

# The Welfare Effects of Supply-Side Regulations in Medicare Part D

Francesco Decarolis, Maria Polyakova, Stephen P. Ryan\*

June 3, 2015

## Abstract

We study the efficiency of the regulatory mechanism through which the government currently administers federal subsidies in Medicare Part D, a large insurance exchange market for seniors. Two economic tensions are important for thinking about subsidization policies in publicly subsidized, privately-administered environments. First, subsidies may lead consumers to select inefficient amounts of coverage by substituting across insurance options with difference relative subsidies. Second, under imperfect competition, which typically characterizes health insurance settings, the exact design of the subsidy mechanism may substantially affect insurers' pricing incentives, and may thus be crucial for the allocating efficiency of the market. In the case of Medicare Part D, we find inefficient sorting of seniors across plans both on the extensive margin - subsidies make stand-alone Part D plans more attractive relative to other insurance sources; and on the intensive margin - subsidies create some distortions in which contracts are purchased. On the supply side, we find that the current subsidization policy uses a reasonably effective decentralized instrument - the current mechanism achieves a level of total welfare close to that obtained under an optimal voucher scheme; however, the decentralized solution is far from the social planner's first-best.

JEL: I11, I18, L22, D44, H57

**Keywords:** Medicare; Prescription Drugs; Health Insurance; Subsidies; Regulation

---

\*Decarolis, Boston University ([fdc@bu.edu](mailto:fdc@bu.edu)); Polyakova, Stanford University ([mpolyak@stanford.edu](mailto:mpolyak@stanford.edu)); Ryan, University of Texas at Austin and NBER ([sryan@utexas.edu](mailto:sryan@utexas.edu)). Decarolis is grateful to the Sloan Foundation (grant 2011-5-23 ECON) for financial support. We also gratefully acknowledge support from the NSF (SES-1357705).

# 1 Introduction

Subsidizing the private provision of health insurance constitutes a large and growing fraction of government's expenditures in healthcare. Most 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 amount to more than \$500 billion in 2014.<sup>1</sup> This development substantially changes how the government spends money in healthcare - from direct reimbursement of physician and hospital services to 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 an expensive program, with federal spending totaling more than \$76 billion annually.

We focus our study on the supply-side of the Part D market. 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 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. In this paper we attempt to assess these distortions.

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 the typical setting due to the presence of two groups of consumers: so-called regular enrollees and low-income (LIS) enrollees. Regular enrollees make unrestricted choices from all plans

---

<sup>1</sup>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% or \$492 billion.

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.<sup>2</sup> Using four years of data on the characteristics and enrollments of all Part D PDP plans across all 34 Medicare Part D markets,<sup>3</sup> we estimate demand for both regular and LIS enrollees using the random coefficients discrete choice framework pioneered by [Berry, Levinsohn, and Pakes \(1995\)](#).

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, 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 which 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 bid below the average bid in their market are eligible for random assignment of LIS enrollees. Consequently, there is 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 low-income enrollees and the prices and choices of regular enrollees.

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. 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

---

<sup>2</sup>As of 2011, about one-third of LIS enrollees had opted out of the random assignment system.

<sup>3</sup>Medicare combines smaller states into the same so-called Medicare Part D regions.

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). Indeed, the evidence from the consumer level data indicates that of all PDP enrollees that switch plans, two thirds select another PDP, but one third move into 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% of the PDP 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 33 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 \$2.22 per dollar of opportunity cost. This is one of our primary findings; the positive welfare effect of Part D PDP is driven exclusively by opportunity costs. On its own merits, the total cost of providing subsidized goods exceeds their benefits; expenditures of \$9.4 billion generated \$4.0 billion of consumer surplus and \$529 million of producer profit. However, we estimate that foregone 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 \$3.12 billion in surplus.

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. In this counterfactual, we re-simulate the current subsidy mechanism without the influence of LIS enrollees.<sup>4</sup> 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. As we

---

<sup>4</sup>This simulation also removes the enrollment weights from the MA-PD market in determining the base subsidy; details are provided below.

are unable to compute counterfactual equilibria for mechanisms that have the LIS enrollees as part of the market, we consider this our baseline counterfactual.

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 in both situations. Under a monopoly, the loss is driven by decline in product market surplus, dominating the increase in producer profits. Under atomistic competition, the changes are less dramatic, but still result in negative welfare as the marginal benefits of serving additional consumers are exceeded by the social costs of providing the goods.

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, prices are equal to social marginal costs. In the last, 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 large 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 private marginal costs. The situation becomes even more extreme under social marginal costs, which incorporate the fact that the government has expenditures on plans that are unrelated to the subsidy directly. In this case, enrollment decline to only five percent of the market.

Interestingly, the social marginal cost counterfactual has lower welfare than the private marginal cost mechanism. The reason is that both mechanisms are ignoring an important component of welfare, which is 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 \$5.3 billion. This is approximately 70 percent higher than the current mechanism. Enrollment in Part D PDP under the social planner is nearly 50 percent of the market. Consumer surplus is nearly identical to the observed mechanism, but the distribution of equilibrium prices is completely different. Average prices are lower than all other mechanisms that we consider; the social planner prices where straight producer profits are negative.

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. 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 which 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. 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 explored 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). Einav et al. (2015) explore 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) Ericson (2014); Miller and Yeo (2014); 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.

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 (Duggan, 2004; 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 D.P. (2014) 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  $\delta_j$  is a measure of desirability of the products. The outside option gives zero utility. The idiosyncratic taste shock,  $\epsilon_{ij}$ , 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,  $\lambda$ , 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 considerations. 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 state of residence. Beneficiaries in traditional Medicare that are eligible for additional low-income subsidies (LIS), on the other hand, are automatically assigned to plans by CMS; these individuals can subsequently change their random assignment by making an active choice. The latter group is known as “LIS

choosers.”<sup>5</sup> 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. Unlike 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, the participating insurance companies are highly regulated, and continuously audited by Medicare. An important regulatory instrument is the annually set Standard Defined Benefit (SDB), 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% co-insurance rate and the infamous donut hole, which is a gap in coverage at higher spending levels. As long as actuarial minimum is satisfied, insurers are allowed to adjust and/or top up the SDB contract design, which generates empirical variation in contracts’ financial characteristics. In addition, contracts may be differentiated horizontally on the quality of insurer’s pharmacy networks, lists of covered drugs, 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.<sup>6</sup> For individuals, who are eligible for low-income-subsidies, these subsidies can go up to 100 percent. The intricate policies governing the program’s subsidy system are the focus of our paper. Below, we briefly outline the details of the subsidy mechanism for regular and LIS enrollees.

To determine the level of subsidies, 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 they want to offer. Similar plans in different markets count as separate plans. Part D program is divided into 34 geographic markets, some of which follow state boundaries, and some combine the states with smaller populations. 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.<sup>7</sup> Medicare takes the bids submitted by insurers for each of their plans and

---

<sup>5</sup>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.

<sup>6</sup>See Table IV.B11 of 2012 Trustees of Medicare Annual Report.

<sup>7</sup>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

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 plans nationwide, weights them by lagged enrollment shares of the plans and takes an 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 consumer’s premium.<sup>8</sup> In our counterfactual analyses we explore welfare properties of this subsidization mechanism, asking whether simple adjustments to the mechanism could improve the efficiency of the program.

Add in sentence about how the MA-PD plans are given weight, zero prices, in average bid calculation.

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 and thus, by definition, zero premium for the LIS enrollees qualify to enroll the randomly assigned population (Decarolis, 2015).

## 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 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

---

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.

<sup>8</sup>The per capita subsidy payment from Medicare is further adjusted by the risk score of each enrollee, while the premium may also include an additional payment for enhanced benefits if the plan offers them.

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.

Where are co-pays in anything we do?

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 state of residence.

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 component and a random shock. The deterministic indirect utility function of a regular enrollee  $i$  who chooses plan  $j$  in market  $t$  is given by:

$$v_{ijt} = -\alpha_i p_{jt} + \beta x_{jt} + \xi_{jt}, \tag{1}$$

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,  $\sigma$ , can be larger than a particular plan’s desired per capita revenue, the premium is then equal to  $p_{jt} = \max\{0, b_{jt} - \sigma_{jt}\}$ .  $b_{jt}$  denotes the supply-side price or the per capita revenue that is the insurers’ choice variable. Adopting Medicare’s terminology, we refer to  $b_{jt}$  as insurer’s bid to distinguish between supply and demand-side prices.  $x_{jt}$  contains observable characteristics of plan  $j$  in market  $t$ ,  $\xi_{jt}$  is a plan-specific fixed effect that captures unobserved plan quality. Each choice is also subjected to a random shock,  $\epsilon_{ijt}$ , distributed as a Type I Extreme Value:

$$u_{ijt} = v_{ijt} + \epsilon_{ijt}. \quad (2)$$

We define the market to be one of 34 statutory Part D geographic regions in years 2007 to 2010, for a total of 136 well-defined markets. 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.

We further include the vintage of the plan to 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.<sup>9</sup>

---

<sup>9</sup>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 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

fix this inconsistency in notation for the random coefficients

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 coefficients on premium and gap coverage are specified as:

$$\ln \alpha_i = \alpha + \sigma_\alpha \nu_i \quad (3)$$

where  $\nu \sim \mathcal{N}(0,1)$ , and  $\alpha$ ,  $\beta_{gap}$ ,  $\sigma_\alpha$ , and  $\sigma_{gap}$  are parameters of interest that guide the distribution of taste heterogeneity.

We next proceed to formulating a preferences model for the population eligible for low income subsidies. The institutional design of this part of the program posits substantial challenges for estimation. Typically, individuals eligible to receive low-income subsidies are automatically assigned to plans by the government rather than choose their plans. At the same time, however, individuals are eligible to change their assignment to a plan of their own choice after the random assignment took place. As the number of the so-called “LIS choosers” is substantial, competition for this part of the market potentially plays an important role in the pricing decisions of the insurers. In order to include this part of the market into the supply side part of the model, we need to estimate the elasticity of LIS demand.

We use observations on the choices of the “LIS choosers” as well as a set of assumptions about the structure of the outside option to recover the elasticity of demand in this part of the market. We posit that the demand of low-income beneficiaries can be described by a random utility very similar to the one we use for the 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  $i$  who chooses plan  $j$  in market  $t$  be given by:

$$v_{ijt} = -\alpha_i^{LIS} p_{jt}^{LIS} + \beta^{LIS} x_{jt}^{LIS} + \xi_{jt}^{LIS}, \quad (4)$$

where  $p_{jt}^{LIS}$  is the plan’s premium for the low-income population. This premium is computed

---

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% higher or lower than the “true” dynamic values. Recognizing this concern in our setting, we report the key counterfactual results in Section 6 for 20% interval around our marginal cost estimates.

as the remainder of the difference between the insurers’ bid and the state-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.

To close the demand model for the LIS enrollees, we assume that the potential market for the LIS population is defined as all LIS individuals enrolled in stand-alone PDP plans. Since many LIS enrollees are assigned to plans rather than choose plans, it would be unreasonable to assume that these “choices” represent individual preferences. Therefore, we say that all LIS-eligible individuals that are enrolled in plans that are eligible for Medicare’s automated LIS assignment are choosing the outside option. We thus estimate preferences of the LIS-eligible population from the choices of LIS “choosers” that enrolled in plans not eligible for random assignment.

## 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 key regulatory distortions and set-up a general profit function that can incorporate these distortions. We then discuss our strategy of arriving at an empirically tractable version of the supply-side model.

Consider one insurance plan  $j$  offered by a one-plan-insurer in one market. We assume that all characteristics of plan  $j$  are pre-determined and the only decision variable for this insurer is which bid  $b_j$  to submit to Medicare for plan  $j$ .<sup>10</sup> For each individual that plan  $j$  enrolls, the insurer collects an enrollee premium,  $p_j$ . The premium is a function of the bid  $b_j$ , as well as of the enrollment-weighted average of all other bids in the whole country,  $\bar{b}$ . Indeed, recall that premiums are determined as a residual between the insurer’s bid and the baseline subsidy, which is equal to a pre-set fraction of the average bid  $\bar{b}$ .

The subsidy payment  $\sigma_i$  from the government is different for each enrollee, as it is adjusted for individual risk profiles. 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. The level of the baseline subsidy depends on the average bid,  $\bar{b}$ . In other words, we can write the subsidy as a function of the average bid and individual-specific health risk:  $\sigma_i(\bar{b}, r_i)$ .

---

<sup>10</sup>In practice insurers that offer enhanced PDPs decide on both the bid and the “enhanced” premium. We take this aspect into account in the estimation, but abstract from it in the description of the model.



On the cost side, the ex-post costs of a plan differ for each enrollee and depend on individual drug expenditures. Some of the 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. For an individual with a given total annual drug expenditure amount, the costs of the plan will also depend on the cost-sharing characteristics of the plan, denoted by  $\phi$ . 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.<sup>11</sup> We denote the function which adjusts a plan’s ex-post profit with  $\Gamma$ .

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

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

For each individual, 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:  $\sigma_i = \sigma + \tilde{\sigma}_i$  and  $c_{ij} = c_j + \tilde{c}_{ij}$ . Denote the individual-specific difference in the subsidy and cost as  $\eta_{ij} = \tilde{c}_{ij} - \tilde{\sigma}_i$ . 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,  $\eta_{ij}$  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:

$$\pi_j(b) = \Gamma \left[ N(\mathbf{p})(p_j(\bar{b}, b_j) + \sigma(\bar{b}) - c_j(\bar{r}, \phi_j)) + \left( \sum_i \eta_{ij}(\phi_j) \right) \right]. \quad (6)$$

Denoting  $\eta_{ij}(\phi_j)$  with  $H_j(\phi)$ , 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.

---

<sup>11</sup>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))"

Note, however, that the premium that an insurer collects together with the baseline level of the subsidy is by construction equal to the bid submitted by insurer to Medicare, i.e.  $p_j(\bar{b}, b_j) + \sigma(\bar{b}) = b_j$ . We can then re-write the pre-risk corridor profit of plan  $j$  as:

$$\pi_j(b_j) = (b_j - c_j)s_j(p_j, p_{-j})M - H_j(\phi), \quad (7)$$

where premiums are functions of insurers' bids:  $p_j(\bar{b}, b_j)$  and  $p_{-j}(\bar{b}, b_{-j})$ .

We now expand this expression to allow for multi-plan insurance organizations as well as to allow for different prices, marginal costs, and demand elasticity for the Low-Income-Subsidy 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:

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

where (ignoring type superscripts)  $M_t$  is the population in the market (defined as region-year),  $s_{jt}(p_j(\bar{b}_t, b_{jt}), p_{-jt}(\bar{b}_t, b_{-jt}))$  is the share of plan  $j$  given the vector of all bids and the bid-averaging rule that translates bids into premiums,  $b_{jt}$  is the firm's bid for plan  $j$  in market  $t$ , and  $c_{jt}$  is the marginal cost. Firms maximize profits by choosing bid  $b$  for each insurance plan in each market. While similar, Equation 8 is more complex than a standard profit function in a differentiated products market. The key difference lies in 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:

$$p_{jt}^R = \max \{0, b_{jt} - \zeta \bar{b}_t\}, \quad (9)$$

where  $\bar{b}_t$  is the enrollment-weighted average bid of all plans in the entire US and  $\zeta$  is the share of the average bid covered by the federal subsidy.  $\zeta$  is set every year by CMS and is governed by fiscal considerations and the Part D statutes. 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. For the latter group, the share again depends on premiums that are non-linear functions of bids and subsidies. [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. Section 5.3 outlines how we deal with the piece-wise structure in the estimation. Here, we derive the first-order conditions for an insurer that it is not eligible for random assignment, but rather can only enroll LIS “choosers.” In this case the share function for the LIS population is differentiable.

Then, for a contract  $j$  offered by firm  $J$ , the Nash-Bertrand first order condition for setting bid  $b_j$  is as follows (omitting market subscripts):<sup>12</sup>

$$\begin{aligned} \frac{\partial \pi_J}{\partial b_j} = & M^R s_j^R(b) + (b_j - c_j^R) M^R \frac{\partial s_j^R(b)}{\partial b_j} + \sum_{k \neq j \in J} (b_k - c_k^R) M^R \frac{\partial s_k^R(b)}{\partial b_j} \\ & + M^{LIS} s_j^{LIS}(b) + (b_j - c_j^{LIS}) M^{LIS} \frac{\partial s_j^{LIS}(b)}{\partial b_j} + \sum_{k \neq j \in J} (b_k - c_k^{LIS}) M^{LIS} \frac{\partial s_k^{LIS}(b)}{\partial b_j} \end{aligned}$$

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 and LIS enrollees  $c_j^R$  and  $c_j^{LIS}$ , we need to make an additional assumption to close the model. As Medicare specifically increases its risk-adjustment payments to plans for each LIS enrollee, we will assume that those payments make the marginal cost of these two groups the same from the point of view of the insurer. In other words, we assume that  $c^R = c^{LIS}$ . Imposing this assumption and collecting terms in vector notation, we arrive at:

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

where

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

Note again the role of the relative market size for regular and LIS enrollees.

---

<sup>12</sup>Note that we do not take into account the effect that each bid has on the demand of plans in other markets through the national bid averaging system. Since bid averaging is weighted by national enrollment weights, each plan has a very small impact on this average.

### 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 ( $\Pi$ ), and the deadweight loss associated with taxation used to fund government subsidies ( $G$ ):

$$W = CS + \Pi - \lambda G, \quad (12)$$

where  $\lambda$  is the social cost of raising revenues to cover government expenditures,  $G$ .

Since utility is ordinal, we need to impose a normalization that would allow us to measure surplus levels in dollars. A natural normalization is to consider all three pieces of the welfare function as being defined relative to the outside option. For consumer surplus  $CS$  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 the choice of not purchasing any Part D coverage. For producer surplus, or profits  $\Pi$ , 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 have to consider only the extra government spending in PDP as compared to what it would have spent on subsidizing the same individual in MA-PD. We conservatively assume that the outside option for the government is subsidizing MA-PD, excluding the possibility that some individuals could leave subsidized insurance altogether.

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

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

where  $\gamma$  is Euler’s constant, and  $v_{ij}$  is the deterministic component of utility <sup>13</sup> for person

---

<sup>13</sup>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

$i$  from contract  $j$  as given in Equation 1.<sup>14</sup> We integrate out over the unobserved taste heterogeneity to obtain consumer surplus:

$$CS = \int CS(\theta)dF(\theta). \quad (14)$$

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. For each plan  $j$ , we thus measure the profit as follows and then add up the profits of all plans in each market.

$$\pi_{jt}(b) = (b_{jt} - c_{jt})s_{jt}(b_{jt}, b_{-jt}, \bar{b}_t)M_t^R. \quad (15)$$

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  $\lambda = 1.3$ . 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:

$$W = \overbrace{\int \frac{1}{\alpha} \left( \gamma + \ln \left[ 1 + \sum_{j=1}^J \exp(v_j(\theta)) \right] \right) dF(\theta)}^{\text{Consumer Surplus (CS)}} + \overbrace{\sum_{j=1}^J (b_j - c_j)s_j(p)}^{\text{Producer Profit (II)}} - \underbrace{\lambda \left( \sum_{j=1}^J (G_j^{PDP} - G^{MAPD})s_j(p) \right)}_{\text{Social Cost of Government Spending (G)}}. \quad (16)$$

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 ac-

---

different conclusions in our counterfactuals (which is intuitive, since keeping or removing the vintage effect from the utility function affects all price counterfactuals symmetrically).

<sup>14</sup>Euler's constant is the mean value of the Type I Extreme Value idiosyncratic shock under the standard normalizations in the logit model, and is approximately equal to 0.577.

count the cost of public funds for doing that in the social planner’s problem. Another way 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 16, except that the “product market profit” is included under the  $\lambda$ -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):

$$p^{SocialPlanner} = \operatorname{argmax} W^{SP}(p) = c + \Delta G + \Omega(p)^{-1} \frac{(1 - \lambda)}{\lambda} s(p) \quad (17)$$

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 surplus, which is captured by the social planner and recycled into government revenues (if positive).

## 5 Data and Estimation

### 5.1 Data and Descriptive Facts

Our primary data set combines a variety of aggregate plan-level information released annually by CMS.<sup>15</sup> 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.

Table 1 provides summary statistics for the PDPs analyzed. Although we observe between 1,500 and 1,800 PDPs offered each year, the supply-side is rather concentrated. A total of around 50 insurer parent organizations appears in the data, with on average 17-19 separate organizations competing in each of the 34 regional markets in which CMS divides the US. The table also shows that the average plan premiums for regular enrollees increased quite substantially in the time frame we are considering. The unweighted average premium went

---

<sup>15</sup>All of the data is publicly available at <http://www.cms.gov/Research-Statistics-Data-and-Systems>.

up from \$439 per year in 2007 to \$559 in 2010. This growth in premiums was accompanied with increased dispersion in plan premiums and in particular with a higher number of very expensive plans. The growth in premiums between 2007 and 2010, differed dramatically across regions. The increase in the non-weighted average premium over the four years ranges from 13 percent in New Mexico to 61 percent in the California market. The baseline subsidy levels grew slower than bids and premiums. The baseline subsidy for an average risk enrollee went from \$637 in 2007 to \$677 in 2010, which is an increase of 6% as compared to an increase in unweighted average bids and premiums of 10% and 27% respectively.

Figure 5 documents in the top left panel 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.<sup>16</sup>

The three other plots of Figure 5 further show the 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. Interestingly, the plans that offered coverage in the donut hole are consistently more expensive, up to three times more expensive, than plans with only the minimally required coverage. We also see that the dispersion in premiums is relatively similar across different markets in the 2010 cross-section. The low-income subsidy is on average \$388, but with variations 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 have market share close to zero, some plans cover as many as 20% 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 choose the “outside option” if they are enrolled in plans that are eligible for LIS random assignment.

## 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 (1) reports the estimates of the random coefficient logit model described earlier. These estimates will be used for the counterfactual analysis in the next section. For comparison purposes, we also report in columns (2) and (3) estimates for the [Berry \(1994\)](#) logit model that does not involve random coefficients.

---

<sup>16</sup>??, ?? and CBO (2014) describe how competition and subsidy-related distortions are associated with the evolution of premiums in Part D.

Column (2) contains OLS estimates, while column (3) contains 2SLS estimates.

These three sets of estimates indicate rather similar results. We estimate a negative and significant coefficient on plans premium and deductible. We also estimate positive coefficients on plan generosity features: Beneficiaries appear to like plans that offer coverage in the gap, cover more of common drugs and include more pharmacies in their networks. 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. Considering the magnitude of the estimated effects, we note that the random coefficient estimates are typically closer to the 2SLS than to the OLS. In particular, the mean price coefficient amounts to -13.69 for the random coefficient, relative to -10.44 for the 2SLS and -2.74 for the OLS. The random coefficient model shows substantial dispersion in the sensitivity to price. At the same time, we do not find significant dispersion for the other variables for which we allow random coefficients (the inner option and the dummy for gap coverage). In terms of price elasticity of demand, the random coefficient estimates imply a modal own and cross price elasticity of -5.98 and 0.06 respectively.

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 premium support 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 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. 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.

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).<sup>17</sup> The concern regarding the bias in the coefficient on premiums is that they might

---

<sup>17</sup>For example, if an insurer offers a contract with some coverage in the gap in the first years of the



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 “same” plan in other geographic markets.<sup>18</sup> The idea of this instrument is 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 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.

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).<sup>19</sup> 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

---

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.

<sup>18</sup>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.

<sup>19</sup>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).

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

The key step in the supply-side estimation is the recovery of plan-level marginal costs. As argued in Section 4, the first order condition linking marginal costs to bids and demand parameters has several non-standard features. 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 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).<sup>20</sup> 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:

$$mc_{jt} = X_{jt}\beta + \delta_t + \tau_j + \epsilon_{jt}, \quad (18)$$

---

<sup>20</sup>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.

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 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% 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  $(\text{bid}-\text{marginal cost})/(\text{marginal cost})$ . 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.<sup>21</sup> This gives us an estimate of direct costs for basic benefit of \$861. Using the latter number - which certainly is within the scope of our estimated marginal costs - and the average revenue reported above, we would arrive at a mark-up of  $(\$1,136-\$861)/\$861=32\%$ . 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% markup, we would have to assume that administrative costs are around \$170 on average per person -  $(\$1,136-\$861-\$170)/(\$861+\$170)=10\%$  - or 70% of the number that CBO reports as an estimate of administrative costs plus profits, which appears entirely plausible. Overall, our estimates appear realistic and consistent with the actual cost data.

## 5.4 Welfare Estimates

Using the demand and supply estimates, we compute consumer surplus, producer profits, government transfers, and total surplus for the observed market allocation, using the welfare function in Equation 16. We restrict our calculations to regular enrollees. The calculations are reported in the first column of Table 5. We estimate that at the observed premiums, the total annual consumer surplus is about \$2.5 billion.<sup>22</sup> This value is estimated relative to the outside option, as the utility model is inherently ordinal. Insurer profit that we estimate without the multitude of re-insurance transfers from the government, is \$500 million.

On the cost side, we calculate that the government spends about \$6 billion on premium subsidies for regular enrollees - equal to about \$680 a year for each individual. Moreover, using CMS data on average non-premium level of subsidies, we calculate that the total amount of direct reinsurance subsidies is on the order of \$3.5 billion. To take into account that the government would still have to pay subsidies if individuals were to leave PDP and switch to MA-PD, we also estimate how much the government would have spent on the same individuals were they to enroll in MA-PD.<sup>23</sup> The difference between the total PDP subsidies

---

<sup>21</sup>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.

<sup>22</sup>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.

<sup>23</sup>We utilize public data realized annually by CMS on the average levels of subsidies for different Part D plans to calculate these estimates.

and what these would have been in MA-PD, allows us to calculate the “net” government spending on PDPs, which is the value we need to calculate total surplus as in Equation 16.

Multiplying the net government subsidy by 1.3 to account for a 30 cents a dollar cost of public funds, and subtracting it from consumer and producer surplus, we arrive at the total surplus calculation for the regular Part D PDP market to be about \$1.5 billion. In relative terms, we estimate that a dollar of public funds generates only 0.17 dollars of surplus per dollar directly spent in the PDP program if we do not take into account the government’s opportunity cost of paying MA-PD subsidies. 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.

In the next section of the paper, we will conduct several simulations of counterfactual market structure and regulatory regimes. We will evaluate these counterfactuals using the same approach as we just outlined for calculating welfare in the non-counterfactual data. Since evaluating welfare in the counterfactual settings will require several additional assumptions, we outline additional details of how we calculate government spending on regular enrollees.<sup>24</sup> In each counterfactual, we calculate government spending on premium subsidy for stand-alone prescription drug plans as the sum of per capita simulated subsidies multiplied by counterfactual enrollment predicted by the demand model. We assume that the average reinsurance subsidy for each plan 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 data point for the non-premium subsidy.<sup>25</sup> In 2010, for example, the unweighted mean per capita reinsurance payment among PDP plans was \$503 per plan with a standard deviation of \$297. 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

---

<sup>24</sup>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. We are further not explicitly calculating the selection component of the profit function  $H(\phi)$ .

<sup>25</sup>For raw CMS data see <http://www.cms.gov/Medicare/Medicare-Advantage/Plan-Payment/Plan-Payment-Data.html>.

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. We use this average spending together with enrollment predictions for inside and outside option 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.

## 6 Simulation Results

**A) Removing linkages to other parts of the market.** 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. We proceed in two steps. First, we remove the LIS pricing incentive, while maintaining unaltered other elements of the current formula. Hence, we set prices to be:  $p_{jt} = b_{jt} - 0.68 * (Average(b_{PDP}, b_{MAPD}))$ .<sup>26</sup> Second, we remove the presence of MA-PD bids in the mechanism. Thus, we obtain:  $p_{jt} = b_{jt} - 0.68 * (Average(b_{PDP}))$ . The latter 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. Relative to the allocation observed in the data (column 1), removing the LIS pricing incentives from the mechanism results in slightly lower enrollment-weighted average premiums (\$471 vs \$502 in the data), but also higher consumer surplus. Removing LIS incentives of submitting lower bids, leads to slightly higher average bid and on net higher subsidies for regular enrollees. Therefore, even though consumer surplus increases once we detach the LIS

---

<sup>26</sup>We do not observe MA-PD bids directly as MA-PD premiums contain not only subsidies whose levels are known, but also unobserved Part C rebates. We back out the contributions of MA-PD to bid average by calculating averages that would have resulted from PDP bids only and the CMS data on realized bids.

part of the market, insurer profits and government expenditures also go up, resulting in a lower total surplus of \$1.47 billion. Next, we also remove the MA-PD portion of the average bid that is currently used by CMS to calculate subsidies. 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 PDPs' strategic response. Consequently, the subsidy, which is set at 68% of the average bid also increases. In net, removing MA-PD part of the mechanism results in lower average premiums, with the average premium going down to \$418. The change in premiums increases enrollment in PDP up to 55%. Higher subsidies, however, also imply higher government spending, and thus the net result is a lower total surplus amounting to \$0.92 billion. Given the increase in government spending, the per dollar ratios go down to about 6 cents on a dollar and 21 cents on an opportunity cost dollar.

**B) Effects of altering 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. 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. Compared to the baseline counterfactual, we find the expected pattern that profits increase greatly and consumer surplus declines under the monopolistic regime (consumer surplus goes down to \$2.8 billion and profits go up to \$2.2 billion), with the opposite pattern under atomistic competition (consumer surplus at \$3.4 billion and insurer profits at \$1.2 billion). Interestingly, total surplus is higher under monopoly, as the loss driven by consumer surplus is dominated by the increase in producer profits. Under atomistic competition, the changes are less dramatic, but result in negative welfare relative to the observed ownership structure. The motive is that at the low premiums observed under full competition the marginal benefits of serving additional consumers are exceeded by the social costs of providing the goods.

**C) Alternative subsidy systems: proportional vs flat subsidies.** In the next set of counterfactuals we ask whether deviations in the subsidy mechanism from the averaging rule currently used to proportional or flat subsidies could improve total welfare.

We start with a proportional subsidy mechanism in which premiums are given by:  $p_{jt} = x * b_{jt}$ , where  $(1 - x)$  with  $x \in [0, 1]$  is the proportional subsidy. The simulation results reveal the poor performance of this system. To illustrate this, we report in Table 6 the findings for three levels of  $x$ : 5%, 32% and 95%. The idea of the 5% counterfactual is to

test how insurers would respond if the government almost entirely bore consumer premiums. This counterfactual removes most of consumer price sensitivity, as the government is not price sensitive in the model and consumers bear only 5% of the premiums. Insurers dramatically increase their bids: we find a threefold increase in bids. Consequently, even though individuals now pay only 5% 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% in the PDP plans from 55% in the baseline. Consumer surplus increases to \$13 billion. The change in insurers' profits is, as expected, very large. Given the dramatic increase in bids, government spending increases dramatically. The result is a stark drop in welfare levels to negative \$37 billion. The per-dollar efficiency measures are small and negative. In this counterfactual, we essentially generate a large transfer from the taxpayers to the insurers with a less than one-to-one pass-through to consumers.

Similar results are obtained with a 32% subsidy. However, in counterfactual where we decrease premium subsidies to be only 5% of the price, we observe that consumer prices are about three-fold of the baseline level. At this level of prices, very few individuals (1%) are willing to purchase PDP plans and switch to the outside option. Despite the drop in enrollment, the program generates almost the same amount of total surplus as in the benchmark case, and a very high surplus per dollar spent if we do not take into account government' spending on the outside option. Such high per dollar surplus is not surprising, as the government is now paying little and only the beneficiaries with the highest willingness to pay participate in the program. 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, by which, for example, a very generous subsidy would decrease the overall level of prices. 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% of the extra premium in more expensive plans, competitive forces are significantly muted and insurers pass through substantially higher expenditures to the (by construction) inelastic federal budget.

The second alternative subsidy system that we explore is a flat voucher mechanism. The premiums are set to equal to:  $p_{jt} = b_{jt} - F$ . We assume that insurers and consumers know the levels of the flat subsidy in advance and adjust their behavior accordingly. In calculating the new equilibrium bids, we check corner solutions, where the insurers may decide to bid exactly at the subsidy level. In Table 6, we report results on consumer and producer welfare, as well as government spending for three flat subsidy levels  $F$ : \$0, \$721 and \$1340. The value of \$721 per year corresponds to the officially set fraction of the average bid, while



\$1340 is twice that amount. The fourth column of Table 6 illustrates the components of total welfare for  $F = 0$ .

The second alternative subsidy system that we explore is a flat voucher mechanism. The premiums are set to equal to:  $p_{jt} = b_{jt} - F$ . We assume that insurers and consumers know the levels of the flat subsidy in advance and adjust their behavior accordingly. In calculating the new equilibrium bids, we check corner solutions, where the insurers may decide to bid exactly at the subsidy level. In Table 6, we report results on consumer and producer welfare, as well as government spending for three flat subsidy levels  $F$ : \$0, \$721 and \$1,340. The value of \$721 per year corresponds to the officially set fraction of the average bid, while \$1340 is twice that amount. The fourth and sixth columns illustrate the two most extreme counterfactuals. In the fourth column of Table 6, 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,  $F = 1,340$  turns out to be a sufficiently generous subsidy to guarantee 100% enrollment in PDPs. Albeit full enrollment in PDPs was attained also with the generous proportional subsidy, the total welfare in the two cases is remarkably different. Although both entail negative welfare, the flat subsidy does better: a loss of \$6.7 billion relative to one of \$37.4 billion under the 95 percent proportional subsidy. Insurers utilize the increase in subsidy and increase their bids, but not dramatically, as consumers still face higher prices for more expensive plans. Consumer surplus increases almost three-fold to \$13 billion. Producer profit increases substantially as well. At the same time, government spending increases too, driving down total welfare.

The third case,  $F = 721$  is particularly noteworthy. Although  $F = 721$  implies that the subsidy is nominally the same as the observed subsidy, incentives are very different under a flat voucher relative to the observed mechanism. While in the observed mechanism, the subsidy is determined after the bidding process as a fraction of the average bid, here we set the subsidy ex ante and it does not depend on the submitted bids. We find that the mechanism matters, as consumer surplus, producer surplus and allocations change relative to our benchmark. First, we find that prices that consumers face increase by about \$100. This leads to a drop in enrollment to 43% (relative to the benchmark 55%) and an accompanying decrease in total consumer surplus. Government spending patterns, however, change starkly in response to decreased enrollment and thus overall, the total surplus generated in the program is substantially higher under the ex ante set voucher than in the benchmark case (\$1.7 billion vs. \$0.9 billion in the benchmark). Importantly, the efficiency per dollar, and especially per opportunity-cost dollar spend jumps dramatically, from 24 cents per opportunity-cost dollar to 77 cents.

We repeat the calculation of total welfare at a range of vouchers from \$0 to \$1400, in

order to identify the optimal uniform voucher. The first Panel in Figure 7 summarizes the outcome of these calculations graphically. We find that the total welfare is the highest at  $F = 600$ . Setting higher vouchers significantly reduces welfare. So do lower 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.

**WHAT TO DO WITH THIS PARAGRAPH?** Recognizing that due to concerns about inertial consumers in this market, our static marginal costs estimates may under or over-estimate the “dynamic” marginal costs (see Miller (2014)), we repeat the key voucher counterfactuals for MC values in a symmetric 20% interval around our baseline estimates. The results of this robustness check are reported in Figure 9. We again show the total welfare values calculated at the range of vouchers from \$0 to \$1,400 at three levels of marginal costs. The values marked with green dots are baseline estimates - repeating the values from Figure 7. The estimates above and below are for counterfactual simulations that assume MC to be 20% lower or 20% higher than baseline, respectively. We observe that welfare estimates for all voucher levels shift up (by about XX%) or down (by about XX%), depending on whether we increase or decrease the marginal costs. 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.

In looking at welfare outcomes at different voucher levels, we held constant the assumption that subsidy levels are the same in each market across the country. At the same time, we document substantial heterogeneity in demand, supply, and prices across 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 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.

**D) Second best benchmarks: marginal cost pricing and 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 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 through the optimal flat subsidy (either uniform across all markets, or market specific) relative to marginal cost pricing and the social planner. Table 7 reports the full set of welfare estimates for all the second best scenarios that we explore as well as, in the last two columns, the welfare values associated with the optimal flat voucher and the region-specific optimal flat voucher.

As regards the second best benchmarks, we start by setting premiums equal to the estimated marginal costs for each insurance plan:  $p_{jt} = MC_{jt}$ . The results are reported in Table 7. Facing premiums as high as marginal costs, consumers leave the PDP program in favor of the outside option with enrollment dropping to 1%. 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. 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 the following column of Table 7, are similar, albeit starker. In this case, enrollment and total surplus fall even more, with nearly 0% of the market choosing the inside option of PDP plans.

Next we consider the social planner’s problem. As expected, the social planner’s problem generates the highest total surplus of \$3.6 billion. This surplus comes at a cost 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” at \$10 billion. As the algebraic expression for social planner’s prices in Equation 26 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.

**FIX THIS WITH NEW NUMBERS: One caveat of the analysis, however, is**

that it is not unreasonable to think that a change in the PDP subsidy will affect MA-PD subsidies. Then 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 (e.g. in real life). We can implement this by just mechanically imposing a penalty on the utility of the outside option such that the relative market share of the inside and outside option remain constant. The results that we obtained are reported in the fourth column of Table 7. The simulation reveals a substantially higher total welfare [check numbers].

These four second best benchmark can be contrasted with the optimal voucher results reported in the last two columns of Table 7. The 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.

## 7 Conclusion

In this paper we have analyzed the welfare effects of the subsidization mechanism in Medicare Part D, focusing on incentives created by subsidies on the supply-side. We draw several conclusions. First, we find that the current program as it is, for the market of regular enrollees in stand-alone prescription drug plans, generates surplus per one dollar of government spending 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.

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 Prescrip-

tion 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 affect 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 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 through 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, or 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.
- 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). Competition 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.
- 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. P. (2015). Risk Adjustment and Low Income Subsidy Distortions in Medicare Part D. *Clemson University Working Paper*.
- Miller, D. P. and J. Yeo (2014). 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*.



- Miller D.P., Yeo, J. (2014). The Consequences of a Public Health Insurance Option: Evidence From Medicare Part D. *Mimeo, Clemson University and Singapore Management University*.
- Nevo, A. (2001). Measuring Market Power in the Ready-to-Eat Cereal Industry. *Econometrica* 69, 307–342.
- 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*.
- 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.
- 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.

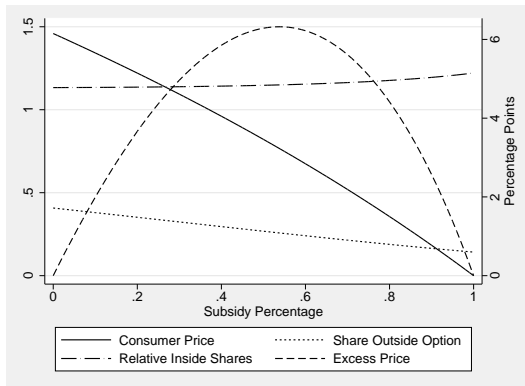


Figure 1: A Simple Model of Proportional Subsidies

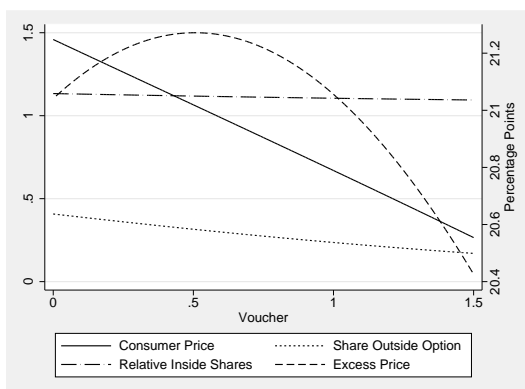


Figure 2: A Simple Model of Vouchers

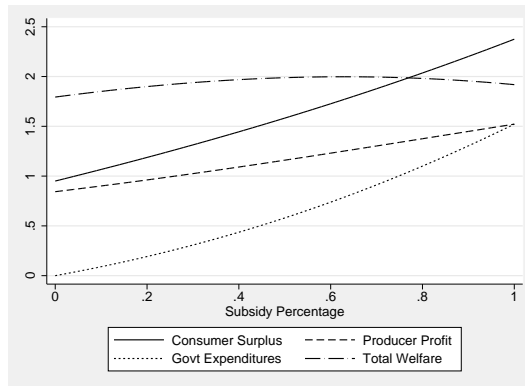
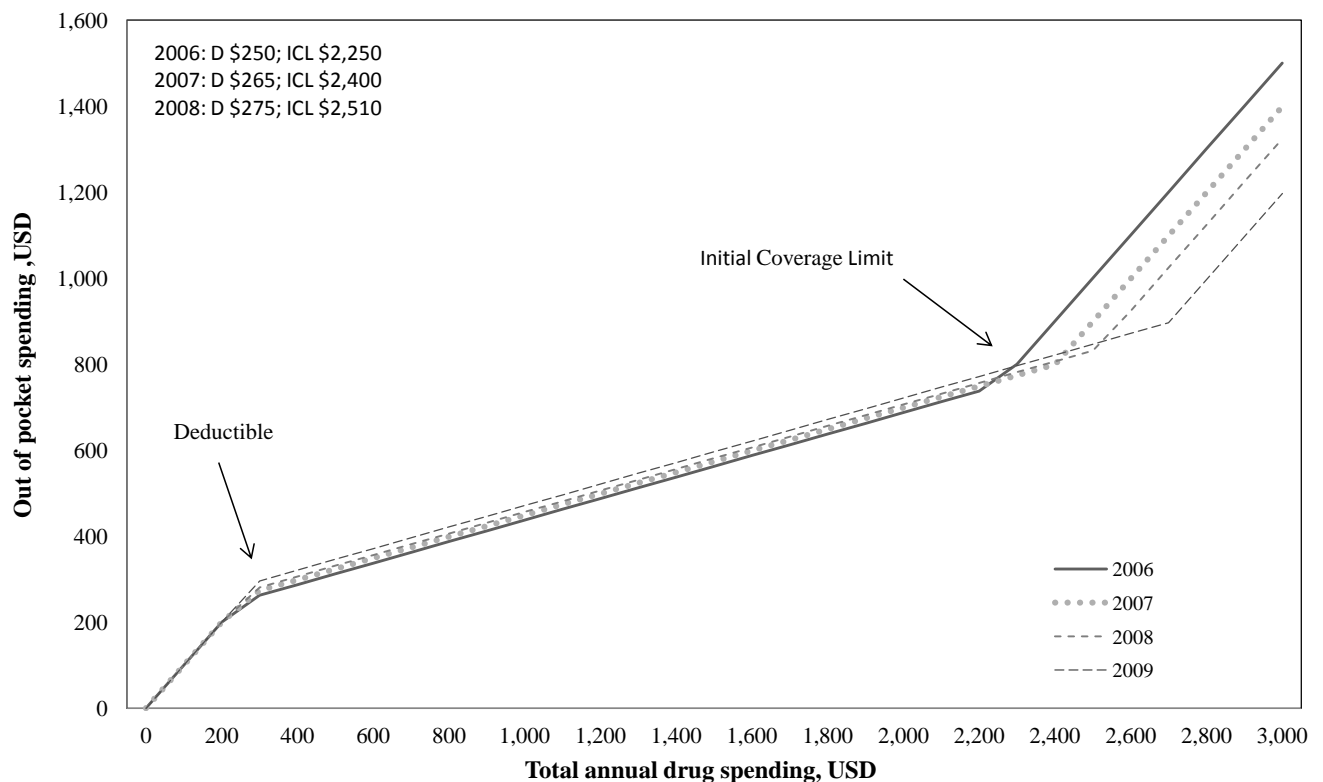


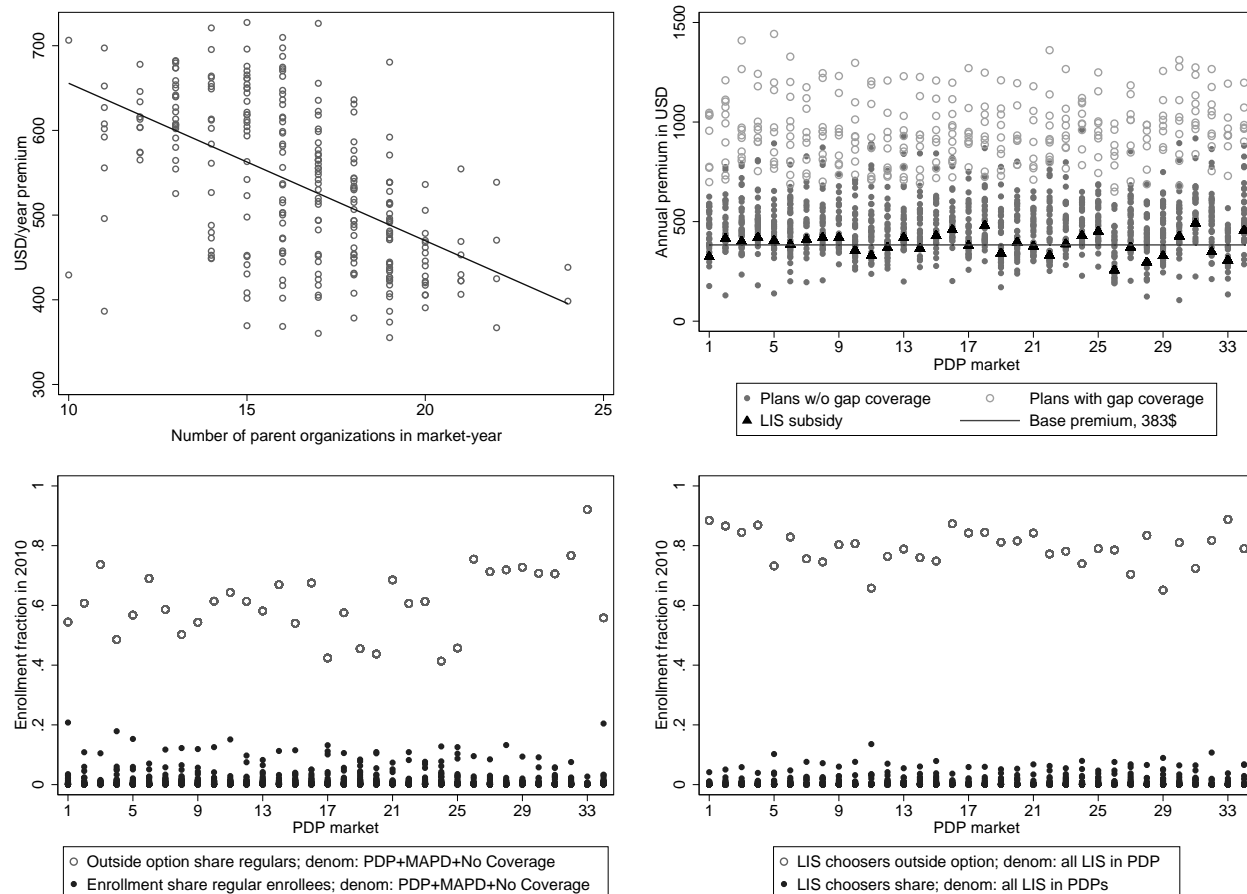
Figure 3: Welfare Decomposition of Proportional Subsidies

Figure 4: Minimum coverage requirements in Part D



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% up to the initial coverage limit (ICL) and the subsequent donut hole that has a 100% 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. Consider an individual, who in 2006 was in an SDB contract, and purchased prescription drugs for a total of \$3,000. Out of this amount, the individual would pay the deductible of \$250, then 25% of the next \$2,000 up to the ICL of \$2,500, and then 100% of the next \$750 in the gap, for a total out of pocket spending of \$1,500. As the figure illustrates, the generosity of the SDB changed over time. For example, an individual spending \$3,000 on drugs in 2009 would face the out-of-pocket expenditure of less than \$1,200.

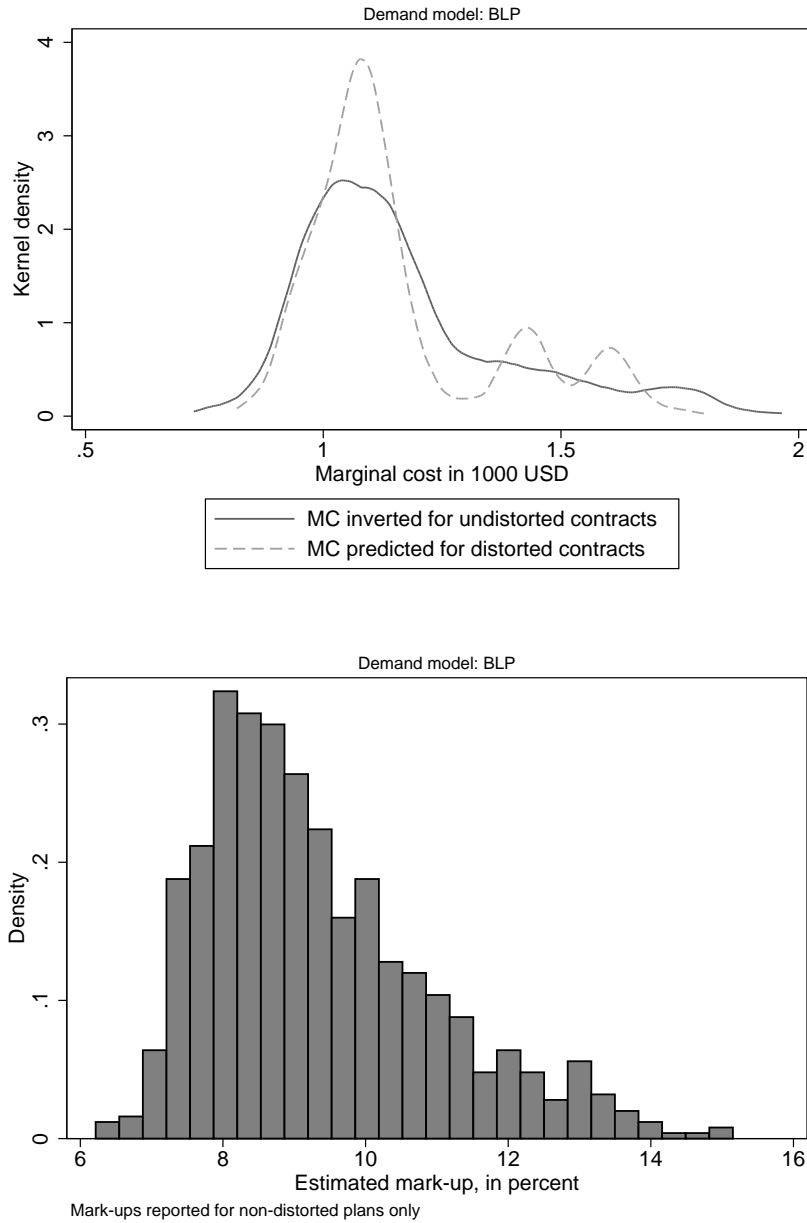
Figure 5: Contract prices and market structure



Panel 1. Correlation between the number of competing insurers and average premiums in the market. The premiums are not enrollment-weighted. The data includes years 2006-2010. Panel 2. For year 2010, the figure depicts cross-sectional variation in premiums, in the levels of low income subsidies, and in "base premiums". Plans with premiums below LIPSA are eligible for the random assignment of LIS beneficiaries. Panels 3 and 4. The distribution of market shares of the outside and inside options across 34 Medicare Part D regions, separately for regular and LIS-eligible enrollees. According to CMS, in 2010, almost 47 million individuals in the US were eligible for Medicare Part D coverage. This includes beneficiaries of all incomes that were eligible both due to old age and disability (ca. 9 million). Out of 47, about 42 million had coverage that satisfied minimum requirements. Out of the 42 million with creditable coverage, 18 million were enrolled in stand-alone PDPs, 10 million were enrolled in MA-PDs, about 6 million in employer-sponsored plans, and about 8 million had other coverage, such as federal or VA insurance. In the graphs, we restrict the definition of the market to include PDP options, MA-PDs and no coverage choices. The latter two comprise the outside option.

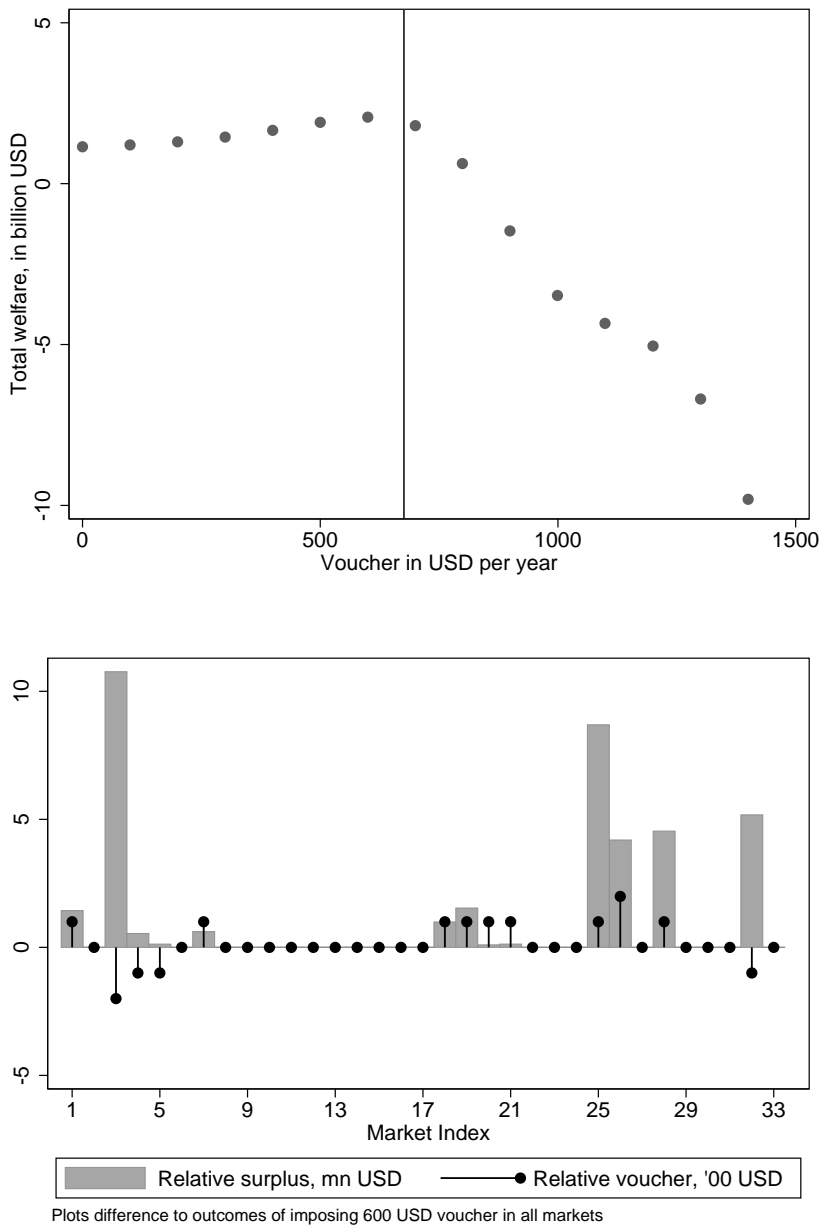


Figure 6: Recovering marginal costs



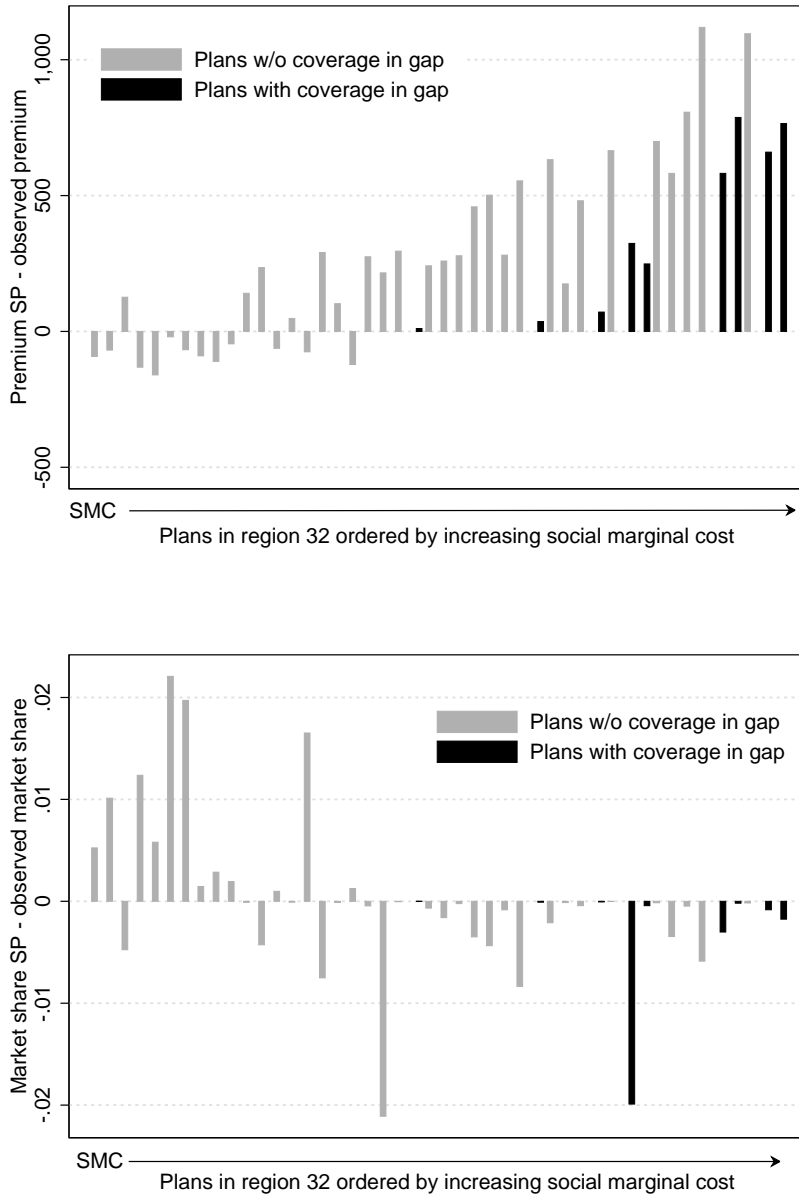
Panel 1. Marginal cost estimates from inversion (for undistorted contracts) and from hedonic projection (for distorted contracts). Plan characteristics used in the hedonic projection include deductible, coverage in the gap and enhanced plan indicators, measures of formulary generosity, pharmacy networks, vintage, as well as estimated unobserved plan quality, and region and insurer fixed effects. Panel 2. Mark-ups are 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. The measure of price in this case is insurer's bid - basic bid submitted to Medicare for baseline coverage plus enhanced premium that covers additional benefits.

Figure 7: Welfare with flat subsidies: uniform and market-specific optimal vouchers



Panel 1. Plots estimated total welfare in full equilibrium counterfactuals at different levels of voucher-like subsidies. Panel 2. From Panel 1 it follows that optimal uniform voucher lies at \$ 600. In Panel 2 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 solution: illustration of change in allocation



Example market: California in 2010. Panel 1. Plots the difference between premiums set by social planner and observed premiums. Panel 2. Plots the difference between shares in the social planner counterfactual and observed shares. Note that the social planner increases prices in contracts with higher social marginal cost, which shifts enrollment out of these contracts to plans with lower social marginal cost.



Table 1: Summary Statistics

Year	2007	2008	2009	2010
<b>Plans</b>				
Total number of PDP plans	1,742	1,791	1,674	1,565
Average number of PDP plans per market	51	53	49	46
<b>Firms</b>				
Total number of PDP parent organizations	56	56	50	50
Average number of PDP parent organizations per market	19	19	17	17
<b>Premiums</b>				
Unweighted average annual PDP consumer premium	\$439	\$477	\$545	\$559
<b>Subsidies</b>				
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 (LIS) benchmark threshold	\$341	\$333	\$353	\$388
<b>Enrollment</b>				
All Part D Eligible (in millions)	43.3	44.4	45.5	46.6
PDP enrollment, non LIS (in millions)	8.3	8.6	8.9	9.4
PDP enrollment, LIS (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 (0.15)	-10.44 (1.09)		-2.08 (0.09)	-7.58 (1.27)
$\mu_\alpha$			2.58 (0.36)		
$\sigma_\alpha$			0.27 (0.09)		
$\sigma_{inner}$			0.00 (0.08)		
Annual Deductible	-3.25 (0.21)	-6.72 (0.56)	-7.11 (1.52)		
Coverage in the Gap	0.18 (0.07)	2.93 (0.40)	2.89 (0.94)		
$\sigma_{gap}$			0.00 (0.08)		
Number of Top Drugs Covered	0.24 (4.36)	31.58 (7.30)	30.4 (12.1)	-3.05 (4.62)	10.63 (7.15)
Pharmacy Network Measure	0.31 (0.04)	0.29 (0.06)	0.31 (0.01)	0.23 (0.05)	0.22 (0.08)
Number Years Plan on Market	0.61 (0.02)	0.88 (0.05)	0.90 (0.12)	0.76 (0.3)	0.88 (0.04)

Standard errors in parentheses. Includes firm, market, and year fixed effects.

Sample: Medicare Part D stand-alone prescription drug plans in years 2007 to 2010. Columns (1)-(3) report estimates for the regular enrollees. Columns (4) and (5) report estimates 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; number of APIs in formulary; number of drugs placed in Tiers 1-2 of the formulary (i.e. having low cost-sharing). In the regressions for the LIS enrollees, we assume that LIS choosers and randomly assigned LIS beneficiaries choose the “outside option” if they are enrolled in plans that are eligible for LIS random assignment. Columns (1), (3) and (5) use the set of instrumental variables described in the text. In column (1), the computed mean is reported for the distribution of the price parameter in the random coefficients specification. The mean is computed as  $exp(\mu + \frac{1}{2}\sigma^2)$ , where  $\mu$  is estimated at 2.58 and  $\sigma$  is estimated to be 0.27.

Table 3: First stage - BLP and Hausman instruments for plan premiums

	(1)	(2)
	Regular FS	LIS FS
No. of PDP plans in a region-year by same PO	-9.631*** (2.375)	-3.318 (4.352)
No. of MA plans in a region-year by same PO	-0.184 (0.242)	0.0285 (0.474)
Deductible of MA plans in the same region-year	-0.147 (0.173)	0.0294 (0.344)
Hausman IV	0.371*** (0.0326)	0.275*** (0.0551)
Observations	6675	4561

Standard errors in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

The table reports the first stage for a collection of variables that are used as instruments for premiums. 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. PO stands for Parent Organization. The regressions also include all plan characteristics that are used in demand estimation, including a constant, fixed effects for geographic markets, parent organizations, and years.

Table 4: Marginal cost projection

	(1) Berry MC inversion	(2) BLP MC inversion
Annual deductible	-0.365*** (0.0440)	-0.361*** (0.0429)
No. of common APIs	0.142 (0.130)	0.151 (0.127)
Has coverage in the gap (1/0)	0.422*** (0.0105)	0.411*** (0.0102)
Enhanced plan (1/0)	-0.0352** (0.0118)	-0.0305** (0.0115)
No. of top drugs in Tier 1 and 2	-0.569 (0.380)	-0.540 (0.370)
No. of top drugs covered	-8.696*** (2.476)	-8.686*** (2.410)
Pharmacy network measure	-0.188*** (0.0509)	-0.187*** (0.0495)
Number of years the plan is on the market	46.09*** (3.170)	43.78*** (3.086)
Mean dep.var.	1.171	1.171
Std. dev. dep. var.	0.239	0.229
R-squared	0.868	0.864
Number of observations	756	756

Standard errors in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

We select a subset of “undistorted” insurance plans in 2010. These are defined as plans offered by insurance companies that had not qualified for LIS random assignment with any of their plans in a given market. For example, if Humana had qualified to enroll LIS beneficiaries in California, we would exclude all plans offered by Humana in California in 2010, calling them “distorted” plans. We next take the sample of undistorted plans and project the MC estimates obtained through the inversion of the first-order conditions (Equation 10) onto characteristics of plans. The result of this regression is reported in this table separately for the Berry IV and BLP demand models. Note that for the model marked with BLP, we use the Berry IV specification for the LIS part of the market in the inversion procedure. The MC projection regression has exactly the same characteristics as the utility function. Fixed effects for markets and insurers are included in the regression but not reported.

Table 5: Welfare Estimates: Observed Allocation and Simulations of CMS Mechanism

	Data	no LIS link	no LIS no MA-PD link	Independent plans	Monopoly ownership
Consumer Surplus, \$mn	2,517	2,892	3,413	3,350	2,814
Insurer Profit, \$mn	459	1,063	1,279	1,172	2,161
Premium subsidy, \$mn	5,936	7,807	10,125	9,742	8,884
Reinsurance subsidies, \$mn	3,444	4,096	5,026	4,993	4,107
Inside option, enrollment, '000	8,772	10,565	12,948	12,640	10,592
Inside option, %	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
Premium sbsd if out. opt., \$mn	6,018	7,248	8,882	8,671	7,266
Reinsurance sbsd if out. opt., \$mn	2,281	2,747	3,366	3,286	2,754
<b>Total surplus, \$mn</b>	<b>1,570</b>	<b>1,474</b>	<b>918</b>	<b>912</b>	<b>1,114</b>
Surplus per dollar PDP subsidy, \$	0.17	0.12	0.06	0.06	0.09
Surplus per opp. cost dollar, \$	1.12	0.59	0.24	0.25	0.29

Table 6: Counterfactual Welfare Estimates: Decentralized Mechanisms

	P= 5% Bid	P= 32% Bid	P= 95% Bid	P= Bid-0	P= Bid-\$721	P= Bid-\$1,340
Consumer Surplus, \$mn	13,344	4,536	1,081	1,073	2,844	13,028
Insurer Profit, \$mn	50,403	4,846	20	13	960	4,272
Premium subsidy, \$mn	84,144	18,087	10	-	7,301	31,290
Reinsurance subsidies, \$mn	15,695	8,191	59	39	3,932	9,236
Inside option, enrollment, '000	23,280	14,506	134	90	10,127	23,351
Inside option, %	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
Premium sbsd if out. opt., \$mn	15,970	9,951	92	62	6,947	16,019
Reinsurance sbsd if out. opt., \$mn	6,053	3,772	35	23	2,633	6,071
<b>Total surplus, \$mn</b>	<b>(37,414)</b>	<b>(6,939)</b>	<b>1,177</b>	<b>1,147</b>	<b>1,654</b>	<b>(6,667)</b>
Surplus per dollar PDP subsidy, \$	(0.37)	(0.26)	17.17	29.38	0.15	(0.16)
Surplus per opp. cost dollar, \$	(0.37)	(0.43)	(15.50)	(19.11)	0.77	(0.28)

Table 7: Counterfactual Welfare Estimates: Normative Analysis

	P=MC	P=SMC	Social Planner	SP end. out. opt.	Optimal Voucher	t-specific Voucher
Consumer Surplus, \$mn	1,094	1,064	2,937	1,942	1,824	1,912
Insurer Profit, \$mn	-	-	(10,055)	(7,148)	516	547
Premium subsidy, \$mn	-	-	-	-	2,870	3,235
Reinsurance subsidies, \$mn	93	-	-	-	1,866	1,894
Inside option, enrollment, '000	217	35	11,192	8,786	4,783	5,136
Inside option, %	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
Premium sbsd if out. opt., \$mn	149	24	7,677	5,711	3,281	3,523
Reinsurance sbsd if out. opt., \$mn	57	9	2,910	2,284	1,244	1,335
<b>Total surplus, \$mn</b>	<b>1,241</b>	<b>1,107</b>	<b>3,628</b>	<b>3,044</b>	<b>2,066</b>	<b>2,106</b>
Surplus per dollar PDP subsidy, \$	13.30	-	0.28	0.33	0.44	0.41
Surplus per opp. cost dollar, \$	(8.49)	(25.93)	(5.25)	(2.76)	7.54	5.97

## 8 Appendix

### 8.1 Derivation of Social Planner's Pricing Problem

Social planner's problem:

$$\begin{aligned}
 W^{SP}(p) = & \overbrace{\int \frac{1}{\alpha} \left( \gamma + \ln \left[ 1 + \sum_{j=1}^J \exp(v_j(\theta), p_j) \right] \right) dF(\theta)}^{\text{Consumer Surplus (CS)}} + \\
 & \lambda \left[ \overbrace{\sum_{j=1}^J (p_j - c_j) s_j(p)}^{\text{Product Market Surplus}} - \overbrace{\left( \sum_{j=1}^J (G_j^{PDP} - G^{MAPD}) s_j(p) \right)}^{\text{Cost of Subsidies (G)}} \right]. \quad (19)
 \end{aligned}$$

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:

$$\frac{\partial CS(p)}{\partial p_j} = \int \frac{1}{\alpha} \left[ \frac{-\alpha \exp(v_j(\theta))}{1 + \sum_{k=1}^J \exp(v_k(\theta))} \right] dF(\theta) = -s_j(p). \quad (20)$$

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

$$\frac{\partial \Pi(p)}{\partial p_j} = \lambda s_j(p) + \lambda \sum_k (p_k - c_k) \frac{\partial s_k(p)}{\partial p_j}. \quad (21)$$

The derivative of government spending with respect to  $p_j$  is:

$$\frac{\partial GS(p)}{\partial p_j} = -\lambda \left[ \sum_k (G_k^{PDP} - G^{MAPD}) \frac{\partial s_k(p)}{\partial p_j} \right], \quad (22)$$

$$= -\lambda \left[ \sum_k \Delta G_k \frac{\partial s_k(p)}{\partial p_j} \right]. \quad (23)$$

Summing these terms, we obtain:

$$\frac{\partial W(p)}{\partial p_j} = (\lambda - 1)s_j + \lambda \sum_k (p_k - c_k) \frac{\partial s_k(p)}{\partial p_j} - \lambda \left[ \sum_k \Delta G_k \frac{\partial s_k(p)}{\partial p_j} \right], \quad (24)$$

$$= (\lambda - 1)s_j + \lambda \sum_k (p_k - c_k - \Delta G_k) \frac{\partial s_k(p)}{\partial p_j}. \quad (25)$$

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:

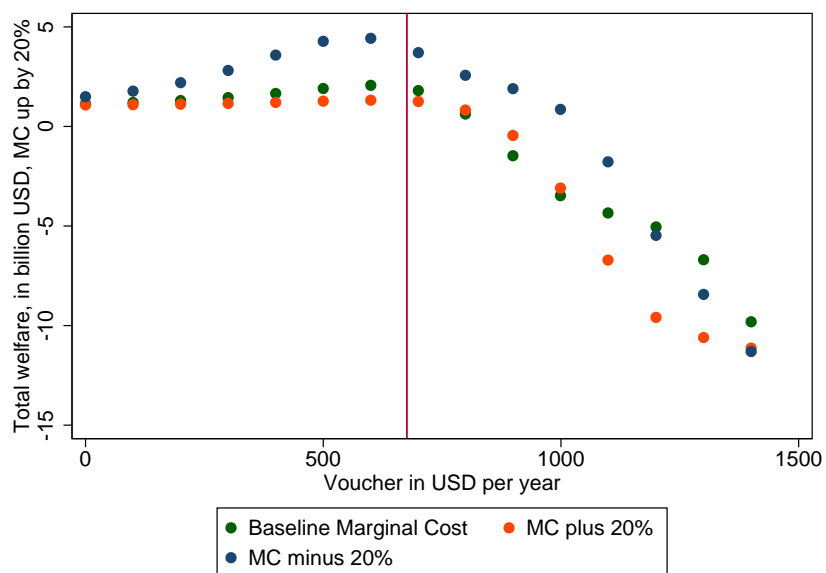
$$(\lambda - 1)s(p) + \lambda \Omega(p)(p - c - \Delta G) = 0, \quad (26)$$

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

$$\Omega_{ij}(p) = \frac{\partial s_j(p)}{\partial p_i}. \quad (27)$$

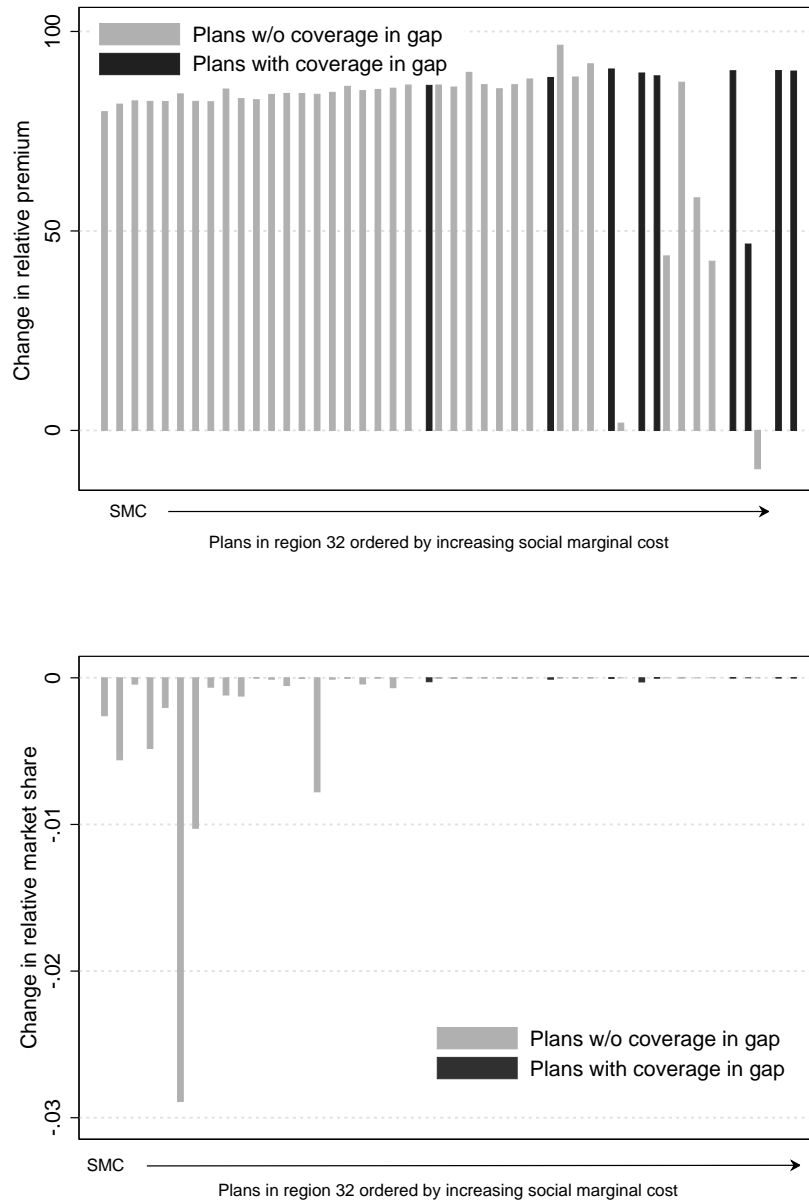


Figure 9: Welfare with flat subsidies: uniform optimal vouchers with  $\pm 20\%$  MC interval



Plots estimated total welfare in full equilibrium counterfactuals at different levels of voucher-like subsidies using three sets of marginal cost estimates: 1) baseline - green dots; 2) baseline plus 20%; 3) baseline minus 20%.

Figure 10: Social planner solution with endogenous outside option



Example market: California in 2010. Panel 1. Plots the diff-in-diff between premiums set by social planner with endogenous outside option and observed premiums, vs premium difference between baseline social planner and observed premiums, as plotted in Figure 8. Panel 2. Plots the diff-in-diff between enrollment shares in allocations by social planner with endogenous outside option net of observed shares, vs share difference between baseline social planner and observed premiums, as plotted in Figure 8.