsidy which is set as a higher price by firms. For example, if the proportional subsidy is 10 percent, and the effective prices that consumers face decline by only 8 percent, we define the excess price to be $100*(0.10 - 0.08)/0.10$, or 20 percent. The excess price is plotted as a percentage against the right y-axis. The values range from zero when the subsidy is zero and one, to a little over six percent when the subsidy is just under 60 percent. This is the supply-side response to a demand-side subsidy analog of pass-through when a firm faces higher costs.

The form of the subsidy also matters. Figure 2 shows the same figure with a flat subsidy instead of a proportional subsidy. Consumers are given a voucher ranging from zero to one and a half to spend on an inside option if they choose to do so. Increasing the generosity of the voucher decreases the share of the outside option as the equilibrium price that consumers

face falls. While the relative shares of the two options changes, it does so in the opposite direction from the proportional subsidy; the market share of the first good falls as the voucher becomes more generous. We redefine the excess price in this case to be the percentage of the voucher that is not expressed in the decline of the consumer-facing price. For example, if the voucher is equal to 1, but the price only drops by 0.7, the excess price is 30 percent. In our simple model, firms are able to raise prices by approximately 20 percent in response to the voucher; interestingly, the relationship between price increases and the generosity of the voucher is non-monotonic.

This discussion highlights that the government faces a complex, nonlinear set of economic forces when considering both how generously to subsidize markets and also how to provision those subsidies. Figure 3 shows the constituent components of the social planner's welfare maximization problem—consumer surplus, producer profits, government expenditures—along with total surplus. The optimal subsidy occurs in the interior, where the social planner trades off increases in consumer surplus and producer profit against socially-costly government expenditures. The figure also highlights an essential economic rationale for subsidization: market power. In our model, oligopolists set prices above marginal cost. Consistent with the theory of the second best, subsidizing demand leads to increased demand in equilibrium, which increases total welfare from the no subsidy policy. This is also an essential rationale for the idea behind managed competition: instead of having a single supplier, multiple firms competing in the market will help drive prices towards marginal cost. In our simple model, if we impose that firms price at marginal cost, the optimal subsidy goes to zero in both mechanisms. However, competition also comes with a potential drawback: if the subsidy is set too generously, additional competition can actually lead to welfare decreases, as prices may be set below social marginal cost in equilibrium.

To summarize, our simple model highlights the essential points of our empirical investigation that follows. First, increasing subsidies increases consumer surplus but may induce marginal consumers to purchase goods at a social cost higher than their valuations. Second, within the market, subsidies can distort the relative prices of goods, which can lead to allocative distortions. Third, firms can capture some rents from the subsidy by exerting market power and raising prices. Fourth, government expenditures have a social cost due to the deadweight loss of taxation. The social planner's solution balances all of these factors when constructing optimal prices. We note that our model sidesteps an important issue regarding why the government subsidizes markets. In real world contexts, there may be additional reasons for subsidization, such as consumer-side externalities or political consid-

9

erations. Throughout the paper, we do not take a stand on why there are subsidies, but rather take them as given and consider how the present mechanism could be adjusted to obtain more efficient outcomes.

3 Institutional Environment

Medicare is a public health insurance program for the elderly and disabled in the United States that covers over 50 million beneficiaries and costs the government about \$500 billion annually. The program is administered by CMS, and consists of several pieces. Parts A and B cover hospital and outpatient services, respectively, under a fee-for-service model of "traditional" Medicare. Part C, commonly known as Medicare Advantage, was introduced in 1997 and allows consumers to switch from fee-for-service to managed care plans administered by private insurers that are highly subsidized by the government. In 2006, Congress expanded Medicare program to include prescription drug coverage via Medicare Part D. In 2014, approximately 37 million individuals benefited from the Medicare Part D program and the Congressional Budget Office estimates that the government currently spends over \$76 billion on Part D annually. This new part of the Medicare program is the institutional setting of our study.

For beneficiaries in traditional Medicare coverage, who are not eligible for additional low-income assistance, buying a Part D prescription drug insurance plans is voluntary and requires an active enrollment decision. These so-called "regular" enrollees may choose one of about 40 stand-alone PDP contracts offered in their region; the Part D program is divided into 34 geographic markets, some of which follow state boundaries, and some combine the states with smaller populations. Beneficiaries in traditional Medicare that are eligible for additional low-income subsidies, on the other hand, are automatically assigned to eligible plans by CMS; these individuals can subsequently change their random assignment by making an active choice. The latter group is known as "LIS choosers."[4] Once enrolled, regular beneficiaries pay premiums on the order of \$400-\$500 (see Table 1) a year, as well as deductibles, co-payments or co-insurance. LIS-eligible enrollees receive additional support to cover premiums and cost-sharing.

The supply-side of the Part D program has a unique, and controversial, design.[5] Un-

---

[4]If either regular or LIS-eligible beneficiaries choose to enroll in private Medicare Advantage plans rather than traditional fee-for-service Medicare, their Part D coverage will be provided within the MA plans, known as MA-PD. The majority of MA plans include MA-PD coverage.

[5]Oliver, Lee, and Lipton (2004) discusses the political origins of Part D and its mixed reception in the

like the rest of Medicare, the drug insurance benefit is administered exclusively by private insurance companies. At the same time, the setting differs from more conventional private insurance markets in two key ways. First, firms are highly regulated and product selection is restricted; CMS sets an annual Standard Defined Benefit (SBD), which defines the minimum actuarial level of insurance that the private plans are required to provide. The SDB has a non-linear structure illustrated in Figure 4; it includes a deductible, a 25 percent co-insurance rate and the infamous "donut hole," which is a gap in coverage at higher spending levels. As long as an actuarial minimum is satisfied, insurers are allowed to adjust and/or top up the SDB contract design, which generates variation in contracts' financial characteristics. In addition, contracts may be differentiated by the quality of insurer's pharmacy networks, which drugs are covered, and other non-pecuniary quality measures. The second way in which Part D environment differs from more conventional insurance markets is that consumers bear only a fraction of the cost in the program, as 90 percent of insurer revenues come from the government's per capita subsidies.[6] For individuals, who are eligible for low-income-subsidies, these subsidies can go up to 100 percent.

Subsidies are determined through a complex system that depends on firm behavior. First, the government administers an annual "simultaneous bidding" mechanism. According to this mechanism, the insurers that want to participate in the program submit bids for each insurance plan in each region they want to offer. By statute, the bids are supposed to reflect how much revenue the insurer "needs," including a profit margin and fixed cost allowances, to be able to offer the plan to an average risk beneficiary.[7] Medicare takes the bids submitted by insurers for each of their plans and channels them through a function that outputs which part of the bid is paid by consumers in premiums and which part is paid by Medicare as a subsidy. This function takes the bids of all Part D PDP plans nationwide, adds in the prices of plans in MA-PD, weights them by lagged enrollment shares of the plans, and takes the average. Less than 75 percent of this average is set as Medicare's subsidy. The remaining 25 percent of the national bid average together with the difference between the plan's bid and the national average is set as the consumer's premium.[8]

---

first years of the program, particularly among consumers.

[6]See Table IV.B11 of 2012 Trustees of Medicare Annual Report.

[7]There are several nuances buried in the set-up of the bidding procedure that are important for insurers' incentives and will enter the insurers' profit function in our empirical model. First, Medicare sets a minimum required actuarial benefit level that plans have to offer. Plans are allowed to offer more coverage ("enhance" the coverage), but that enhanced portion is not subsidized. Thus, when submitting their bids plans are supposed to only include the costs they expect to incur for the baseline actuarial portion of their benefit. The incremental premium for the enhanced coverage in the plans has to be directly passed on to the consumers.

[8]The per capita subsidy payment from Medicare is further adjusted by the risk score of each enrollee,

A related feature of the subsidy mechanism concerns the role of low income beneficiaries in the Part D program. Medicare utilizes the insurers' bids described above to also determine which insurance plans qualify to enroll randomly assigned LIS beneficiaries. For each geographic market, Medicare calculates the average consumer premium. This average constitutes the subsidy amount that low-income beneficiaries receive, known as LIS benchmark or LIPSA. Plans that have premiums below the LIS benchmark qualify for random assignment of LIS enrollees (Decarolis, 2015).

This somewhat Byzantine system determines prices, and thus allocations, for plans in Part D PDP and, by extension, alternatives such as MA-PD plans. We emphasize three elements of the current system: first, the level of subsidies is endogenous to firm behavior. If all firms increased their bids, the subsidy would be more generous. This generates a complex feedback between firm pricing and the subsidy, as the two are inextricably tied together. In such an environment, it is unclear ex ante whether competition for consumers, which lowers prices, or the softening of demand through the subsidy, which raises prices and the subsidy, will dominate in equilibrium. Second, the mechanism links together the prices of plans from MA-PD to the subsidy; this is important because some regions have very low MA-PD prices. This will lower the generosity of the subsidy in those markets, which may be appropriate given the low price of close substitutes in MA-PD. Third, firms may lower their prices strategically in order to qualify for LIS enrollees; this shading down will also have the effect of lowering the average bid, which decreases the generosity of the subsidy. When considered in full, this complex system may or may not lead to socially-desirable outcomes; the remainder of this paper attempts to take first steps in answering this question by understanding the fundamental forces of supply and demand that express themselves through this mechanism in equilibrium. We next turn to describing our model to tackle these questions.

4 Model

We propose an empirical model of demand and supply of insurance contracts in Medicare Part D that will help us evaluate the market structure and the efficiency of the subsidization mechanism in the program. We start with a model of demand for insurance contracts that follows the approach of Berry (1994) and Berry, Levinsohn, and Pakes (1995) (hereafter referred to as BLP). We then move to a supply-side model that allows us to estimate the

---

while the premium may also include an additional payment for enhanced benefits if the plan offers them.

marginal costs of insurers.

4.1 Demand

We consider two separate demand systems. First, we estimate demand of regular enrollees, who choose their plans, pay full enrollee premiums, and also pay full cost-sharing through deductibles, co-insurance, and co-pays. Second, we estimate a separate demand system for enrollees that are eligible for low income subsidies and face different plan characteristics.

We start with the enrollment decisions of regular beneficiaries. We define the potential market as all Medicare beneficiaries that are not eligible for low income subsidies, and did not receive their Part D coverage through their employer or through special groups like Veteran Affairs. This leaves us with non-LIS Medicare beneficiaries that chose to enroll into a stand-alone prescription drug plan (PDP), or a Medicare Advantage Prescription Drug plan (MA-PD), or did not have any Part D coverage. We let the choice of not enrolling into any part of the Part D program or enrolling through a Medicare Advantage plan comprise the outside option. The utility from this outside option is normalized to zero. Within the inside option, individuals are choosing among 40 to 50 stand-alone Prescription Drug Plans (PDPs) that are available in their region. We estimate demand for plans in the 34 statutory Part D geographic regions in years 2007 to 2010, for a total of 136 well-defined markets.

We posit that individuals select insurance contracts among PDP plans by choosing a combination of pecuniary and non-pecuniary plan characteristics that maximizes their indirect utility. We take the characteristics-space approach and project all plans into the same set of characteristics. This approach allows us to make fewer assumptions about how individuals perceive the financial characteristics of plans, but also implies that we remain agnostic about the objective actuarial efficiency of choices, and also do not recover deeper structural parameters such as risk aversion. Despite the fact that we are estimating demand for insurance and thus preferences may depend on risk aversion, we argue that the model of linear index utility with unobserved heterogeneity is suitable for our goals. The risk protection quality of an insurance plan is represented by financial characteristics other than premiums. We can think about the linear utility index as a reduced-form way of capturing revealed valuation of different financial characteristics of plans that are generated by underlying concave utility functions over the distributions of expected spending. In the simulations of the model in Section 6, we will be interested in capturing the demand response to changes in premiums, while keeping the plans' actuarial properties fixed.

With these modeling choices in mind, we let the utility consist of a deterministic compo-

nent and a random shock, ϵ_{ijt}, distributed as a Type I Extreme Value:

$$u_{ijt} = v_{ijt} + \epsilon_{ijt}. \tag{1}$$

The deterministic indirect utility function of a regular enrollee i who chooses plan j in market t is given by:

$$v_{ijt} = -\alpha_i p_{jt} + \beta_i x_{jt} + \xi_{jt}, \tag{2}$$

where p_{jt} is the plan's enrollee premium. Note that unlike in standard product markets, the premium that enrollees pay in Part D is not equivalent to the per capita revenue that firms receive, since there is a large part paid in federal subsidies to insurers. Allowing for the possibility that the government subsidy, z, can be larger than a particular plan's desired per capita revenue, the premium is then equal to $p_{jt} = \max\{0, b_{jt} - z_{jt}\}$, where b_{jt} is the amount of money the firm receives for the enrollee, or its "bid" in Medicare's terminology.

The observable characteristics of plan j in market t, x_{jt}, includes the annual deductible, a flag for whether the plan has coverage in the donut hole, whether the plan is enhanced, and several generosity measures of drug formularies. We also include fixed effects for parent organizations that capture individuals' preferences for brand names of large insurance companies and insurer-level quality characteristics of plans, such as pharmacy networks, while ξ_{jt} is a plan-specific fixed effect that captures unobserved plan quality.

We also include the vintage of the plan as a approximate proxy for switching costs. The intuitive idea is that the longer the plan has been around, the larger the proportion of its subscribers have been from previous years. If there are switching costs or inattention involved in the re-optimization of an insurance plan choice for some consumers, those consumers will appear to be less price sensitive than those choosing a plan for the first time; this reduced elasticity of demand should translate into higher prices on the part of insurers, all else equal. Since in this paper we are interested in the effects of the subsidy mechanism on pricing, rather than the effects of switching costs, we do not develop a dynamic model, instead pursuing a slightly ad hoc approach of including vintage into the utility function. One way to think about our approach is to assume that any changes in subsidization policies would not result in the reduction of switching costs or consumer inattention, and thus the vintage measure is sufficient to account for reduced demand elasticity and insurers' static re-pricing responses that are conditional on the existing enrollment pool.[9]

---

[9]We note that the coefficient on plan vintage can be interpreted as a structural parameter only under very specific circumstances—for example, if consumers are unaware of their switching costs. A complete characterization of the influence of switching costs on demand and pricing would require an equilibrium

Unobserved consumer heterogeneity enters the model through random coefficients on the premium, coverage in the gap, and overall inside option. The unobserved heterogeneity may capture differences in income, as well as individuals' differences in risk and risk aversion. We choose a log-normal distribution for random coefficients on premiums that is only defined on the positive quadrant and reflects typically log-normally distributed income. The random coefficient on premium is composed of a common component, α, and an individual-level random shock, $\nu \sim \mathcal{N}(0,1)$, which is scaled by σ_α:

$$\ln \alpha_i = \alpha + \sigma_\alpha \nu_i. \tag{3}$$

The parameters governing coverage in the gap coverage, β_{gap} and σ_{gap}, and the inside option, β_{inner} and σ_{inner}, are specified analogously but lack the logarithmic transformation.

Estimating demand for LIS choosers is slightly different due to the institutional setting. While LIS enrollees are randomly assigned to eligible plans by the government by default, they are allowed to change their assignment to a plan of their own choice after the random assignment. We model the demand of low-income beneficiaries closely to what we do for regular enrollees. The key difference is that low-income beneficiaries face different characteristics of plans, as their cost-sharing is largely covered by the government. Let the deterministic indirect utility function of a low-income subsidy enrollee i chooses plan j in market t be given by:

$$v_{ijt} = -\alpha_i^{LIS} p_{jt}^{LIS} + \beta_i^{LIS} x_{jt}^{LIS} + \xi_{jt}^{LIS}, \tag{4}$$

where p_{jt}^{LIS} is the plan's premium for the low-income population. This premium is computed as the remainder of the difference between the insurers' bid and the region-level LIS subsidy (LIPSA), which is higher than the subsidy for regular enrollees. x_{jt}^{LIS} contains observable characteristics of plan j in market t as faced by the low-income population. The difference in the plan characteristics that regular and LIS enrollees face lies primarily in cost-sharing: to the first order, the LIS population does not face a deductible or coverage in the gap or co-payments above certain thresholds, as this cost-sharing part is picked up by the government.

---

model as in Klemperer (1995), Dubé et al. (2009), or Ho et al. (2015). We note that this literature has conflicting predictions about the sign of pricing effects in response to switching costs: Klemperer (1995) concludes that prices are likely to be higher in equilibrium, Dubé et al. (2009) demonstrates that prices can be lower in equilibrium. In the setting of Medicare Advantage plans, Miller (2014) argues that in insurance markets that are characterized by inertial demand, the marginal cost estimates from a static Bertrand model may be around 20 percent higher or lower than the "true" dynamic values. Recognizing this concern in our setting, we report the key counterfactual results in Section 6 for a 20 percent interval around our marginal cost estimates.

The final empirical challenge is that we cannot distinguish LIS enrollees who are in LIS-eligible plans due to random assignment or by choice. Therefore, we make the assumption that all LIS enrollees in LIS-eligible plans are there by random assignment, and thus estimate preferences of the LIS-eligible population from the choices of LIS choosers that enrolled in plans not eligible for random assignment. We define the outside option for LIS enrollees as staying in a randomly assigned plan in Part D PDP.[10]

4.2 Supply

Modeling the supply side in Medicare Part D market presents a considerable challenge, as the decision-making of the insurers is affected by a complex set of regulatory provisions. We start with a description of the flow of payments in Part D and set-up a general profit function that can incorporate these features. We then discuss our strategy of arriving at an empirically tractable version of the supply-side model.

Firms receive revenues across a variety of channels. For each individual that plan j enrolls, the insurer collects an enrollee premium, p_j. This is augmented with an individual-specific subsidy, z_i, from the government, which is composed of the baseline subsidy for the plan, as described above, and an adjustment for the enrollees ex-ante health risk, denoted by r_i. For example, an individual with average risk level will only receive baseline subsidy, while an individual with costly chronic conditions may generate twice the amount of the baseline subsidy in insurers' revenues. This individual-level risk adjustment is intended to ensure that all consumers look equally profitable to firms in order to reduce incentives for risk-based selection. Recalling that the level of the baseline subsidy depends on the average bid, $\bar{b}$, we can write the subsidy as a function of the average bid and individual-specific health risk as $z_i(\bar{b}, r_i)$. For an average-risk beneficiary, the sum of the premium and government subsidy is equal to the bid that the firm submitted for that plan.

On the cost side, the ex-post costs of a plan differ for each enrollee and depend on individual drug expenditures. Some of these costs are mitigated by the government through catastrophic reinsurance provisions, according to which the government directly pays about 80 percent of individual's drug spending for particularly high spenders. Throughout the empirical results we will refer to these reinsurance provisions as "non-premium subsidies" or reinsurance payments. For an individual with a given total annual drug expenditure amount,

---

[10] We have also estimated a version of this model where the outside option additionally includes LIS enrollees in MA-PD plans (about 15% of LIS-eligible beneficiaries are enrolled in MA-PD plans). This results in practically identical parameter estimates, as the primary identifying variation in the model comes from choices of LIS choosers across different PDP plans.

the costs of the plan will also depend on the cost-sharing characteristics of the plan, denoted by ϕ_j. These include characteristics such as the deductible level, co-pays and co-insurance, as well as coverage in the donut hole if any. We let individual-level ex-post costs be the function of these cost-sharing characteristics of a plan as well as the individual's measure of health risk, r_i; that is we let the cost be $c_{ij}(r_i, \phi_j)$.

The final piece of a plan's ex-post profit are risk corridor transfers between insurers and the federal government. These transfers that happen at the end of the year, and restrict the downside (but also upside) risk of enrolling extremely costly individuals for the insurers.[11] We denote the function which adjusts a plan's ex-post profit with Γ.

The ex-post profit for plan j as a function of its bid b_j is then:

$$\pi_j(b_j; b_{-j}) = \Gamma\left[\sum_{i \in j}\left(p_j(\bar{b}, b_j) + z_i(\bar{b}, r_i) - c_{ij}(r_i, \phi_j)\right)\right], \tag{5}$$

where the summation is taken over all individuals enrolling in the plan. For each individual i, the subsidy and the cost can be expressed as an individual-specific deviation from the baseline subsidy and an average plan-specific cost of coverage: $z_i = z + \tilde{z}_i$ and $c_{ij} = c_j + \tilde{c}_{ij}$. Denote the individual-specific difference in the subsidy and cost as $\eta_{ij} = \tilde{c}_{ij} - \tilde{z}_i$. This function allows us to capture adverse or advantageous selection from the point of view of the insurance plan. Given the empirical evidence in Polyakova (2015) on the selection patterns in Medicare Part D, η_{ij} mostly depends on whether or not a plan offers coverage in the gap. We thus let this individual-specific component be a function of plan characteristics: $\eta_{ij}(\phi_j)$. Using this notation, we can re-write the profit function as:

$$\pi_j(b_j; b_{-j}) = \Gamma\left[\sum_{i \in j}(p_j(\bar{b}, b_j) + z(\bar{b}) - c_j(\bar{r}, \phi_j)) + \left(\sum_{i \in j} \eta_{ij}(\phi_j)\right)\right]. \tag{6}$$

Letting $H_j(\phi) = \sum_{i \in j} \eta_{ij}(\phi_j)$, we obtain a profit function that does not have individual-specific terms and can be written using the market share notation that is useful for the empirical analysis.

As the sum of the premium and the baseline level of the subsidy is by construction equal

---

[11]See more details in Medicare Part D Manual. As CMS describes in Chapter 9 of Prescription Drug Benefit Manual, risk corridors are: "Specified risk percentages above and below the target amount. For each year, CMS establishes a risk corridor for each Part D plan. Risk corridors will serve to decrease the exposure of plans where allowed costs exceed plan payments for the basic Part D benefit." (See 42 C.F.R, 423.336(a)(2))

to the bid submitted by insurer to Medicare, $p_j(\bar{b}, b_j) + z(\bar{b}) = b_j$, we can re-write the pre-risk corridor profit of plan j as:

$$\pi_j(b_j; b_{-j}) = M s_j(b)(b_j - c_j) - H_j(\phi), \tag{7}$$

where $s_j(b)$ is the market share of plan j and M is market size. We emphasize that s_j encapsulates all of the regulatory details involved in turning bids, b, into plan-specific market shares.

We now expand this expression to allow for multi-plan insurance organizations as well as to allow for different prices, marginal costs, and demand from the LIS segment. The structure of profit from LIS enrollment is specified as entirely symmetric to the regular enrollees. We denote quantities related to regular enrollees with superscript R, and quantities related to the LIS part of the market with superscript LIS. The profit function for insurer J offering a portfolio of $j \in J_t$ plans across markets $t \in T$ is:

$$\pi_J(\mathbf{b}) = \sum_{t \in T} \sum_{j \in J_t} \Gamma \left[M_t^R s_{jt}^R(b)(b_{jt} - c_{jt}^R) - H_{jt}^R(\phi) + M_t^{LIS} s_{jt}^{LIS}(b)(b_{jt} - c_{jt}^{LIS}) - H_{jt}^{LIS}(\phi) \right]. \tag{8}$$

Firms maximize profits by choosing bid b for each insurance plan in each market.

Equation 8 is more complex than a standard profit function in a differentiated products market due to how the share equation $s_{jt}(b)$ is constructed. For regular enrollees, the share depends on the plan's premium, p^R, which is not set directly by insurers, but rather depends on the bids of other insurers in a non-linear fashion:

$$p_{jt}^R = \max\left\{0, b_{jt} - \zeta \bar{b}_t\right\}, \tag{9}$$

where $\bar{b}_t$ is the enrollment-weighted average bid of all plans in the entire US and ζ is the share of the average bid covered by the federal subsidy. The adjustment ζ is set every year by CMS and is governed by fiscal considerations and the Part D statutes; in 2010, this number was 0.68. The share equation for the low-income segment of the market is substantially more complex. It can be thought about as a piece-wise function with two components: random assignment of low-income enrollees by CMS for those plans that are eligible for random assignment, and enrollment choices by LIS choosers. While LIS choosers are easily modeled in the standard discrete choice demand system, the eligibility requirement for random assignment introduces a discontinuity into the share function. Only plans below the average premium are eligible for random assignment, so for some choices of b_j, the share

function for that portion of the market discontinuously jumps to zero.[12]

Section 5.3 outlines how we deal with the discontinuity in the share function in the estimation. Here, we assume that the firm does not anticipate random LIS enrollees and derive the first-order conditions for profit maximization. We also assume that the firm ignores the effect of its bidding behavior on the average bid, $\bar{b}$; this seems reasonable in light of the over 1,500 PDP plans that, along with the MA-PD plans, determine the average bid. We assume that firms are small enough so that they take the weighted average as given.[13] Then, for a each plan j in firm i's portfolio of J_i plans, the Nash-Bertrand first-order condition is:

$$\frac{\partial \pi_i}{\partial b_j} = M^R s_j^R(b) + (b_j - c_j^R)M^R \frac{\partial s_j^R(b)}{\partial b_j} + \sum_{k \neq j \in J_i} (b_k - c_k^R)M^R \frac{\partial s_k^R(b)}{\partial b_j}$$

$$+ M^{LIS} s_j^{LIS}(b) + (b_j - c_j^{LIS})M^{LIS} \frac{\partial s_j^{LIS}(b)}{\partial b_j} + \sum_{k \neq j \in J_i} (b_k - c_k^{LIS})M^{LIS} \frac{\partial s_k^{LIS}(b)}{\partial b_j}.$$

This expression differs from the more familiar first-order condition in the differentiated product literature in that the market size now plays an important role for the firm's decision-making. The market size affects the relative effects on profit from enrolling regular beneficiaries versus LIS choosers. As we now have one equation in two unknowns, marginal costs for regular enrollees, c_j^R, and marginal costs for LIS enrollees, c_j^{LIS}, we make an additional assumption to close the model. As Medicare specifically increases its risk-adjustment payments to plans for each LIS enrollee, we assume that those payments make the marginal cost of these two groups the same from the point of view of the insurer, so that $c_j^R = c_j^{LIS}$. Imposing this assumption and collecting terms in vector notation, we arrive at:

$$c = b(p) - \Omega^{-1}(M^R s^R(p^R(b)) + M^{LIS} s^{LIS}(p^{LIS}(b))). \tag{10}$$

where

$$\Omega_{kj} = \begin{cases} -M^R \frac{\partial s_k^R(p)}{\partial p_r} - M^{LIS} \frac{\partial s_j^{LIS}(p)}{\partial p_k} & \text{if } \{j, k\} \in J, \\ 0 & \text{else.} \end{cases} \tag{11}$$

As in Nevo (2001), we use Equation 10 to solve for marginal costs in the empirical application

---

[12]Decarolis (2015) discusses the piece-wise structure of the share function and the incentives generated by the LIS random assignment mechanism in much greater detail.

[13]We have experimented with relaxing this assumption and found that the biases were minimal.

below.

4.3 Welfare Metrics

In our counterfactual exercises, we will focus on measuring welfare levels and changes for regular enrollees. For these enrollees, total welfare in the Medicare Part D PDP market is comprised of three pieces: consumer surplus (CS), insurer profits (Π), and the deadweight loss associated with taxation used to fund government subsidies (G):

$$W = CS + \Pi - \lambda G, \tag{12}$$

where λ is the social cost of raising revenues to cover government expenditures, G.

All three pieces of the welfare function are calculated relative to the outside option. For consumer surplus the normalization to the outside option follows directly from the utility model. In Section 4.1 we had defined utility from enrolling in stand-alone Part D prescription drug plans as being relative to the choice of an MA-PD plan or to not purchasing any Part D coverage. For producer surplus, the insurer pricing decision as formulated in Section 4.2 implicitly takes into account the opportunity cost of serving the outside option. In other words, the marginal cost as recovered from the inversion of the first-order conditions incorporates the opportunity costs of potentially serving each consumer in the MA-PD market or not serving the consumer at all. Consequently, the profit function is defined relative to profits that could have been made in the MA-PD program or elsewhere. Finally, since the government subsidizes both PDP and MA-PD parts of the market, we consider government spending in PDP net of what it would have spent on subsidizing the same individual should they leave the PDP market. We conservatively assume that the outside option for the government is the level of average subsidies spent in MA-PD. This assumption excludes the possibility that some individuals could leave subsidized insurance altogether.

Following Williams (1977) and Small and Rosen (1981), surplus for consumer i with marginal utilities θ_i from plan characteristics, including the premium, takes the following form:

$$CS(\theta_i) = \frac{1}{\alpha_i} \left[\gamma + \ln \left[1 + \sum_{j=1}^{J} \exp(v_{ij}(\theta_i)) \right] \right], \tag{13}$$

where γ is Euler's constant, and v_{ij} is the deterministic component of utility for person i from plan j as given in Equation 2.[14] We integrate out over the unobserved taste heterogeneity

---

[14]Euler's constant is the mean value of the Type I Extreme Value idiosyncratic shock under the standard

to obtain consumer surplus:

$$CS = \int CS(\theta) dF(\theta). \tag{14}$$

The second piece of the welfare calculation is producer surplus that we approximate using the pre-risk-corridor version of the profit in Equation 8.[15]

The last piece of net welfare calculations is the deadweight loss associated with raising revenue to cover government transfers to insurance firms and regular Part D beneficiaries enrolled in stand-alone prescription drug plans. In our welfare calculations, we weigh the government spending with the shadow cost of public funds, commonly estimated at $\lambda = 1.3$.[16] Similarly to the outside option reasoning in the case of consumer and producer surplus, we consider how much extra government spending the PDP part of the Part D program (G^{PDP}) generates relative to the outside option of subsidizing the beneficiaries in Medicare Advantage prescription drug plans (G^{MAPD}).

Adding the three parts of the welfare function back together, we have the following measure of average total surplus:

$$W = \overbrace{\int \frac{1}{\alpha} \left(\gamma + \ln \left[1 + \sum_{j=1}^{J} \exp(v_j(\theta)) \right] \right) dF(\theta)}^{\text{Consumer Surplus (CS)}} + \overbrace{\sum_{j=1}^{J} (b_j - c_j) s_j(p)}^{\text{Producer Profit } (\Pi)}$$
$$- \lambda \overbrace{\left(\sum_{j=1}^{J} (G_j^{PDP} - G^{MAPD}) s_j(p) \right)}^{\text{Social Cost of Government Spending (G)}}. \tag{15}$$

While this welfare function describes the surplus for the private market, where firms administer insurance contracts, it does not correspond to the welfare function that a social planner would maximize. If we had the government setting prices for insurance contracts and in effect hypothetically administering these contracts, we would need to take into account the cost of public funds for doing that in the social planner's problem. Another way

---

normalizations in the logit model, and is approximately equal to 0.577. The welfare interpretation of the vintage variable is open to debate; on one hand people may learn to like the plan the longer they are in it, in which case the vintage variable is truly related to utility. On the other hand, if it simply captures inattention, it may not be appropriate to include this as part of welfare. We performed the computation using both interpretations, and did not find that this leads to qualitatively different conclusions in our counterfactuals (which is intuitive, since keeping or removing the vintage effect from the utility function affects all price counterfactuals symmetrically).

[15]This assumes transfers between firms and the government are welfare neutral.

[16]See, for example, Hausman and Poterba (1987).

of thinking about this problem is to imagine that the government dictates prices to private insurers that administer the plans, but then taxpayers cover any shortfall in insurers' profits. Algebraically, both of these interpretations imply that surplus or loss generated in the product market should be weighted with the deadweight loss of taxation. Hence, the social planner's objective function, which we denote with $W^{SP}(p)$ looks the same as equation 15, except that the product market profit is included under the λ-weighted term. The vector of prices that maximizes this version of the welfare function is the social planner's solution. Note that we use prices in the social planner's case, as the distinction between insurer bids and consumer premiums is not meaningful in this case. Using our notation, optimal prices are (see Appendix Section 8 for derivation details):

$$p^{SocialPlanner} = \text{argmax } W^{SP}(p) = c + \Delta G + \Omega(p)^{-1}\frac{(1-\lambda)}{\lambda}s(p). \tag{16}$$

Price is set to marginal cost plus an additional term which adjusts for the opportunity cost of government spending across the inside and the outside options. The final term represents the trade-off between lost consumer surplus and additional product market, a margin which is captured by the social planner.

5 Data and Results

5.1 Data and Descriptive Facts

Our primary data set combines a variety of aggregate plan-level information released annually by CMS.[17] Table 1 reports several summary statistics of the sample. In 2010, about 47 million individuals in the US were eligible to purchase Medicare Part D coverage. Out of these individuals, 13 million obtained coverage through their employer or through other sources such as Veteran Affairs. Out of the remaining 34 million, about 6 million did not purchase any Part D coverage and about 10 million chose to buy drug plans bundled with Medicare Advantage. We consider the latter two groups as choosing the outside option and focus on demand for PDPs for the years 2007-2010.[18]

Nationally, consumers in traditional Medicare chose among 1,500 and 1,800 PDPs each

---

[17]All of the data is publicly available from CMS. CMS tabulates the depository of the data sources at www.cms.gov/Research-Statistics-Data-and-Systems/Research-Statistics-Data-and-Systems.html.

[18]Detailed Medicare Part D enrollment numbers are recorded in Tables 14.x of the annual Medicare and Medicaid Statistical Supplement published by CMS at www.cms.gov/Research-Statistics-Data-and-Systems/Research/MedicareMedicaidStatSupp/Overview.html

year. In this period, there were a total of around 50 insurer parent organizations, with on average 17–19 separate organizations competing in each of the 34 regional markets in which CMS divides the US. Plans were heavily subsidized; while the national average bid was around $1000 per year, consumers generally paid only a third of that, with the remainder covered by government subsidies.[19] Consumers that chose to opt out of the traditional Medicare program and purchase private Medicare Advantage plans received bundled drug coverage in that program. MA-PD plans received the same types of subsidies as PDP plans for each enrolled consumer.

The top left panel of Figure 5 documents a stark downward slope between the level of premiums and the number of competing parent organizations in a market. This suggests that part of the different premium dynamics observed across markets might be due to differences in insurers' market power. A complementary explanation is the different effect of policy-design distortions across markets. According to the literature both motives are likely present and our structural analysis will help disentangling their relative importance.[20]

The other panels of Figure 5 further show three measures of heterogeneity across the 34 markets. The top right panel shows both the premium and the type of plans offered, as well as the level of the LIS subsidy and direct premium. Notably, the plans that offered coverage in the donut hole were consistently more expensive, by a factor of up to three, than plans with only the minimally required coverage. Dispersion in premiums is relatively similar across different markets in the 2010 cross-section. The low-income subsidy averaged $388, but with substantial variation across markets. The bottom two panels report market shares. The left panel reports market shares for regular enrollees and reveals major heterogeneity in plans market shares, both within and across markets. While many plans had market share close to zero, some plans covered as many as 20 percent of eligible beneficiaries within a market. The right panel reports the same information for LIS choosers. We assume that LIS choosers and randomly assigned LIS beneficiaries chose the outside option if they were enrolled in plans that are eligible for LIS random assignment.

In the remainder of the paper, we repeatedly calculate welfare, both under the observed mechanism along with several simulations of counterfactual market structure and regulatory regimes. The evaluation of welfare requires several assumptions on how we calculate government spending on regular enrollees. We calculate government spending on the premium

---

[19]CMS does not directly report bids; however, they can be backed out from the formula for the subsidization mechanism and the data on premiums and subsidies reported by CMS.

[20]Ericson (2014), Decarolis (2015), and CBO (2014) describe how competition and subsidy-related distortions are associated with the evolution of premiums in Part D.

subsidy for stand-alone prescription drug plans as the sum of per capita (simulated) subsidies multiplied by (counterfactual) plan enrollment. We assume that the average reinsurance, or what we call "non-premium" subsidy for each plan is fixed and does not change across counterfactuals. Thus, total reinsurance subsidies only change across counterfactuals due to enrollment changes. We use CMS annual reporting on average reinsurance payment for each Part D plan as the measurement of the non-premium subsidy.[21] In 2010, for example, the unweighted mean per capita reinsurance payment among PDP plans was $503 per plan.

In addition to calculating the premium and non-premium subsidies on stand-alone prescription drug plans, we also estimate the government's opportunity cost of having individuals enroll in PDP plans. Based on the choices of individuals switching plans observed in the consumer-level data, we assume that if individuals switch from the inside option of PDP plans to the outside option, they switch to the MA-PD program rather than leave drug insurance altogether. Thus, the government is still likely to incur subsidy spending for these individuals through the MA-PD program. To account for the MA-PD spending, we use CMS data to calculate average observed level of government spending on premium and non-premium (re-insurance) subsidies in the MA-PD program. We observe that the average per capita premium subsidy in the MA-PD program is $686, while the average non-premium subsidy is $260. This amounts to a total of $946 government spending per capita on individuals enrolled in the MA-PD program.[22] We use this average spending together with enrollment predictions for inside and outside options in each counterfactual to calculate the total opportunity cost for the government of having individuals enroll in PDP rather than MA-PD program. It is crucial to emphasize that we assume this number does not change across our counterfactuals, as we are focusing on the mechanisms of determining subsidy levels within the PDP part of the program. This implies that while at the PDP subsidy levels observed in the data, the per capita government spending is higher on the PDP plans rather than MA-PD plans, this relationship can reverse in counterfactuals where we increase the PDP premium subsidy.

---

[21]For raw CMS data, see http://www.cms.gov/Medicare/Medicare-Advantage/Plan-Payment/Plan-Payment-Data.html.

[22]These numbers are calculated using annual Reconciliation Files released by CMS at http://www.cms.gov/Medicare/Medicare-Advantage/Plan-Payment/Plan-Payment-Data.html?DLSort=0&DLPage=1&DLSortDir=ascending. For each plan that participated in Medicare Part D, these files report monthly "Average Part D Direct Subsidy" as well as monthly "Average Reinsurance Payment". We take unweighted averages of each of these numbers for MA-PD plans and multiple them by 12 to arrive at the subsidy levels reported in the text.

5.2 Demand Parameters

Table 2 reports the demand estimates. Columns (1)–(3) are for regular enrollees, while columns (4)–(5) are for LIS enrollees. Column (3) reports the estimates of the random coefficient logit model described earlier, which is our preferred specification and the demand system used in the computation of the counterfactual scenarios in the next section. For comparison purposes, we also report estimates for the Berry (1994) logit model; column (1) contains OLS estimates, while column (2) contains 2SLS estimates. Columns (4) and (5) report analogous OLS and IV estimates for the LIS enrollees.[23]

Models (1), (3) and (5) are estimated using instrumental variables. We instrument for plan premiums and assume that other characteristics of the contracts are exogenous in the short run. We motivate this by observing that, while bids for a given plan vary substantially over time, insurers offer a rather stable portfolio of contract types over time (Polyakova, 2015).[24] The concern regarding the bias in the coefficient on premiums is that they might be correlated with an unobserved quality aspect of plans that we fail to capture with the observed characteristics. While we include a rich set of observed plan features, we may not be fully capturing insurer-plan specific customer service or advertising efforts, as well as issues such as drug prices driven by insurers' bargaining power. Some of the variation will be insurer-specific rather than plan specific and so will be captured by insurer fixed-effects. For the remaining variation, we rely on four instrumental variables. Three are BLP-style instruments, measuring the number of PDP or MA-PD contracts that the same insurer offers in the same or different regions. The other instrument is a version of the Hausman instrument measuring prices charged for the similar plans in other geographic markets. Specifically, we construct the instrument by including the lagged enrollment-weighted average of prices of plans offered in other regions in the same macro region and in the other macro-regions by the same company, where macro-regions are defined as three large geographic areas in the US. The idea of this instrument is particularly appealing in our setting due to the regulatory structure of the market, where markets are separated by CMS. Instrumenting the price in one region with the prices of the same contract in other regions, allows us to isolate the variation in prices that is common across these contract due to, for example, particular agreements

---

[23]We also estimated a BLP-style estimator for LIS enrollees, but omit the results here since we did not find any statistically significant random coefficients; as such, the point estimates in the BLP estimator are identical to those of the IV logit.

[24]For example, if an insurer offers a contract with some coverage in the gap in the first years of the program, this insurer is likely to continue offering a contract with some coverage in the gap. The amount of coverage may change, but the dummy-measure that we use for whether there is any coverage in the gap does not appear to respond to short-term demand shocks.

of a given insurer with pharmaceutical producers, and is thus not correlated with market-specific unobserved quality due to, for example, local marketing. The IV first stage estimates are reported in Table 3. They indicate a positive coefficient for the Hausman instrument, suggesting that plans that cost more in other regions are indeed likely to be priced higher in a given region. The first stage is jointly statistically significant with an F-statistic of 246.

All specifications result in significant and negative coefficients for premiums; as is usual, the estimated price coefficients are biased downward in specifications without instruments. The BLP specification suggests that there is a significant but relatively tight dispersion of the price coefficient in the population. This is an important finding, as heterogeneity in consumer tastes is one of the motivating principles for allowing choice in public insurance programs, as consumers can increase efficiency by aligning themselves with plans that are better-suited to their tastes. To make the comparison between the first three estimates comparable, we note that the estimated modal price coefficient is -12.30, which is close to the IV logit estimate.[25] The average own-price elasticity for regular enrollees across all plans in Medicare Region 1 in 2010 are -1.19, -4.31, and -5.98, for the OLS logit, IV logit, and BLP specifications, respectively. The elasticity of the BLP specification, which is our preferred estimate, is economically reasonable and similar to those reported in Lucarelli et al. (2012).

The remainder of the plan characteristics are estimated to have coefficients with signs according to intuition. Consumers dislike higher plan deductible, but enjoy measures of plan generosity: coverage in the gap, broader coverage of common drugs, and pharmacies in their networks all give higher utility. We also note an economically and statistically significant positive coefficient on the vintage of plans, suggesting that plans that entered earlier in the program were able to capture a larger beneficiary pool. We do not find significant dispersion for two other variables for which we allow random coefficients: the inner option and the dummy for gap coverage.

Columns (4) and (5) report OLS and 2SLS estimates of the Berry logit model applied to LIS choosers. To estimate LIS demand, we shut down the deductible and gap coverage characteristics of plans, as individuals eligible for low-income subsidies receive additional support from the government that helps cover these out of pocket expenditures. We also adjust premiums to reflect the additional subsidies for the LIS enrollees. As described in Section 4.1, we have to make several additional assumptions to formulate a meaningful demand system for the LIS market. The key assumption is that all individuals that we observe in plans that are eligible for LIS random assignment are considered to have chosen

---

[25]The average own-price elasticity is -13.69 and the median is -13.20.

the outside option. The results of the OLS and IV specifications for this part of the market are quite similar to the demand estimates for regular enrollees. Individuals prefer plans with more generous formularies and larger pharmacy networks, and plans that have existed on the market for longer time are also more likely to attract beneficiaries. The price coefficient in the OLS specification is almost identical to the one for regular enrollees. In the IV specification, it is slightly lower at -7.6, suggesting lower price sensitivity to prices that, recall, are about $400 lower per year for the LIS enrollees.

We extensively explored the robustness of the above estimates. In particular, we evaluated three sets of robustness checks. First, we estimated alternative versions of the random coefficient model using alternative model specifications as well as different identification strategies based either on market-level demographics (Nevo, 2001) or on demographics and micro-moments constructed from the consumer-level data (Petrin, 2002).[26] Second, we explored the effects of using alternative instruments like those proposed in Train and Winston (2007) and in Gandhi and Houde (2015). Third, we evaluated alternative estimation methods for the LIS demand. First, we applied to LIS choosers the same random coefficient model described above. Second, we estimated LIS demand as an out of sample prediction based on regular enrollees demand. Namely, we used the random coefficient model inclusive of demographics for the regular enrollees and then predicted LIS demand using the LIS enrollees demographics. Overall, these checks broadly confirmed the results in Table 2. Albeit not reported in the paper, all the results are available from the authors upon request.

5.3   Marginal Cost Estimates

Following Nevo (2001), we use the demand estimates for regular and LIS enrollees with Equation 10 to recover plan-level marginal costs. As shown in Section 4, the first-order condition linking marginal costs to bids and demand parameters has several non-standard features, particularly the discontinuous share function for LIS enrollees. Therefore, in order to proceed with the estimation of marginal costs, we make several important assumptions. First, we assume that the multitude of risk-adjustment and reinsurance mechanisms implemented in Medicare Part D imply that insurers de facto face constant expected marginal costs. Second, we select a subset of plans that were plausibly not distorted by LIS gaming.

In essence, the idea is to select a group of plans for which we find the Bertrand-Nash

---

[26]We also assessed the reliability of our random coefficient estimates to various computational issues, including the estimation algorithm (nested fixed point relative to MPEC), the optimization routine (gradient relative to non-gradient based) and the integration method (quadrature relative to sparse grid).

assumption acceptable for describing the pricing behavior of the insurer. We construct a group of such plans by selecting all contracts of those insurers that within a given market (year-region) were not eligible to enroll randomly assigned LIS individuals into any of their plans. Even if the assumption that this group of "non-manipulating" plans does not express premium distortions related to the LIS benchmark, we may be still be worried that these plans are not comparable to plans qualifying for low-income enrollees. Empirically, this does not seem to be the case. There has been substantial variation in the low-income subsidy across regions and there are many insurers who never qualified for low-income enrollees in at least one region. This variation is mostly due to the different penetration of Medicare Advantage: where in 2006 enrollment in Medicare Advantage was high, MA-PD received a high weight in the calculation of the low-income subsidy and, since their premium is typically close to zero, they induced a small low-income subsidy (Decarolis, 2015).[27] Hence, we invert the first-order conditions only for these "non-distorted" plans. Since the plans offered by the same insurer across different regions are remarkably similar, the marginal cost estimates of the "non-distorted" plans through the inversion of the first-order condition can be used to predict the cost of similar plans in other regions for which we could not directly infer costs.

In the first step, we recover the marginal costs for "non-distorted" plans using Equation 10. We use 756 plans offered in 2010 data, as our counterfactuals will focus on this year only. We then proceed to relate the estimated marginal costs to the observed characteristics of non-manipulating plans by estimating the following hedonic-style linear regression:

$$mc_{jt} = X_{jt}\beta + \delta_t + \tau_j + \epsilon_{jt}, \tag{17}$$

where X_{jt} includes the same non-premium characteristics of plans that we had included in the utility function. We add the unobserved quality estimate for each plan as an additional explanatory variable in X. We condition the regression on firm and market fixed effects to account for inherent differences in marginal costs across insurers and geographic regions.

Table 4 reports the coefficients for the hedonic regression. We note that the most important determinants of marginal costs appear to be, as expected, the plans' coverage limits, as well as the generosity of their drug formularies. For example, we estimate that offering coverage in the gap increases a plan's marginal cost by $405 a year, which is a large increase relative to the average marginal cost estimate of about $1,000 from the inversion procedure. This estimate of the additional marginal cost from coverage in the gap roughly corresponds

---

[27]The variation in the total weight assigned to MA-PD in 2006 is substantial ranging from almost 60 percent in Arizona and Nevada to less than 4 percent in Mississippi and Maine.

to the premium add-ons that are charged by insurers that offer coverage in the gap.

We use the estimates of how plan characteristics translate into marginal costs to predict marginal costs for all plans that we did not include in the inversion procedure. This exercise hinges on the assumption that all plans have a similar "production function." In other words, we assume that the plans that manipulate the LIS threshold manipulate the premiums, but do not have different marginal costs conditional on a set of non-price characteristics. This appears reasonable, as the main source of costs in the insurance market is determined by individual risk spending; therefore, it is conceivable to assume that plans with the same financial characteristics and formulary generosity will have similar marginal costs.

Figure 6 plots the distribution of predicted marginal costs and compares it to the estimated distribution via the inversion procedure. We estimate substantial heterogeneity in the marginal costs across plans. We find this heterogeneity both through the inversion procedure and in the hedonic projection exercise. This indicates that our hedonic-style regression captures the key drivers for the differences in marginal costs. The manipulating plans are estimated to have slightly lower marginal costs on average, which is intuitive if we believe that cheaper plans are the ones that would try to compete for LIS enrollment. The marginal cost estimates in both cases are centered around \$1,100 and range from about \$750 to about \$1,900.

Our marginal cost estimates imply mark-ups in the order of 9 percent on average. The bottom panel of Figure 6 plots the full distribution across plans. These mark-ups are fairly low, suggesting that the environment is reasonably competitive. We explored several external sources that discuss the profitability of the Part D PDP plans to verify that our estimates of marginal costs and mark-ups appear plausible. A CBO report from July 2014 (Congressional Budget Office, 2014) provides, to our knowledge, the most detailed publicly available external analysis of revenues and costs in stand-alone prescription drug plans for years 2007-2010. According to Table 2-1 of this report, in 2010, PDP stand-alone plans collected on average \$1,136 in direct subsidies and premiums per person. This measure is similar to our measure of bid in the mark-up formula (bid-marginal cost)/(marginal cost). To construct a proxy for marginal cost from the CBO report, we take the estimate of the "Net Drug Spending for the Basic Benefit," which is reported to be \$1,382 and subtract the reinsurance payments that are reported to be \$521 on average per person.[28] This gives us an estimate of direct costs for basic benefit of \$861. Using the latter number—which certainly is within the scope of

---

[28]Since our marginal cost estimates come from the inversion of the insurers' first-order conditions, they are net of reinsurance transfers between the plans and the government.

our estimated marginal costs—and the average revenue reported above, we would arrive at a mark-up of (\$1,136-\$861)/\$861, or 32 percent. These numbers, however, do not include administrative costs (which CBO reports to be \$242, but it is calculated as a residual and includes profits and administrative costs), and further does not include "enhanced" benefits that are covered by additional premiums and are included in our calculations. To arrive at around a 10 percent markup, we would have to assume that administrative costs are around \$170 on average per person: (\$1,136-\$861-\$170)/(\$861+\$170)=10 percent, or 70 percent of the number that CBO reports as an estimate of administrative costs plus profits, which appears plausible. Overall, our estimates appear realistic and consistent with the actual cost data.

5.4   Welfare Estimates of the Current Mechanism

Using the demand and marginal cost estimates from above, we next compute consumer surplus, producer profits, government transfers, and total surplus for the observed market allocation, as outlined Equation 15. We restrict our calculations to regular enrollees for for consistency across our counterfactuals below, which remove the LIS enrollees from the market. For expositional clarity, we report results for a single year (2010). The calculations are reported in the first column of Table 5.

We estimate that total annual consumer surplus generated by Part D PDP was about \$2.5 billion, or about \$280 per enrollee.[29] We also estimate that insurer profits, before risk corridor adjustments, were about \$460 million.[30] In sum, the Part D PDP market generated a little less than \$3 billion of consumer and producer surplus. This surplus came at a price, however, as CMS reports that the government spent about \$9.5 billion in subsidies in Part

---

[29]This is calculated as the sum across markets of average consumer surplus in each market multiplied by the market size of each market. Note that the consumer surplus estimate relies solely on the demand model for regular enrollees and does not depend on any assumptions or specification of the supply side.

[30]In the estimation of insurer profits from regular enrollees in the counterfactuals, we let the per capita revenue from regular enrollees be equal to the bid of the plan as simulated in the counterfactual plus the enhanced component of the premium as observed in the data if the plan is enhanced. We thus assume that the enhanced component of the premium does not change across counterfactual mechanisms and does not enter bidding. We then take the difference in estimated marginal cost and counterfactual per capita revenue for each plan and multiply it by the counterfactual share of each plan in each market scaled by the market size of the regular enrollees' market. Relative to the profit function formulation in Equation 8 we are not explicitly calculating the effects of risk-corridors that may alter profits at the end of the fiscal year. Our welfare estimates would be negatively affected if risk corridor payments to firms were positive, as the transfer would be adjusted by the deadweight loss of raising taxes. On the other hand, transfers from firms to the government would be welfare neutral. We are further not explicitly calculating the selection component of the profit function $H(\phi)$, which is zero in expectation by construction.

D PDP, split between $6 billion on premium subsidies and $3.5 billion on non-premium subsidies. Under our assumption that the government would still have to pay subsidies if individuals were to leave PDP and switch to MA-PD, we also estimate the government would have spent $8.2 billion on the same individuals were enrolled in MA-PD. The difference between the total PDP subsidies and what these would have been in MA-PD, a little more than $1 billion, is the opportunity cost of government spending on PDPs, which, along with our assumption that the deadweight loss of government taxation is 30 cents on the dollar, gives us the social cost of government spending of $1.4 billion.

We estimate the total surplus generated by the Part D PDP market was about $1.5 billion. This is one of the primary findings of the paper, along with the observation that the vast majority of this surplus came from foregone government expenditures. That finding follows directly from our assumption that all enrollees in Part D PDP would have been enrolled in MA-PD had the Part D PDP market not existed. This is likely an overestimate of government spending, as it is possible that some of the enrollees in Part D PDP would have decided not to obtain drug coverage of any variety. However, we think that this assumption provides a plausible upper bound for the welfare gains of the Part D PDP program. We also compute a lower bound by assuming that the opportunity cost of government spending was zero. In that case, we estimate that a dollar of public funds would have generated only 0.17 dollars of surplus per subsidy dollar. That is, welfare would have been improved dramatically by shutting down the program entirely. However, if we do take the government's payments for the outside option into account, we arrive at a more encouraging calculation of 1.12 dollars per opportunity-cost-dollar spent by the government. The break-even opportunity cost of government spending is $261 per enrollee; if government spending outside Part D PDP exceeds this number, the Part D PDP market generates non-negative social surplus.

While the total amount of surplus generated in the market is positive, this calculation does not provide any guidance about how efficient the current subsidy mechanism is relative to the first-best, or how its various components contribute to the overall welfare outcomes. To address those questions, we next turn to assessing the welfare properties of counterfactual mechanisms which change the way that subsidies are determined.

6 Counterfactual Results

We are interested in understanding how supplier incentives, the subsidy mechanism, and their interplay determine market outcomes. Recalling the framework in Section 2, the essential

economic ingredients are consumer willingness-to-pay, the exercise of market power, and the generosity and form of the subsidy mechanism. We establish a baseline by first removing the link in the Part D PDP subsidy mechanism tying regular and LIS enrollees. There are two reasons for doing this. First, removing the link between the regular and LIS enrollees has been a long-considered policy proposal; theory suggests that tying the subsidies, prices, and product choices across the two markets potentially introduces distortions and reduces overall welfare. Understanding the effect of tying the two markets together is of independent interest. The second reason is technical: gaming for the LIS enrollees introduces both dynamic incentives and discontinuous payoffs.[31] By removing the LIS enrollees from the subsidy mechanism for regular enrollees, we avoid the complex modeling issues associated with those two factors. We also sever the link between the subsidies in Part D PDP and prices in the MA-PD market. This is also an independently interesting counterfactual, and doing so also allows us to ignore possible feedback effects in the MA market from changes made in the Part D PDP market.

6.1   Links to Other Markets

To isolate the effect of the current subsidy mechanism for regular enrollees, we begin by simulating the market once we remove the interconnections existing in the subsidy formula with features of the LIS and MA-PD portions of the market. In the first counterfactual, we remove the LIS enrollees to their own market while holding all other aspects of the current subsidy mechanism unchanged. Hence, prices are set as we describe above in Section 4, where the baseline subsidy is a fraction of the enrollment-weighted average of prices from Part D PDP and MA-PD.[32] In the second counterfactual, we remove the MA-PD bids from the computation of the weighted average used in computing the baseline subsidy. In this case, the baseline subsidy is set as 68 percent of the average Part D PDP bid. This second counterfactual gives us a benchmark to compare other subsidy mechanisms for regular beneficiaries, keeping separated issues associated with LIS assignments and MA-PD bids. The effects of these two counterfactuals are described in Columns (2) and (3) of Table 5.

The results of the first counterfactual where the tie between regular and LIS enrollees is severed illustrates the rich equilibrium effects of changes to the subsidy mechanism. The key

---

[31]See Decarolis (2015) for a treatment of these two issues.

[32]We note that we do not directly observe MA-PD bids, as MA-PD premiums contain not only subsidies whose levels are known but also unobserved Part C rebates. However, it is easy to back out the average MA-PD bid because we observe the base subsidy and can compute the average Part D PDP bid; the remainder is the average MA-PD bid.

change is that firms previously had incentives to lower prices of some of their plans in order to qualify for random LIS enrollments. This incentive would lower average prices, which in turn lowers the generosity of the baseline subsidy. Removing this incentive leads to two related equilibrium effects: firms have incentives to raise bids since they are no longer competing to get under the average premium, and the baseline subsidy increases as the average price increases. There is a feedback effect between the two forces, as higher prices lead to more generous subsidies which in turn lead to higher prices and so on; we compute the equilibrium that results when firms' projection of the average subsidy is equal to the one resulting in equilibrium under those beliefs. This is born out in the results in column (2); the average government subsidy per enrollee increases from $1,069 to $1,117. The combination of the more generous subsidy, along with a shift in the relative prices of Part D PDP plans, results in slightly lower enrollment-weighted average premiums ($471 vs $502 in the data), and a seven percentage point increase in the share of inside options. Overall consumer surplus increases, but tellingly the average consumer surplus declines from $287 to $274 per enrollee. The role of the LIS incentives in disciplining firm prices is shown in the dramatic increase in firm profits, which more than double to over $1 billion. Echoing the discussion in Section 2, increasing the generosity of subsidy induces relatively low marginal valuation consumers to enter the market, with some of the rents being captured by firms. The question becomes whether the extra value generated by the change in the mechanism exceeds the social cost of providing the goods. In this case, the answer is a clear no: although consumer surplus and producer profit increase by almost $1 billion, government expenditures increase by more than $2.5 billion. Accounting for the opportunity cost of government spending, total welfare declines by almost $100 million when the regular and LIS enrollee markets are unbundled. This counterfactual mechanism results in an inefficient system: the per dollar return on government spending drops to about 12 cents per nominal dollar and only 59 cents per opportunity cost dollar.

The removal of the MA-PD plans from the enrollment-weighted average bid used to compute the baseline subsidy results in even stronger effects along the same lines. As the MA-PD bids are quite low and thus typically drive down the average bid, removing their bids increases the average bid both mechanically and due to insurers' strategic response. Consequently, the subsidy further increases to $1,170 per enrollee, with the average premium declining to $418. Enrollment in the market increases to 55 percent and producer profits increase by $210 million. While total consumer surplus increases by another $600 million, the marginal consumers drop the average consumer surplus per enrollee to $263. Higher

subsidies also imply higher government spending of $15 billion on Part D PDP subsidies, and thus the net result is a decrease in total surplus of $550 million. The combination of these factors drives the per dollar return on government spending to about 6 cents per nominal dollar and 21 cents per opportunity cost dollar.

The results of these counterfactuals are instructive. They highlight the complex equilibrium interactions arising here between the endogenous-determined subsidies and social welfare along the channels discussed in Section 2. However, they also illustrate the potential for large welfare effects arising from seemingly-unrelated choices in mechanism design. While it is possible that the designers of the Part D PDP subsidy mechanism were able to predict the directional effects of tying prices in the Part D PDP market for regular enrollees to the eligibility of those same plans in the LIS market, and of linking the baseline subsidy to MA-PD prices, it seems unlikely that anyone could ex ante predict the exact magnitudes of those choices. As it turns out here, the designers' choice to link on both of those margins resulted in higher levels of welfare. However, we emphasize that, again, the welfare increase only comes from the foregone opportunity cost of government expenditures in Part D PDP. Both consumer surplus and producer profit are higher when the links are removed. The policy implication revolves solely around the ability of the mechanism to prevent socially-costly government expenditures. That does not have to be the case in general; as we show below, there are a range of alternative counterfactual mechanisms that may do better on all fronts. First, we explore the role that market power plays before considering alternative subsidy mechanisms.

6.2 Market Power

Before proceeding to counterfactuals that change the subsidy mechanisms, we first evaluate the effect of market power. Thus, we consider the current CMS mechanism (without the LIS and MA-PD pricing links) and explore two polar cases: full competition and monopoly. We implement these two counterfactuals by simulating alternative plan ownership structures within each CMS-defined market. In the first, we assume that each PDP is its own firm; in the second, we assume that every PDP in each market belongs to one firm. The results are reported in the last two columns of Table 5. Our results are somewhat surprising on first glance: contrary to the usual introductory microeconomics intuition, social welfare basically remains unchanged under atomistic competition but increases substantially under monopoly ownership. Under atomistic competition, consumer surplus and producer profits actually decline modestly, driven by the fall in the government subsidy per person to $1,166. The net

34

effect of the three changes is essentially zero. However, under monopoly consumer surplus decreases by $600 million to $2.8 billion and producer profits go up almost $900 million to $2.2 billion under the monopoly ownership structure. The real key to increasing social welfare is that the market share of the inside option declines to 45 percent; the consumers who remain in the market are the ones with relatively high valuations, and government subsidies also decline. The combination of these factors pushes welfare higher to $1.1 billion under monopoly. This comports again with the idea that, holding the subsidy mechanism fixed, market power may actually be beneficial in preventing consumers with valuations lower than social cost from entering the market.

The results suggest that, without links to the LIS market and MA-PD, the current ownership configuration delivers outcomes fairly close to that of a purely competitive ownership structure. This is an interesting result, as one of the motivating reasons for using managed competition to deliver publicly-subsidized goods and services was to leverage competition to reduce prices. Our conclusion on this point requires caution, however, as we cannot assess the counterfactual of possible alternatives, which may include a standard government-run program or a regulated monopoly. For one, we do not know anything about the government's, or regulated firms', comparative marginal costs of delivering similar insurance plans in such a scenario, and therefore we cannot make any claims about additional efficiencies introduced by competition. Second, we take the marginal costs of firms as given here; it is likely that a single buyer would be able to exert monopsony power in negotiating with upstream pharmaceutical companies. The combination of the two effects is ambiguous, and we remain conservative in our conclusion that the current ownership structure gives results similar to that if all plans were independent firms.

6.3   Alternative Subsidy Mechanisms

So far, we have considered the effects of linking the current subsidy mechanism to other markets, and the role of market power within that mechanism. We now start to explore alternative subsidy mechanisms. In the next set of counterfactuals we ask whether deviations in the subsidy mechanism form the averaging rule currently used to proportional or flat subsidies could improve total welfare.

6.3.1 Proportional Subsidies

We start with a proportional subsidy mechanism in which premiums are given by: $p_{jt} = xb_{jt}$, with $x \in [0,1]$. Table 6 the findings for three levels of x: 5 percent, 32 percent, and 95 percent. The idea of the 5 percent counterfactual is to test how insurers would respond if the government almost entirely bore the cost of purchase. With the almost-complete reduction in consumer price sensitivity, insurers dramatically increase their bids: we find a threefold increase in bids. Consequently, even though individuals now pay only 5 percent of the bids, the premiums are still relatively high at \$190 a year on average. This drop in premiums, however, is sufficient to increase enrollment to 100 percent in the PDP plans from 55 percent in the baseline. Consumer surplus increases to \$13 billion, and insurers are able to capture huge amounts of the rents generated by the subsidy—profits shoot to over \$50 billion. Given the dramatic increase in bids and enrollment, government spending increases enormously. The result is a stark drop in welfare levels to an astounding negative \$37 billion. Similar results are obtained at 32 percent, which corresponds to the fraction that CMS currently uses after taking the weighted average of bids to determine premiums and subsidies. In both cases, welfare losses are driven by low-valuation consumers entering the market, the subsidy being far too generous, and firms capturing most of the rents. The two more generous proportional subsidies act as extremely expensive mechanisms for transferring funds from taxpayers to firms.

A different story emerges when consumers have to 95 percent of the bid; effective consumer prices triple over the baseline level. At this level of prices, very few individuals, about 1 percent, are willing to purchase PDP plans and switch to the outside option. Despite the drop in enrollment, the program generates even more surplus than the benchmark case, and a very high surplus per subsidy dollar. Such a high surplus per dollar is not surprising, as the government pays little and only the beneficiaries with the highest willingness to pay participate in the program.

This case highlights one of our central empirical findings: there is very little innate willingness to pay for the Part D PDP plans outside of their subsidies. This is likely due to the fact that there are close substitutes readily available at very low (even zero) prices in the form of MA-PD. While traditional Medicare and Medicare Advantage are not identical substitutes, the 5 percent counterfactual highlights that the marginal consumer surplus generated by this program is minimal. Of course, we have kept the subsidies in the MA-PD program unmoved, which are also artificially attractive due to government subsidies; we return to this point below when we introduce a link between the generosity of the subsidy in Part D PDP and

MA-PD.

Overall, these results can be interpreted by considering that proportional subsidies have two effects relative to the observed mechanism. First, there is a price level effect, as a very generous subsidy would decrease the overall level of premiums faced by consumers, even though firms capture much of these rents by increasing their bids. Second, there is a significant change in relative prices that makes the more generous plans relatively more attractive. The counterfactuals illustrate the strong impact of subsidy structure on insurer behavior. In cases where consumers do not face 100 percent of the extra premium in more expensive plans, competitive forces are significantly muted and insurers pass through substantially higher expenditures to the inelastic federal budget.

6.3.2   Flat Subsidies (Vouchers)

The second alternative subsidy system that we explore is a flat voucher mechanism, where F is given to each Part D-eligible consumer to be spent on a Part D PDP plan if they choose. The premiums are set to equal to: $p_{jt} = \max\{0, b_{jt} - F\}$. The last three columns of Table 6 report results for three flat subsidy levels: $0, $721 and $1340.

Column (4) shows that $F = \$0$ is associated with such high premiums that essentially nobody enrolls in PDPs. Not surprisingly, this case is very similar to that in the preceding column of Table 6. At the other extreme, $F = \$1,340$, which is double the subsidy level observed in reality, turns out to be a sufficiently generous subsidy to guarantee 100 percent enrollment in PDPs. While full enrollment in Part D PDP plans was also attained with the 95 percent proportional subsidy, the total welfare in the two cases is remarkably different. Although both entail negative welfare, the flat subsidy does far better: a loss of $6.7 billion relative to $37.4 billion. The key difference is that consumers still must pay every marginal dollar for plans priced above the voucher, so while insures utilize the increase in subsidy and increase their bids, they do not do so dramatically. In fact, almost all firms submit bids at exactly the voucher. Consumers face essentially zero premiums, leading to an increase in consumer surplus to $13 billion, which is essentially the upper bound without giving consumers money. Producer profit increases substantially to $4.2 billion. However, as seen previously, government expenditures overwhelm these welfare gains, resulting in a large welfare loss.

The third case, $F = \$721$ is particularly noteworthy. Although $F = 721$ implies that the subsidy is nominally close to the observed subsidy, incentives are very different under a flat

voucher relative to the observed mechanism.[33] While in the observed mechanism, the subsidy is determined after the bidding process as a fraction of the average bid, here the subsidy is set ex ante and it does not depend respond to endogenous supplier behavior. Compared to the benchmark case, we find that prices that consumers face increase by about $100. This leads to a drop in enrollment to 43 percent, compared to the benchmark 55 percent, and an accompanying decrease in total consumer surplus and producer profit. Government spending patterns, however, change starkly in response to decreased enrollment and thus overall, the total surplus under the voucher, $1.7 billion, is higher than in the benchmark case of $0.9 billion.

The change in welfare across the range of vouchers suggests that there is an optimal voucher above zero. Therefore, we repeat the calculation of total welfare at a range of vouchers from $0 to $1400 in increments of $100. The top panel in Figure 7 summarizes the outcome of these calculations graphically. We find that the total welfare is the highest at $F = 600$. Setting vouchers higher than that level significantly reduces welfare. So do less-generous vouchers, but the welfare gradient and thus the cost of deviation from the optimum is lower. Setting the PDP voucher at zero, still leads to total welfare of $1.1 billion. This is $900 million less in total surplus than at the optimal voucher. Setting the voucher at $600 above the optimum, on the other hand, results in a welfare loss of about $4 billion. Column (5) in Table 7 reports the welfare estimates for the optimal national voucher. Total surplus is $2 billion, or twice the baseline case.

In looking at welfare outcomes at different voucher levels, we held the voucher constant at the national level. At the same time, we have documented substantial heterogeneity in demand, supply, and prices across the 34 geographic markets. Thus, the next dimension of regulatory intervention we explore is allowing the government to set geographically-differentiated subsidies across regions. We focus on the flat voucher mechanism, as regionally-differentiated vouchers would be the simplest policy change to implement. To implement this counterfactual, we compute welfare at different levels of possible vouchers (from $0 to $1400 at $100 steps) within each region, and then for each market select the voucher that results in the highest welfare within that region. The results of this exercise are reported in Panel 2 of Figure 7. We find that in 19 out of 34 markets, it is optimal to set the same voucher subsidy

---

[33]This level of subsidy deviates slightly from the observed level of $677, because it reflects our re-calculation of the subsidy levels by applying formal CMS rules to the bids. The difference comes from slight additional annual adjustments that CMS makes to the multiplier on the average bid as described in note to Figure 3 in www.medpac.gov/documents/payment-basics/part-d-payment-system-14.pdf?sfvrsn=0. We performed the welfare calculations using both voucher levels, and there are only very minor differences in the exact dollar amounts.

that would have been the uniform optimum - at \$600. In other regions, however, it would be welfare-maximizing to deviate from this subsidy. We find that in two markets, it would be optimal to offer higher subsidies of \$700 and \$800, while in the remaining markets it would be welfare-increasing to lower subsidies by \$100-\$200. Figure 7 illustrates welfare gains per market that could be achieved through these adjustments to vouchers. Column (6) of Table 7 reports the welfare calculations for these market-specific optimal vouchers; total welfare would increase by \$40 million to \$2.1 billion.

6.4   Marginal Cost Pricing and the Social Planner

The relatively good performance of the voucher system compared to both the current system and the proportional subsidy indicates a potentially simple policy reform that could improve welfare. Nevertheless, it is interesting to quantify to what extent such a decentralized approach manages to move the market closer to some ideal benchmark. In this final part of the analysis, we thus compare the level of welfare attainable under either marginal cost pricing and the social planner. Columns (1)–(3) in Table 7 reports the full set of welfare estimates for private marginal cost, social marginal cost, and the social planner, respectively.

We start by setting premiums equal to the estimated marginal costs for each insurance plan: $p_{jt} = MC_{jt}$. The results are reported in column (1). Facing premiums as high as marginal costs, consumers leave the PDP program in favor of the outside option with enrollment dropping to 1 percent. The level of total surplus is nevertheless high, since only consumers with the highest willingness to pay enroll in the program, while the government is paying relatively little as it provides only "non-premium" subsidies.[34] We expand upon this counterfactual and let consumers face the full social marginal cost rather than only the marginal cost of insurers that was estimated under the existing reinsurance subsidies. To calculate the premiums, we add average observed re-insurance subsidies (RIS_{jt}) for each plan to the estimated marginal costs: $p_{jt} = MC_{jt} + RIS_{jt}$. The outcomes, reported in Column (2) of Table 7, are similar, albeit starker. In this case, enrollment and total surplus fall even more, with nearly none of the market choosing the inside option of PDP plans.

Finally, we consider the social planner's problem. As expected, the social planner's problem generates the highest total surplus at \$3.6 billion. This surplus comes at a cost

---

[34]The details of how we calculate "non-premium" or reinsurance subsidies are discussed in Section 5. Recall that these represent government payments in the catastrophic coverage part of Part D plans, whereby for the individuals with the highest spending levels the government directly pays 80% of expenditures after a certain threshold. This effectively amounts to additional subsidy dollars that, however, do not directly depend on insurers' bids.

of large subsidies. We calculate that the optimal prices in PDPs are on average lower than in the benchmark case, at \$372. In addition to premium subsidies, the government carries the full cost of the program, including the coverage of insurer losses of \$10 billion. As the algebraic expression for social planer's prices in Equation 25 suggests, the social planner sets prices for each plans as a function of this plan's social marginal cost and a fiscal adjustment term. The latter takes into account how much enrollment in a given plan would cost the government.

Contrasted with the optimal voucher results reported in the last two columns of Table 7, voucher welfare is higher than that generated under (social) marginal cost pricing. However, not even the market-specific optimal voucher manages to get near the welfare value under the social planner. The social planner's solution demonstrates the idea that at flat subsidy rates that are unrelated to the efficiency of individual plans, subsidies distort the allocation of individuals across plans within a given market. Figure 8 shows this point graphically on the example of California's market in 2010. Relative to observed prices, the social planner's solution is to increase premiums in plans with higher social marginal costs. This results in the re-allocation of individuals across plans—market shares of plans with lower social marginal cost increase, while the market shares of plans with higher social marginal costs decrease. Note, however, that the effects are not monotonic. For example, for some plans with coverage in the gap, which have high social marginal cost, prices increase substantially, but the market share almost doesn't change, as there is still enough willingness to pay for at least some plans with generous coverage in the market.

The message of the social planner's solution is clear: there are not enough consumers in Part D PDP, and those that are in the market are systematically choosing plans that are too socially expensive. The social planner ameliorates this shortcoming by increasing subsidies on less-costly plans and increasing premiums for the more expensive plans.

6.5 Robustness Checks

6.5.1 Inertia

Our model of demand includes vintage of the plan as a characteristic in consumer's utility functions. As discussed previously, the interpretation of the coefficient on vintage as a structural parameter will only hold under restrictive circumstances. Rather, we view vintage as potentially capturing some inertia in the consumer market. If firms are aware of this inertia, the static Bertrand-Nash pricing game that we assume that they play will result in

biased estimates of their marginal costs. Recognizing this bias, we perform several robustness checks. Following Miller (2014), who finds that ignoring dynamic incentives can bias marginal costs by a factor of 20 percent, we repeat the key voucher counterfactuals, shifting marginal costs up and down 20 percent relative to our baseline estimates. The results of are reported in Figure 9. The primary result is that the optimal voucher remains at \$600 in both cases. As expected, the case with marginal costs lowered uniformly by 20 percent results in higher peak welfare, with the situation reversed when marginal costs are 20 percent higher. Consequently, while the levels of welfare estimates could change under a model with dynamic supply-side incentives, the qualitative conclusions that compare welfare outcomes across vouchers appear to be fairly robust.

6.5.2 Endogenous Outside Option

We also perform a second robustness check to determine the effect of our assumption that the outside option remains constant while we adjust the mechanism in Part D PDP. This may be unrealistic, as one might think that CMS, who administers both programs, would introduce reforms into both markets simultaneously. For example, a counterfactual mechanism that made Part D PDP much more attractive, say through an increase in the consumer subsidy, would give misleading results if the outside option could also be expected to become more attractive because similar subsidies were increased in MA-PD. Essentially, the basic problem is that the value of the outside option is moving with the inside option. One way of controlling for this is to let the value of the outside option move up and down with the difference in the subsidy in each counterfactual as calculated against the baseline subsidy. We implement a robustness check by imposing a penalty to the inside option such that the share of the outside option remains constant in the social planner's problem.[35] In the random utility framework, this is equivalent to holding the relative attractiveness of the outside option constant while we change the mechanism in MA-PD. We found that a penalty of \$36 to the inside options kept the outside option's share constant in the social planner's solution. The results that we obtained are reported in column (4) of Table 7.

The social planner's solution with the endogenous outside option gives lower overall welfare than with a fixed outside option. The intuition is straightforward: the endogenous option constrains the social planner's solution to lie in the space where the outside option is kept constant at 38 percent of the market. This constraint prevents the social planner

---

[35]This is algebraically equivalent to making the outside option more attractive, but keeps the utility interpretation consistent across all of other counterfactuals where we have held the utility of the outside option fixed at zero.

from implementing the first-best solution, as the social planner would ideally want to shift an additional 10 percent of the market into Part D PDP.

We also plot the differences in plan premiums and market shares, relative to the social planner's unrestricted problem, in Figure 10. The top panel shows that premiums have almost uniformly shifted upward, by an average of $87. The restriction that the outside option remain relatively equally attractive leads the social planner to reduce enrollments primarily in the lower social marginal cost plans, as the bottom panel shows.

The takeaway from this experiment is two-fold: one, the social planner wants to increase enrollment in Part D PDP over the current share, so restricting the outside option to be equally attractive constrains the social planner from maximizing utility; second, even the constrained solution increases total welfare by a factor of two compared to the observed allocation. Looking at the top panel of Figure 10 and comparing it to the top panel of Figure 8, we see that the constrained solution retains the qualitative aspects of the social planner's solution: premiums are increased for more socially-costly plans, and enrollments increase in less socially-costly plans. So while the level of utility changes, the basic underlying message of our story stays the same: the current mechanism subsidizes expensive plans too much, and does not subsidize inexpensive plans enough.

7   Conclusion

In this paper we have analyzed the welfare effects of the mechanism for determining subsidies for Prescription Drug Plans in Medicare Part D, focusing in particular on the supply side of the market. We draw several conclusions. First, we find that the current program is efficient only if we take into account that the government would have had to subsidize enrollees elsewhere as well. Without taking the latter into account, we could conclude that the program only generates a fraction of dollar value that is spent on it from the federal budget. This is due to a number of related factors: PDP plans are highly subsidized; demand for those plans is driven almost exclusively by said subsidies—consumers have very low willingness-to-pay for unsubsized plans, driven by the availability of inexpensive close substitutes; and firms are able to capture some of the rents of the subsidy mechanism.

On the supply-side we find, perhaps surprisingly, that the current structure of the program, where prices for distinct parts of the program, such as Medicare Advantage Prescription Drug coverage, Low Income Subsidies, and market premiums for regular beneficiaries, are tied together into one mechanism, in fact mutes insurers' ability to raise subsidies, and

hence positively affects total welfare. In fact, the current mechanism that weights multiple parts of the program into an average that is used to calculate subsidies, is similar in its incentives to a pre-determined optimal voucher mechanism. We find that providing flat vouchers that are optimally set ex ante could increase the total surplus in levels and relative to federal dollars spent, but not by a large amount (although a flat voucher mechanism could dramatically reduce the cost of administering the program, an effect that we do not include in our calculations). We further find that removing the averaging and just setting proportional subsidies would lead to a rapid upward price spiral, as the competitive pressure on the market—which we do find to have reasonably low margins—is, however, not strong enough to mitigate the "raising-the-subsidy" incentives.

Further, our analysis reveals a close connection between Part D and Medicare Advantage that, although not emphasized in the previous literature, proved to be crucial for our findings. We believe that our approach to the quantification of welfare that gradually removes interlinked parts of the environment - specifically, LIS bidding incentives and MA-PD part of the bid average - can be useful for many other public programs that do not exist in isolation, but, instead, are linked to other programs through the choices of consumers and producers or though government transfers.

While our institutional setting focused on the Medicare Part D program, our findings have broader implications for market design of privately-provided and publicly-subsidized social insurance programs. The motivation of subsidizing these programs is typically redistribution—the government attempts to ensure the affordability of insurance. Inevitably, such subsidy policies will have efficiency costs for the market. One source of such inefficiencies is market power. Subsidies create incentives for imperfectly competitive insurers to raise markups and pass them through to the price inelastic government. In general, we show in this paper that conditional on the decision to subsidize social insurance programs, there are large welfare differences across specific mechanisms that are feasibly at the policy maker's disposal. Depending on whether the policy is guided by the considerations of consumer surplus, total welfare, or government spending, different policies deliver drastically different results across these three measures of surplus. Overall, we argue that contrary to the focus of the literature on consumer choices in social insurance markets, the much less studied supply-side behavior in the presence of regulatory intervention and subsidization plays the key role in determining the efficiency outcomes of social insurance programs.

References

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

Abaluck, J. and J. Gruber (2013). Evolving Choice Inconsistencies in Choice of Prescription Drug Insurance. *NBER Working Paper no. 19163*.

Abaluck, J., J. Gruber, and A. Swanson (2015). Prescription Drug Use under Medicare Part D: A Linear Model of Nonlinear Budget Sets. *NBER Working Paper No. 20976*.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile Prices in Market Equilibrium. *Econometrica 63(4)*, 841–889.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics 25*(2), 242–262.

Center for Medicare and Medicaid Services (2009, February). Guidance to States on the Low-Income Subsidy. http://www.cms.gov/Medicare/Eligibility-and-Enrollment/LowIncSubMedicarePresCov/EligibilityforLowIncomeSubsidy.html. Accessed June 4, 2015.

Chorniy, A., D. Miller, and T. Tang (2014). Mergers in Medicare Part D: Decomposing Market Power, Cost Efficiencies, and Bargaining Power. *Clemson University Working Paper*.

Congressional Budget Office (2014, July). Compettition and the Cost of Medicare's Prescription Drug Program.

Currie, J. and F. Gahvari (2008). Transfers in Cash and In-Kind: Theory Meets the Data. *Journal of Economic Literature 46-2*, 333–83.

Curto, V., L. Einav, J. Levin, and J. Bhattacharya (2015). Can Health Insurance Competition Work? Evidence from Medicare Advantage. *NBER Working Paper No. 20818*.

Dalton, C., G. Gowrisankaran, and R. Town (2015). Myopia and Complex Dynamic Incentives: Evidence from Medicare Part D. *Working Paper*.

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

Dubé, J.-P., G. J. Hitsch, and P. E. Rossi (2009). Do switching costs make markets less competitive? *Journal of Marketing Research 46*(4), 435–445.

Duggan, M. (2004). Does Contracting Out Increase the Efficiency of Government Programs? Evidence from Medicaid HMOs. *Journal of Public Economics 88:12*, 2549–2572.

Duggan, M., P. Healy, and F. Scott Morton (2008). Providing Prescription Drug Coverage to the Elderly: America's Experiment with Medicare Part D. *Journal of Economic Perspectives 22(4)*, 69–92.

Duggan, M. and F. Scott Morton (2006). The Distortionary Effects of Government Procurement: Evidence from Medicaid Prescription Drug Purchasing. *The Quarterly Journal of Economics 121(1)*, 1–30.

Duggan, M. and F. Scott Morton (2010). The Effect of Medicare Part D on Pharmaceutical Prices and Utilization. *The American Economic Review 100*(1), 590–607.

Einav, L., A. Finkelstein, and P. Schrimpf (2015). The Response of Drug Expenditure to Non-Linear Contract Design: Evidence from Medicare Part D. *Quarterly Journal of Economics 130(2)*.

Engelhardt, G. V. and J. Gruber (2011). Medicare Part D and the Financial Protection of the Elderly. *American Economic Journal: Economic Policy 3*(4), 77–102.

Enthoven, A. C. (2011). Reforming Medicare by Reforming Incentives. *New England Journal of Medicine 364*(21), e44. PMID: 21561341.

Ericson, K. (2014). Consumer Inertia and Firm Pricing in the Medicare Part D Prescription Drug Insurance Exchange. *American Economic Journal: Economic Policy 6*(1), 38–64.

Frakt, A. (2011, March, 23). The key difference between exchanges and the Ryan-Rivlin plan. *Web Blog Post TheIncidentalEconomist*.

Gandhi, A. and J. Houde (2015). Measuring Substitution Patterns in Differentiated Products Industries The Missing Instruments. *Mimeo, University of Wisconsin-Madison and University of Pennsylvania*.

Gruber, J. and E. Washington (2005). Subsidies to Employee Health Insurance Premiums and the Health Insurance Market. *Journal of Health Economics 24*, 253–276.

Hausman, J. A. and J. M. Poterba (1987). Household behavior and the tax reform act of 1986. *The Journal of Economic Perspectives 1*(1), pp. 101–119.

Heiss, F., A. Leive, D. McFadden, and J. Winter (2013). Plan Selection in Medicare Part D: Evidence from Administrative Data. *Journal of Health Economics 32*, 1325–1344.

Heiss, F., D. Mcfadden, and J. Winter (2010). Mind the Gap ! Consumer Perceptions and Choices of Medicare Part D Prescription. In D. Wise (Ed.), *Research Findings in the Economics of Aging*, pp. 413–481. The University of Chicago Press.

Ho, K., J. Hogan, and F. Scott Morton (2015). Insurer (In)Efficiency in the Medicare Part D Program. *NBER Working Paper No. 21028*.

Kesternich, I., F. Heiss, D. McFadden, and J. Winter (2013). Suit the Action to the Word, the Word to the Action: Hypothetical Choices and Real Decisions in Medicare Part D. *Journal of Health Economics 32*(6).

Ketcham, J., C. Lucarelli, and C. Powers (2015). Paying Attention or Paying Too Much in Medicare Part D. *American Economic Review 105*(1), 204–233.

Ketcham, J. and K. Simon (2008). Medicare Part D's Effects on Elderly Drug Costs and Utilization. *American Journal of Managed Care*, 14–22.

Ketcham, J. D., C. Lucarelli, E. J. Miravete, and M. C. Roebuck (2012). Sinking, Swimming, or Learning to Swim in Medicare Part D. *The American Economic Review 102*(6), 2639–2673.

Klemperer, P. (1995). Competition when consumers have switching costs: An overview with applications to industrial organization, macroeconomics, and international trade. *The Review of Economic Studies 62*(4), 515–539.

Kling, J. R., S. Mullainathan, E. Shafir, L. C. Vermeulen, and M. V. Wrobel (2012). Comparison Friction: Experimental Evidence from Medicare Drug Plans. *The Quarterly Journal of Economics 127*(1), 199–235.

Laffont, J.-J. and J. Tirole (1993). *A Theory of Incentives in Procurement and Regulation*. The MIT Press.

Lucarelli, C., J. Prince, and K. Simon (2012). The Welfare Impact of Reducing Choice in Medicare Part D: A Comparison of Two Regulation Strategies. *International Economic Review 53*(4), 1155–1177.

Miller, D. and J. Yeo (2014a). The Consequences of a Public Health Insurance Option: Evidence From Medicare Part D. *Mimeo, Clemson University and Singapore Management University*.

Miller, D. P. (2015). Risk Adjustment and Low Income Subsidy Distortions in Medicare Part D. *Clemson University Working Paper*.

Miller, D. P. and J. Yeo (2014b). A Dynamic Discrete Choice Model of Switching Costs in Medicare Part D. *Mimeo, Clemson University and Singapore Management University*.

Miller, K. (2014). Do Private Medicare Firms Have Lower Costs? *Mimeo, University of Minnesota*.

Nevo, A. (2001). Measuring Market Power in the Ready-to-Eat Cereal Industry. *Econometrica 69*, 307–342.

Oliver, T. R., P. R. Lee, and H. L. Lipton (2004). A Political History of Medicare and Prescription Drug Coverage. *Milbank Quarterly 82*(2), 283–354.

Petrin, A. (2002). Quantifying the Benefits of New Products: The Case of the Minivan. *Journal of Political Economy 110*(4), 705–729.

Polyakova, M. (2015). Regulation of insurance with adverse selection and switching costs: Evidence from Medicare Part D. *Mimeo, Stanford University*.

Small, K. A. and H. S. Rosen (1981). Applied welfare economics with discrete choice models. *Econometrica 49*(1), pp. 105–130.

Town, R. and S. Liu (2003). The welfare impact of Medicare HMOs. *RAND Journal of Economics 34*(4), 719–736.

Train, K. and C. Winston (2007). Vehicle Choice Behavior and the Declining Market Share of U.S. Automakers. *International Economic Review 48*, 1469–1496.

Vetter, S., F. Heiss, D. McFadden, and J. Winter (2013). Risk attitudes and Medicare Part D enrollment decisions. *Economics Letters 119*, 128–132.

Williams, H. (1977). On the formation of travel demand models and economic evaluation measure of user benefit. *Environment and Planning A 9*(3), 285–344.

Winter, J., R. Balza, F. Caro, F. Heiss, B. Jun, R. Matzkin, and D. McFadden (2006). Medicare prescription drug coverage: Consumer information and preferences. *Proceedings of the National Academy of Sciences of the United States of America 103(20)*, 7929–7934.

Yin, W. and D. Lakdawalla (2015, May). Insurers' Negotiating Leverage and the External Effect of Medicare Part D. *Review of Economics and Statistics 97*(2), 314–331.
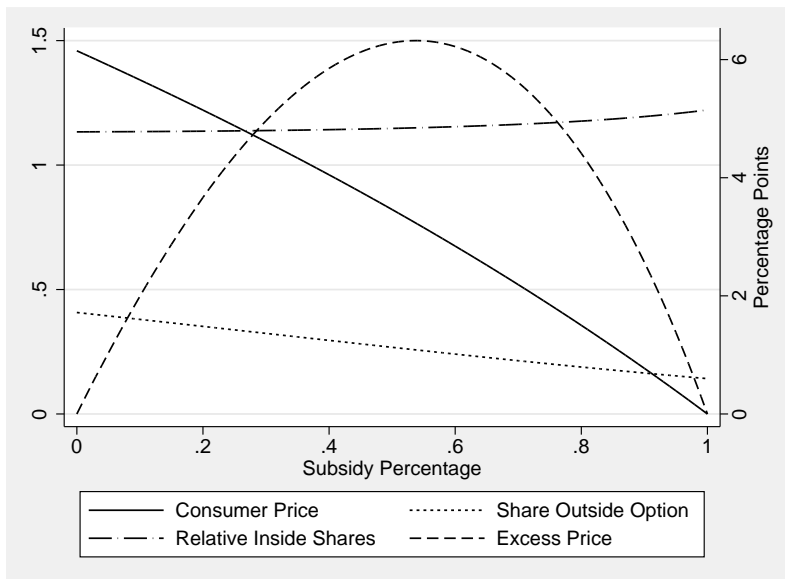
Figure 1: A Simple Model of Proportional Subsidies
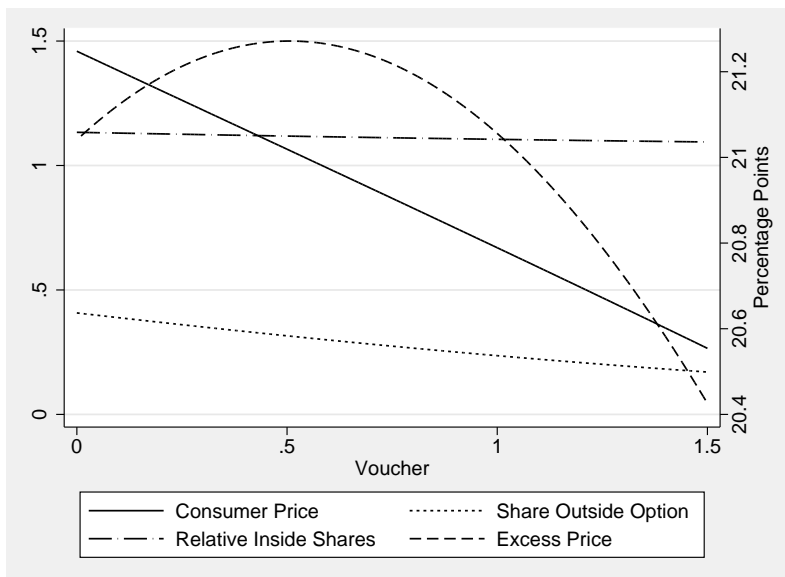

Figure 2: A Simple Model of Vouchers
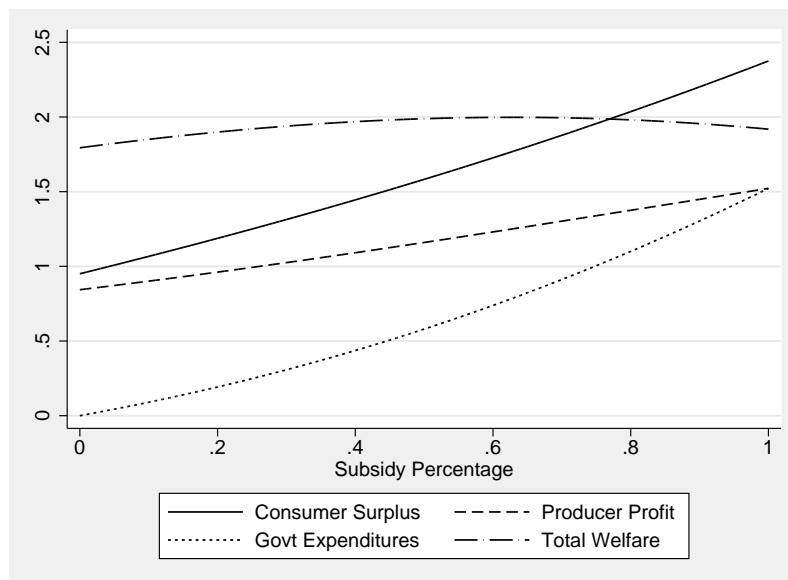

49

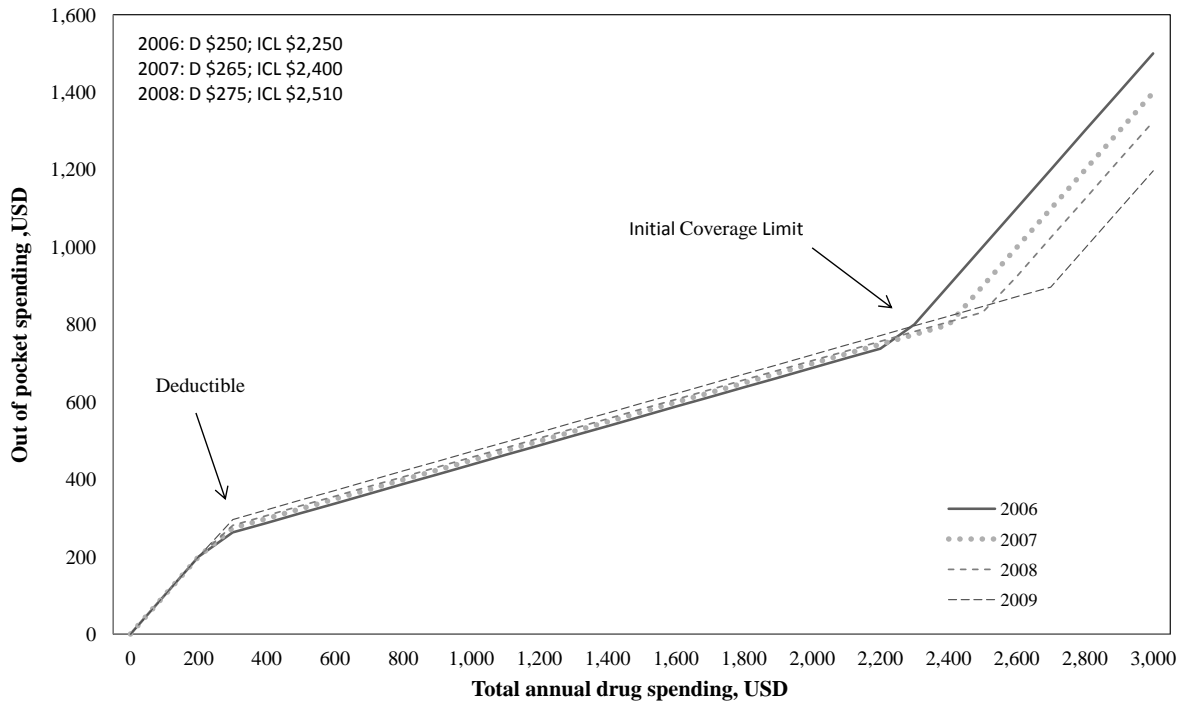Figure 3: Welfare Decomposition of Proportional Subsidies

Figure 4: Minimum Coverage Requirements in Part D

Notes: Insurers in the Medicare Part D program are required to provide coverage that gives at least the same actuarial value as the Standard Defined Benefit (SDB). The SDB design features a deductible, a co-insurance rate of 25 percent up to the initial coverage limit (ICL) and the subsequent "donut hole" that has a 100 percent co-insurance until the individual reaches the catastrophic coverage arm of the contract. The graph illustrates these features of the SDB by mapping the total annual drug spending into the out-of-pocket expenditure. As the figure illustrates, the generosity of the SDB changed over time.

Figure 5: Contract Prices and Market Structure
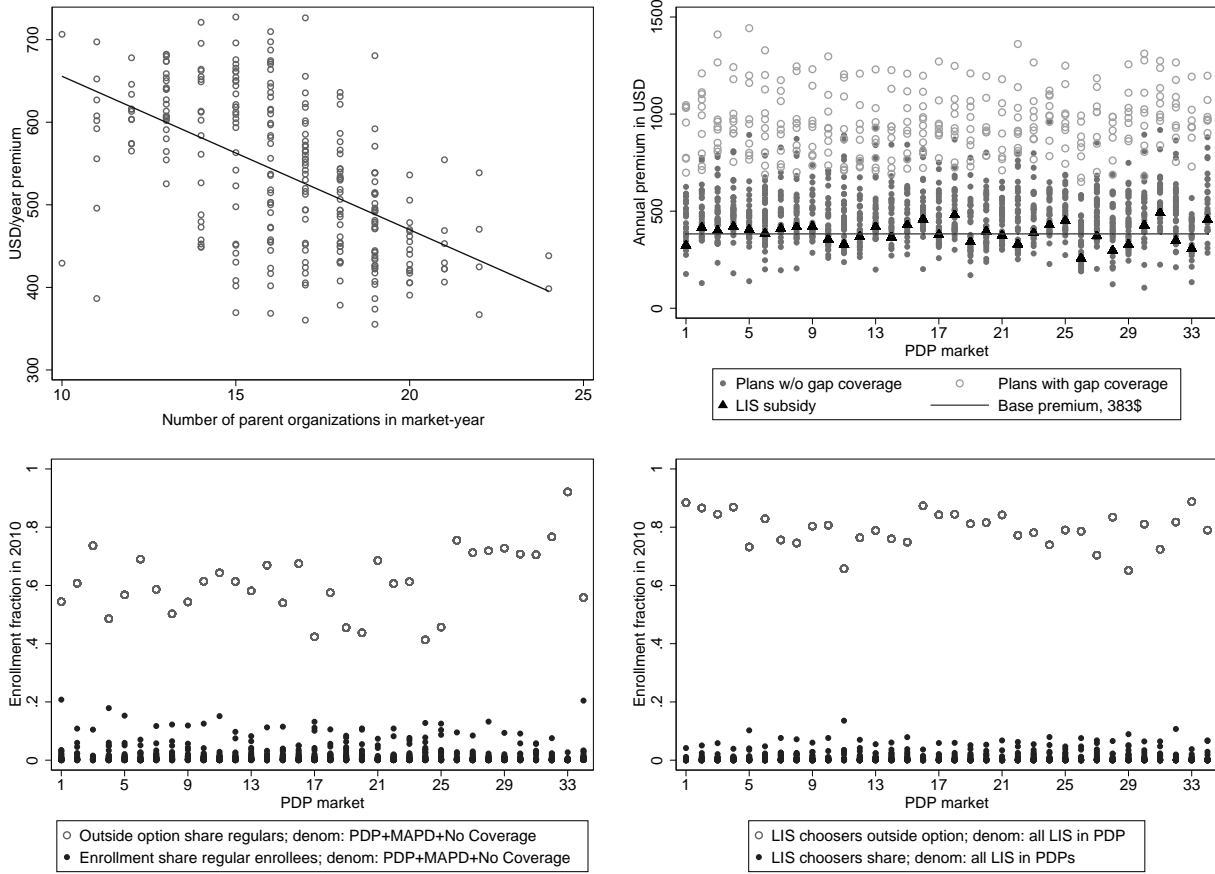
Notes: The top-left panel shows the correlation between the number of competing insurers and average premiums in the market for 2006-2010. The premiums are not enrollment-weighted. The top-right panel depicts cross-sectional variation in premiums, in the levels of low income subsidies, and in "base premiums" in 2010. Plans with premiums below LIPSA are eligible for the random assignment of LIS beneficiaries. The bottom-left and bottom-right panels show the distribution of market shares of the outside and inside options across 34 Medicare Part D regions for regular and LIS-eligible enrollees, respectively. We restrict the definition of the market to include PDP options, MA-PDs and no coverage choices. The latter two comprise the outside option.

Figure 6: Estimated Marginal Costs and Markups

Demand model: BLP

MC inverted for undistorted contracts
MC predicted for distorted contracts

Demand model: BLP

Mark-ups reported for non-distorted plans only
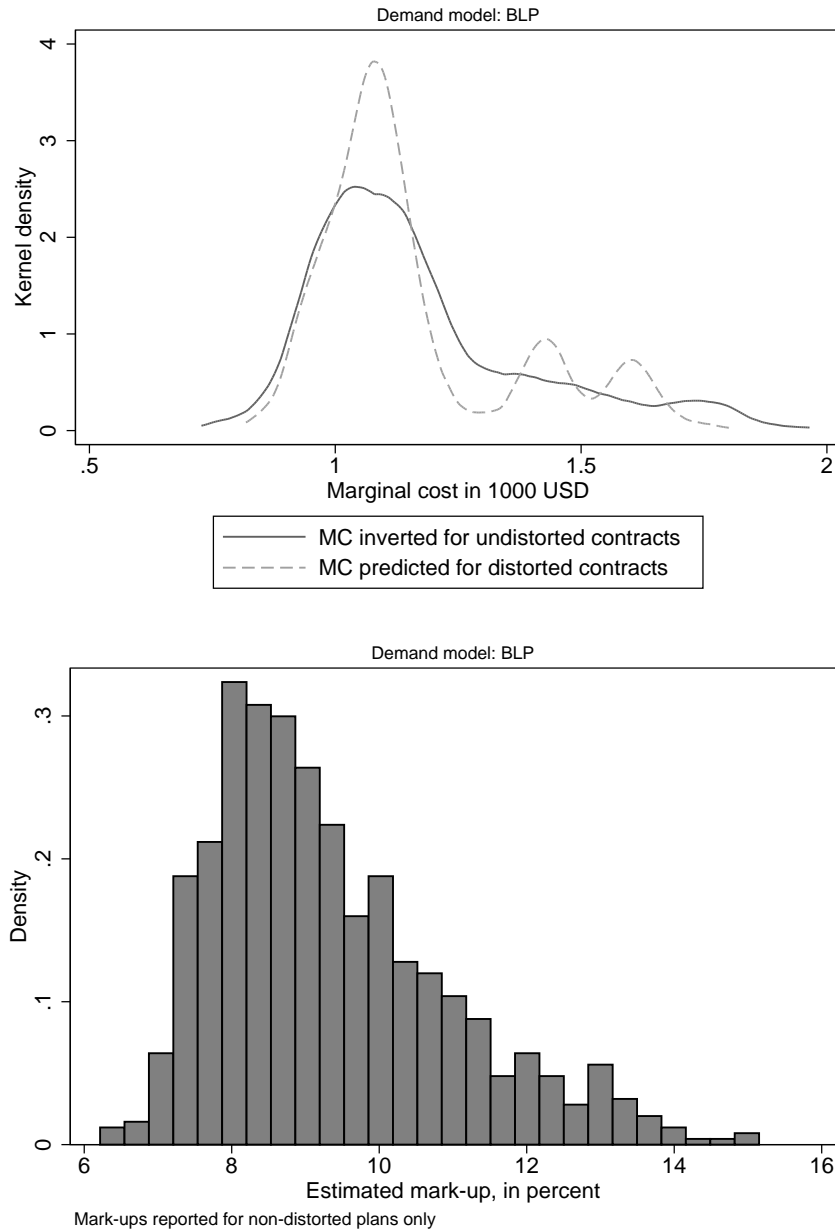
Notes: The top panel plots the marginal cost estimates from inversion (for undistorted contracts) and from hedonic projection (for distorted contracts). Plan characteristics used in the hedonic projection include deductible, coverage in the gap and enhanced plan indicators, measures of formulary generosity, pharmacy networks, vintage, as well as estimated unobserved plan quality, and region and insurer fixed effects. The bottom panel shows the mark-ups computed for non-distorted contracts in 2010, using marginal cost estimates from the inversion of insurers' first-order conditions. The mark-up is computed as price net of marginal cost over marginal cost.

Figure 7: Welfare with Flat Subsidies: Uniform and Market-Specific Optimal Vouchers
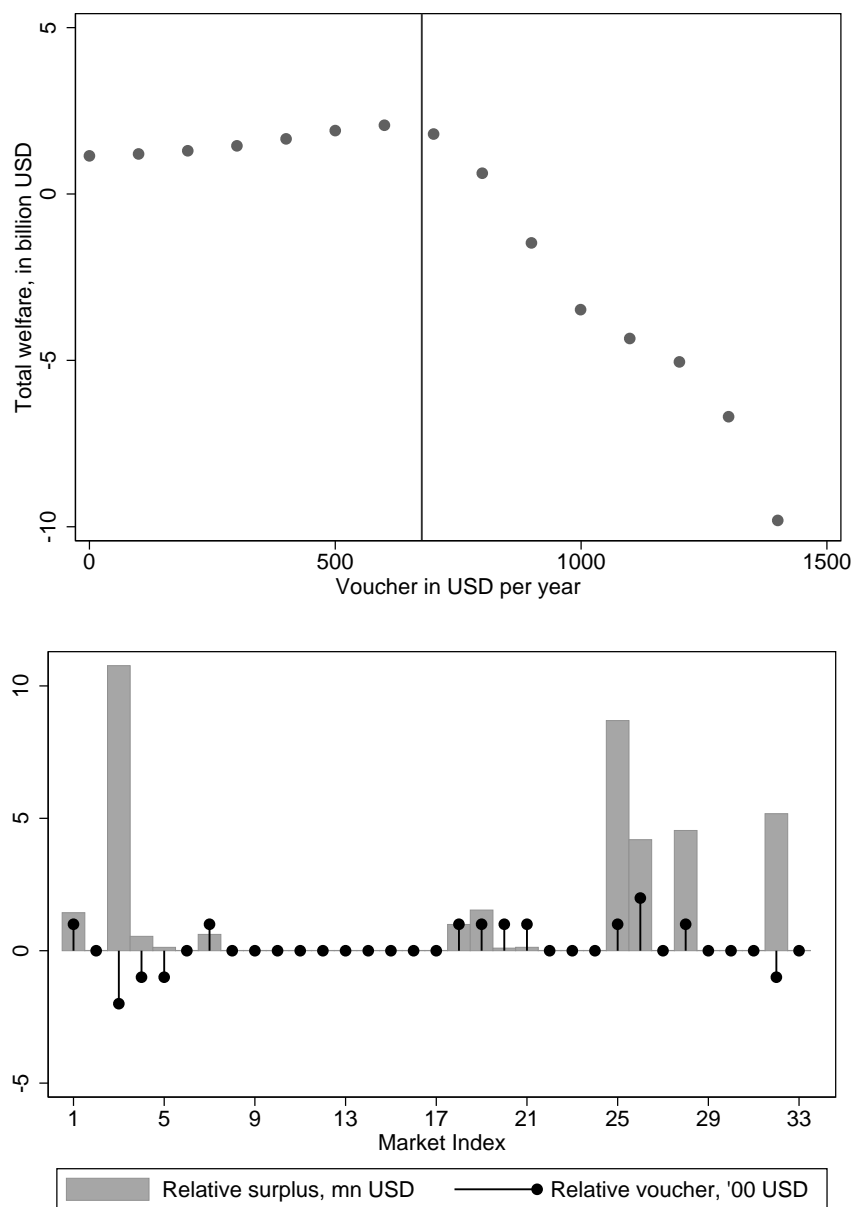
Notes: The top panel plots estimated total welfare in full equilibrium counterfactuals for vouchers in $100 increments. The optimal uniform voucher is $600. In the bottom panel, we calculate optimal vouchers for each market and plot the difference between the market-specific optimal voucher and optimal uniform voucher. We also record the extra welfare that would be gained in each region by implementing the market-specific optimal voucher rather than the uniform voucher.

Figure 8: Social Planner's Solution: Changes in Premiums and Market Shares
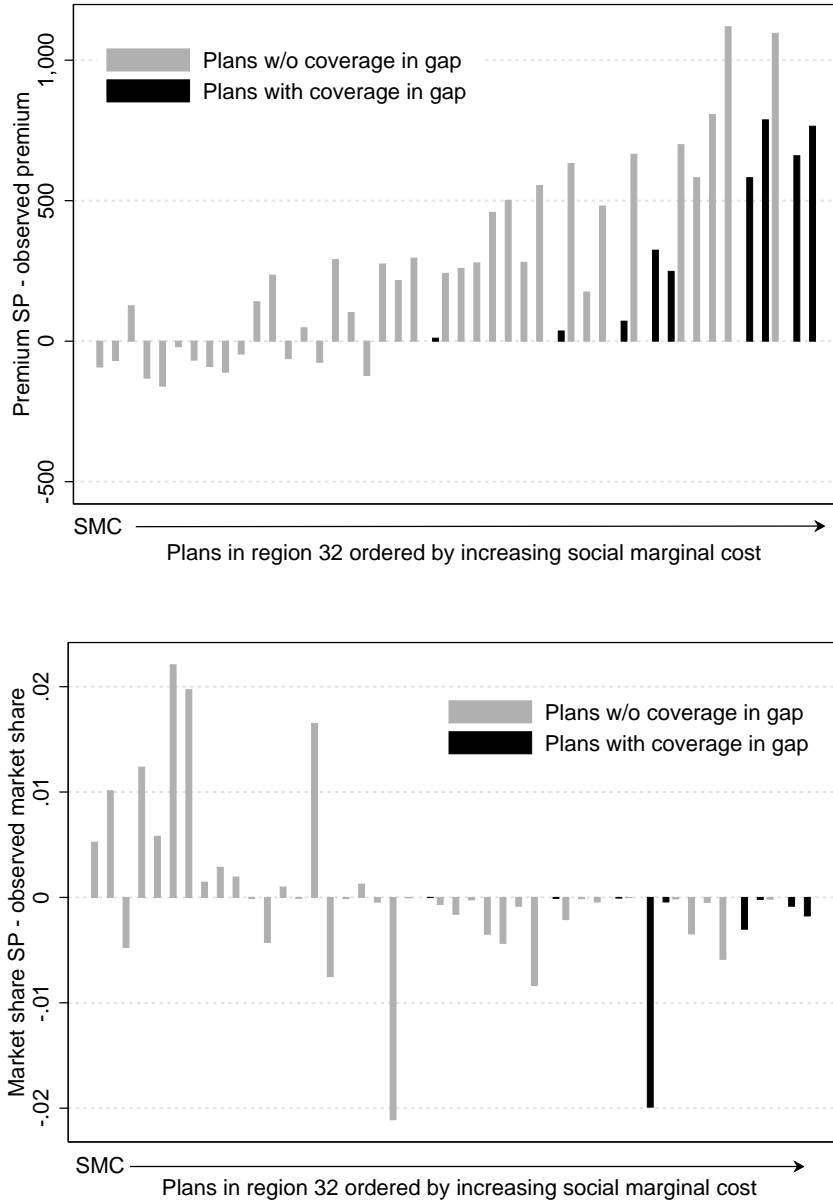
Notes: The figure shows the change in premiums (top panel) and market shares (bottom panel) in California in 2010 for the social planner as compared to the data. Contracts are sorted by their estimated social marginal costs.

Table 1: Summary Statistics

Year	2007	2008	2009	2010
Plans				
Total number of PDP plans	1,742	1,791	1,674	1,565
Average number of PDP plans per market	51	53	49	46
Firms				
Total number of PDP parent organizations	56	56	50	50
Average number of PDP parent organizations per market	19	19	17	17
Premiums				
Unweighted average annual PDP consumer premium	$439	$477	$545	$559
Subsidies				
CMS national average bid (annual)	$965	$966	$1,012	$1,060
CMS base consumer premium (annual)	$328	$335	$364	$383
CMS subsidy for average risk beneficiary	$637	$631	$648	$677
Low income benchmark threshold	$341	$333	$353	$388
Enrollment				
All Part D Eligible (in millions)	43.3	44.4	45.5	46.6
PDP enrollment, regular (in millions)	8.3	8.6	8.9	9.4
PDP enrollment, low-income (in millions)	8.2	8.2	8.2	8.4
Total MAPD enrollment (in millions)	7.5	8.6	9.4	9.8
Employer sponsored coverage RDS (in millions)	7.0	6.6	6.5	6.7
Other coverage sources (in millions)	5.7	5.9	6.0	6.1
No creditable coverage (in millions)	6.6	6.5	6.4	6.2

Table 2: Demand Estimates

	Regular Enrollees			LIS Enrollees	
	OLS Logit (1)	IV Logit (2)	BLP (3)	OLS Logit (4)	IV Logit (5)
Premium	−2.74	−10.44		−2.08	−7.58
	(0.15)	(1.09)		(0.09)	(1.27)
Mean, Premium (α)			2.58		
			(0.36)		
Std. Deviation, Premium (σ_α)			0.27		
			(0.09)		
Std. Deviation, Inner Options (σ_{inner})			0.00		
			(0.08)		
Annual Deductible	−3.25	−6.72	−7.11		
	(0.21)	(0.56)	(1.52)		
Gap Coverage	0.18	2.93	2.89		
	(0.07)	(0.40)	(0.94)		
Std. Deviation, Gap Coverage (σ_{gap})			0.00		
			(0.08)		
Number of Top Drugs Covered	0.24	31.58	30.4	−3.05	10.63
	(4.36)	(7.30)	(12.1)	(4.62)	(7.15)
Pharmacy Network Measure	0.31	0.29	0.31	0.23	0.22
	(0.04)	(0.06)	(0.01)	(0.05)	(0.08)
Number Years Plan on Market	0.61	0.88	0.90	0.76	0.88
	(0.02)	(0.05)	(0.12)	(0.3)	(0.04)

Notes: Standard errors are reported in parentheses. The sample is all Medicare Part D stand-alone prescription drug plans from 2007 to 2010; the sample size is 6,675 plan-year observations for regular enrollees and 4,561 plan-year observations for LIS enrollees. In addition to the displayed coefficients and fixed effects, all models also include a constant and the following plan characteristics: a dummy for an enhanced plan, the number of APIs in the plan's formulary, and the number of drugs placed in Tiers 1-2 of the formulary which have low cost-sharing. Columns (2), (3), and (5) use the set of instrumental variables described in the text.

Table 3: First Stage of Instrumental Variable Regression for Plan Premiums

	Regular FS (1)	LIS FS (2)
Number PDP Plans in a Region-year by Same Firm	−9.63	−3.32
	(2.38)	(4.35)
Number MA plans in a Region-year by Same Firm	−0.18	0.03
	(0.24)	(0.47)
Deductible of MA Plans in the Same Region-year	−0.15	0.03
	(0.17)	(0.34)
Same-Firm Average Plan Prices in Other Markets	0.37	0.28
	(0.03)	(0.06)
Number of Observations	6,675	4,561

Notes: Standard errors in reported in parentheses. Each regression uses data on Medicare Part D stand-alone prescription drug plans in years 2007 to 2010. The first column reports the first stage for the regular premiums, while the second column reports the first stage for low-income-subsidy adjusted premiums. See the text for more details on the construction of the instruments. The regressions also include all plan characteristics that are used in demand estimation, including a constant and fixed effects for geographic markets, parent organizations, and years.

Table 4: Marginal Cost Projection

	Berry MC inversion (1)	BLP MC inversion (2)
Annual deductible	−0.37	−0.36
	(0.04)	(0.04)
No. of common APIs	0.14	0.15
	(0.13)	(0.13)
Has coverage in the gap (1/0)	0.42	0.41
	(0.01)	(0.01)
Enhanced plan (1/0)	−0.04	−0.03
	(0.01)	(0.01)
No. of top drugs in Tier 1 and 2	−0.57	−0.54
	(0.38)	(0.37)
No. of top drugs covered	−8.70	−8.69
	(2.48)	(2.41)
Pharmacy network measure	−0.19	−0.19
	(0.05)	(0.05)
Number of years the plan is on the market	0.05	0.05
	(< 0.01)	(< 0.01)
R-squared	0.87	0.86

Number of observations = 756. For the BLP model, we use the Berry IV specification for the LIS part of the market in the inversion procedure. Fixed effects for markets and insurers are included in the regression but not reported.

Table 5: Observed Allocation; Link and Ownership Counterfactuals

	Data (1)	No LIS Link (2)	No LIS or MA-PD Link (3)	Independent Plans (4)	Monopoly Ownership (5)
Consumer Surplus, $M	2,517	2,892	3,413	3,350	2,814
Insurer Profit, $M	459	1,063	1,279	1,172	2,161
Premium Subsidy, $M	5,936	7,807	10,125	9,742	8,884
Reinsurance Subsidy, $M	3,444	4,096	5,026	4,993	4,107
Inside Option Enrollment, 000's	8,772	10,565	12,948	12,640	10,592
Inside Option, Percent	38	45	55	54	45
Average Weighted Premium, $	502	471	418	426	449
Average Weighted Bid, $	1,123	1,152	1,144	1,141	1,240
MA-PD Premium Subsidy, $M	6,018	7,248	8,882	8,671	7,266
MA-PD Reinsurance Subsidy, $M	2,281	2,747	3,366	3,286	2,754
Total surplus, $M	**1,570**	**1,474**	**918**	**912**	**1,114**
Per PDP Subsidy Dollar, $	0.17	0.12	0.06	0.06	0.09
Per Opportunity Cost Dollar, $	1.12	0.59	0.24	0.25	0.29

Notes: Total surplus is the sum of consumer surplus and producer profit minus the social opportunity cost of government spending. MA-PD premium and reinsurance subsidies are the amount of government expenditures if all PDP enrollees switched to MA-PD. Welfare per opportunity cost dollar is only reported when government expenditures in PDP are higher than in MA-PD.

Table 6: Proportional and Flat Subsidy Counterfactuals

	Proportional Subsidies			Vouchers		
	5 Percent (1)	32 Percent (2)	95 Percent (3)	$0 (4)	$721 (5)	$1,340 (6)
Consumer Surplus, $M	13,344	4,536	1,081	1,073	2,844	13,028
Insurer Profit, $M	50,403	4,846	20	13	960	4,272
Premium Subsidy, $M	84,144	18,087	10	-	7,301	31,290
Reinsurance Subsidy, $M	15,695	8,191	59	39	3,932	9,236
Inside Option Enrollment, 000's	23,280	14,506	134	90	10,127	23,351
Inside Option, Percent	100	62	1	0	43	100
Average Weighted Premium, $	190	587	1,381	1,439	501	34
Average Weighted Bid, $	3,805	1,834	1,453	1,439	1,222	1,374
MA-PD Premium Subsidy, $M	15,970	9,951	92	62	6,947	16,019
MA-PD Reinsurance Subsidy, $M	6,053	3,772	35	23	2,633	6,071
Total surplus, $M	**(37,414)**	**(6,939)**	**1,177**	**1,147**	**1,654**	**(6,667)**
Per PDP Subsidy Dollar, $	(0.37)	(0.26)	17.17	29.38	0.15	(0.16)
Per Opportunity Cost Dollar, $	(0.37)	-	-	-	0.77	(0.28)

Notes: Total surplus is the sum of consumer surplus and producer profit minus the social opportunity cost of government spending. MA-PD premium and reinsurance subsidies are the amount of government expenditures if all PDP enrollees switched to MA-PD. Welfare per opportunity cost dollar is only reported when government expenditures in PDP are higher than in MA-PD.

Table 7: Marginal Cost, Social Planner, and Optimal Voucher Counterfactuals

	Marginal Cost		Social Planner		Optimal Voucher	
	Private (1)	Social (2)	Baseline (3)	MA Adjusted (4)	National (5)	Market (6)
Consumer Surplus, $M	1,094	1,064	2,937	1,942	1,824	1,912
Insurer Profit, $M	-	-	(10,055)	(7,148)	516	547
Premium Subsidy, $M	-	-	-	-	2,870	3,235
Reinsurance Subsidy, $M	93	-	-	-	1,866	1,894
Inside Option Enrollment, 000's	217	35	11,192	8,786	4,783	5,136
Inside Option, Percent	1	0	48	38	20	22
Average Weighted Premium, $	1,282	1,450	372	459	648	608
Average Weighted Bid, $	1,282	1,450	372	459	1,248	1,238
MA-PD Premium Subsidy, $M	149	24	7,677	5,711	3,281	3,523
MA-PD Reinsurance Subsidy, $M	57	9	2,910	2,284	1,244	1,335
Total surplus, $M	**1,241**	**1,107**	**3,628**	**3,044**	**2,066**	**2,106**
Per PDP Subsidy Dollar, $	13.30	-	0.28	0.33	0.44	0.41
Per Opportunity Cost Dollar, $	-	-	-	-	7.54	5.97

Notes: Total surplus is the sum of consumer surplus and producer profit minus the social opportunity cost of government spending. MA-PD premium and reinsurance subsidies are the amount of government expenditures if all PDP enrollees switched to MA-PD. Welfare per opportunity cost dollar is only reported when government expenditures in PDP are higher than in MA-PD.

8 Appendix

8.1 Derivation of Social Planner's Pricing Problem

Social planner's problem is to maximize the following welfare function:

$$
W^{SP}(p) = \overbrace{\int \frac{1}{\alpha} \left(\gamma + \ln \left[1 + \sum_{j=1}^{J} \exp(v_j(\theta), p_j) \right] \right) dF(\theta)}^{\text{Consumer Surplus (CS)}} +
$$

$$
\lambda \left[\overbrace{\sum_{j=1}^{J} (p_j - c_j) s_j(p)}^{\text{Product Market Surplus}} - \left(\overbrace{\sum_{j=1}^{J} (G_j^{PDP} - G^{MAPD}) s_j(p)}^{\text{Cost of Subsidies (G)}} \right) \right]. \quad (18)
$$

The social planner's solution is defined by the set of first-order conditions obtained by differentiating $W^{SP}(p)$ with respect to prices. The derivative of consumer surplus with respect to p_j has a conveniently simple form after some simplifications:

$$
\frac{\partial CS(p)}{\partial p_j} = \int \frac{1}{\alpha} \left[\frac{-\alpha \exp(v_j(\theta))}{1 + \sum_{k=1}^{J} \exp(v_k(\theta))} \right] dF(\theta) = -s_j(p). \quad (19)
$$

The derivative of product market profit with respect to p_j is:

$$
\frac{\partial \Pi(p)}{\partial p_j} = \lambda s_j(p) + \lambda \sum_k (p_k - c_k) \frac{\partial s_k(p)}{\partial p_j}. \quad (20)
$$

The derivative of government spending with respect to p_j is:

$$
\frac{\partial GS(p)}{\partial p_j} = -\lambda \left[\sum_k (G_k^{PDP} - G^{MAPD}) \frac{\partial s_k(p)}{\partial p_j} \right], \quad (21)
$$

$$
= -\lambda \left[\sum_k \Delta G_k \frac{\partial s_k(p)}{\partial p_j} \right]. \quad (22)
$$

Summing these terms, we obtain:

$$\frac{\partial W(p)}{\partial p_j} = (\lambda - 1)s_j + \lambda \sum_k (p_k - c_k)\frac{\partial s_k(p)}{\partial p_j} - \lambda \left[\sum_k \Delta G_k \frac{\partial s_k(p)}{\partial p_j} \right], \qquad (23)$$

$$= (\lambda - 1)s_j + \lambda \sum_k (p_k - c_k - \Delta G_k)\frac{\partial s_k(p)}{\partial p_j}. \qquad (24)$$

Note that a decrease in consumer surplus in response to an increased price $(-s_j(p))$ is offset, up to the cost of transferring public funds, by an increase in profit in the product market $(\lambda s_j(p))$.

The first-order conditions can be expressed in a particularly simple formula in vector notation; the set of equations defining the social planner's solution is:

$$(\lambda - 1)s(p) + \lambda \Omega(p)(p - c - \Delta G) = 0, \qquad (25)$$

where $\Omega(p)$ is a matrix of partial derivatives such that the element in the i-th row and j-th column is:

$$\Omega_{ij}(p) = \frac{\partial s_j(p)}{\partial p_i}. \qquad (26)$$

Figure 9: Robustness Check: Uniform Optimal Vouchers with Alternative Marginal Costs
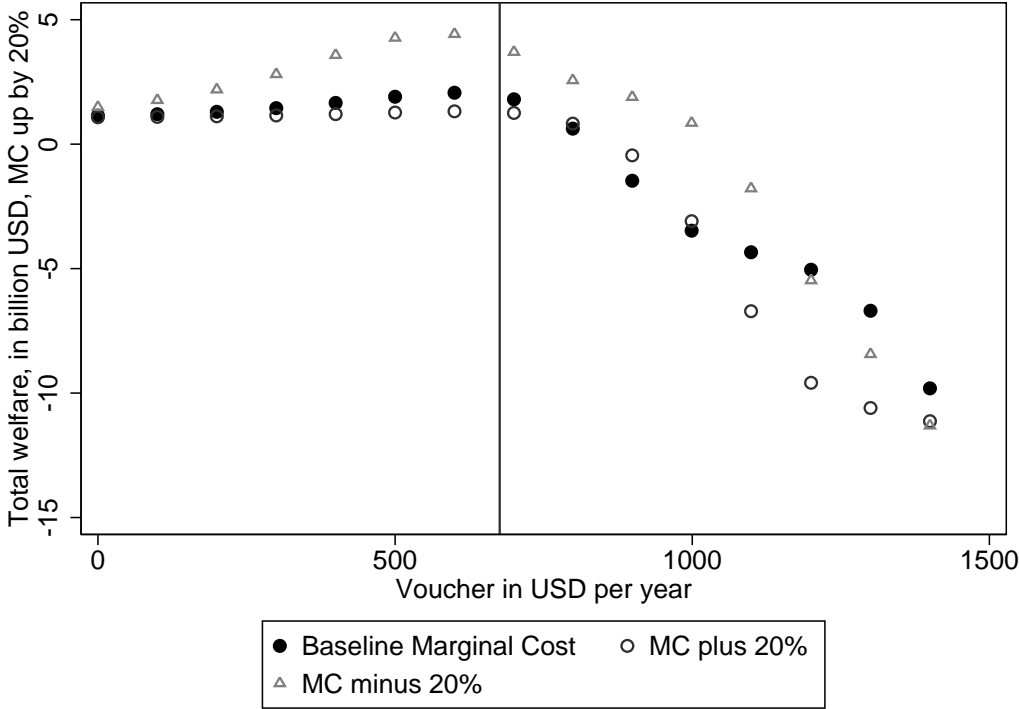
Notes: The figure plots the estimated total welfare in full equilibrium counterfactuals under vouchers of various levels across three estimates of marginal costs: the baseline, baseline plus 20 percent, and baseline minus 20 percent.

Figure 10: Robustness Check: Social Planner Solution with Endogenous Outside Option
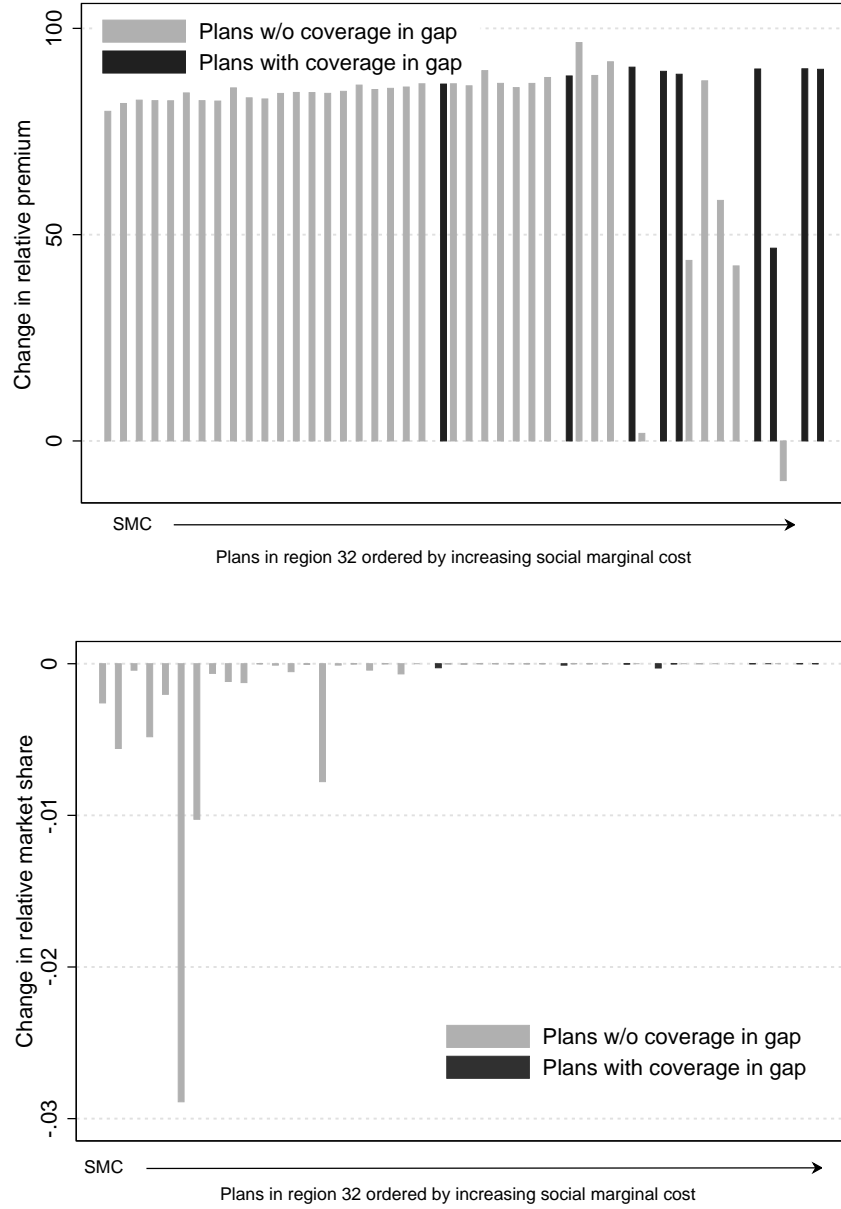
Notes: The figure plots the differences in premiums (top panel) and market shares (bottom panel) between the social planner with and without the endogenous outside option for plans in California in 2010.