

NBER WORKING PAPER SERIES

EXTERNALITIES AND BENEFIT DESIGN IN HEALTH INSURANCE

Amanda Starc
Robert J. Town

Working Paper 21783
<http://www.nber.org/papers/w21783>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2015, Revised April 2018

Previously circulated as "Internalizing Behavioral Externalities: Benefit Integration in Health Insurance." The authors gratefully acknowledge funding from the Leonard Davis Institute. Michael French, Josh Gottlieb, Matt Grennan, Ben Handel, Jonathan Ketcham, Kurt Lavetti, Maria Polyakova, Joshua Schwartzstein, Ashley Swanson, and participants at the American Health Economics Conference, FTC Microeconomics Conference, NBER Insurance Meetings, Kellogg Healthcare Markets Conference, Wharton IO lunch, University of British Columbia, Indiana University, University of Chicago, and Yale provided helpful comments. Hossein Alidaee, Emma Boswell Dean, Jordan Keener, and David Stillerman provided excellent research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at <http://www.nber.org/papers/w21783.ack>

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2015 by Amanda Starc and Robert J. Town. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Externalities and Benefit Design in Health Insurance
Amanda Starc and Robert J. Town
NBER Working Paper No. 21783
December 2015, Revised April 2018
JEL No. I13,L2

ABSTRACT

Insurance plan design has important implications for consumer welfare. In this paper, we model insurance design in the Medicare prescription drug coverage market and show that strategic private insurer incentives impose a fiscal externality on the traditional Medicare program. We document that plans covering medical expenses have more generous drug coverage than plans that are only responsible for prescription drug spending, which translates into higher drug utilization by enrollees. The effect is driven by drugs that reduce medical expenditure and treat chronic conditions. Our equilibrium model of plan design endogenizes plan characteristics and accounts for selection; the model estimates confirm that differential incentives to internalize medical care offsets can explain disparities across plans. Counterfactuals show that strategic insurer incentives are as important as selection in determining endogenous plan characteristics.

Amanda Starc
Kellogg School of Management
Northwestern University
2001 Sheridan Road
Evanston, IL 60208
and NBER
amanda.starc@kellogg.northwestern.edu

Robert J. Town
Department of Economics
The University of Texas at Austin
2225 Speedway
BRB 1.116, C3100
Austin, TX 78712
and NBER
robert.town@austin.utexas.edu

1 Introduction

The welfare generated by private health insurance critically depends on the structure of benefit packages offered by insurers. More generous benefits provide increased enrollee risk protection but are costly to the risk-bearing insurer. Increasing generosity mechanically increases plan expenditures and, more important from a welfare perspective, increases the likelihood of moral hazard and adverse selection. An optimal insurance plan must balance these gains from risk protection against inefficiencies due to asymmetric information; a large theoretical and empirical literature in economics analyzes these issues.

Less well-studied are insurers' strategic responses to other incentives that shape plan design. In this paper, we focus on a potentially important friction in the equilibrium determination of insurer plan characteristics. Patient care often spans different treatment modalities (e.g. inpatient, office visits, outpatient surgery, specialist care, pharmaceuticals), and coverage for one type of care may interact and spill over to other services, creating the potential for an externality (Goldman and Philipson (2007); McGuire (2011); Goldman, Joyce and Zheng (2007)). We focus on a classic example: a substantial body of evidence shows that more generous pharmaceutical coverage increases drug adherence, preventing future inpatient utilization. Unless insurers are responsible for coverage across linked treatment modalities, they will have limited incentives to internalize this externality in their benefit design decisions. The welfare impact of these benefit design decisions are potentially large, as the level and composition of consumption of health care services depend on insurer benefit design. We model plan design and estimate the impact of this externality by studying the mandated separation of covered benefits categories for private insurers providing services in the United States Medicare system in Medicare Part D.

The Medicare Part D program provides prescription drug coverage to beneficiaries through private plans that are publicly financed. In 2015, over 39 million Medicare beneficiaries signed up for a Part D plan, accounting for \$137 billion in drug spending. Under the Medicare Part D program, there are two major categories of drug plans: stand-alone prescription drug plans (PDPs) and Medicare Advantage Prescription Drug (MA-PD) plans. Stand-alone PDPs are mandated to cover only pharmaceutical expenditures, while MA-PD plans cover both drug and medical expenditures. These differences imply that the two types of plans face different benefit design incentives. Stand-alone PDPs have an incentive to minimize drug expenditures, ignoring the impact on medical spending, while MA-PD plans have an incentive to minimize overall medical and drug expenditures taking externalities from drug consumption to medical care utilization into account. As a result, MA-PD plans have an incentive to provide more generous coverage for drugs – particularly for those drugs for which increased adherence reduces medical expenditures.

Our primary data source is the rich Medicare Part D prescription drug event data. We observe every prescription fill for the years 2006-2009 for a random 10% sample of all Medicare eligibles. These data contain information on the specific drug filled, retail price, enrollee out-of-pocket cost (OOPC), and fill date for over 123 million drug

claim events. We supplement these data with information on beneficiary and plan characteristics. The beneficiary data contain information on enrollee demographics and the plan enrollment detail. The plan data contain detailed information on the premiums and benefit design (e.g., which drugs are on each benefit tier and the coinsurance/co-payment structures of each tier).

We begin our empirical analysis by comparing the benefit designs of PDP and MA-PD plans. The comparison is challenging because there are over 35,000 unique pharmaceutical products for which plans need to determine cost-sharing, and plans generally employ a complex nonlinear cost-sharing design. Over our sample period, the Part D standard benefit package included a deductible, an initial coverage region where enrollee cost-sharing is limited, the donut hole where the enrollee is responsible for 100% of the cost of their drugs, and finally the catastrophic region where the enrollee is responsible for only a small fraction of drug cost. Using the detailed prescription level data with over 123 million claims, we estimate that MA-PD plan enrollees on average pay between 11 percent less out-of-pocket on average for the identical bundle of drugs in the donut hole. The differences across plan type are larger for sicker enrollees. Furthermore, consistent with externalities playing a key role, the out-of-pocket price differentials are largely driven by drugs that have been explicitly identified as having large medical care offsets and treat chronic conditions like asthma, diabetes, and high cholesterol. We then examine enrollees' behavioral response to these price differentials and calculate that they lead to MA-PD spending \$99 per enrollee more on prescriptions than their PDP counterparts.

We then turn to our main empirical exercise: specifying and estimating the structural parameters of an oligopoly model of premium and benefit design choice. In order to capture insurer incentives, we model both consumer plan choice and insurer benefit design. Importantly, our model allows for drug expenditures and preferences to vary across consumers, and captures the extent to which differences in generosity by plan type can be rationalized by consumer demand and selection. We also allow strategic insurer incentives to vary by plan type. The model recovers cost and demand side parameters, enabling us to understand the economic rationale behind increased prescription drug benefit generosity in MA-PD plans. Consistent with previous work, the demand side estimates imply that consumer responsiveness to plan generosity when choosing plans is modest. The model parameters imply that the increased generosity of MA-PD plans is driven by both asymmetric information and insurer cost side incentives. Consistent with importance of offsets, MA-PD plans find it less costly to increase generosity than their stand-alone counterparts.

Using the model and parameter estimates, we measure the impact of various economic forces on benefit design, including the effect of plans internalizing the externalities generated by offsets. We find substantial medical care offsets in MA-PD plans: a \$1 increase in prescription drug spending reduces non-drug expenditure by approximately 27 cents.¹ If stand-alone PDPs are forced to account for this externality in their premiums and benefit design behavior,

¹These offsets of medical care costs are viewed as sufficiently important to be included in government budget forecasts of health care expenditures, and are consistent with previous estimates. This estimate aligns with previous work by Chandra, Gruber and McKnight (2010), who examine offsets using demand-side consumption. We cannot employ a similar strategy because we do not observe medical claims for enrollees in MA-PD plans.

they would increase insurer drug spending by 8%. Critically, we show that adverse selection also has an impact on benefit design: absent selection, insurer drug spending would increase by an additional 7%. Based on these estimates, we find that stand-alone Part D plans impose a \$378 million externality on traditional Medicare each year. In contrast to a large literature focused on the dead-weight loss due to moral hazard and over-consumption of medical services, our paper finds evidence of potential under-consumption, but shows that insurers will internalize offsets if incentivized to do so.

In addition to explicitly modeling incentives to internalize offsets, our model allows for both asymmetric information and behavioral biases. Specifically, we account for selection and screening incentives (Geruso, Layton and Prinz (2016); Carey (2017); Lavetti and Simon (2014)) and moral hazard (Finkelstein, Einav and Polyakova (2016)). We also explore the impact of behavioral biases including inertia (Ho, Hogan and Scott Morton (2015); Polyakova (2016)), and explore the impact of choice frictions and the under-utilization of cost-effective care (Abaluck and Gruber (2011); Ketcham et al. (2012)) and Manning et al. (1987), Brot-Goldberg et al. (2017)). Critically, our modeling approach accounts for imperfect competition among insurers (Decarolis, Polyakova and Ryan (2015)) and extends existing models that endogenize prices but hold product characteristics fixed (Handel (2013); Lustig (2010); Starc (2015); Town and Liu (2003); Tebaldi (2017); Ericson and Starc (2015)). By incorporating all of these features, we can explore the implications of endogenous product benefit design in the Medicare Part D market, discuss policy-relevant counterfactuals, and provide a framework for future researchers.

Our work also expands on the recent literature examining insurer competition in private Medicare markets (e.g. Decarolis, Polyakova and Ryan (2015); Curto et al. (2015)); more broadly, we contribute to a recent and growing literature on endogenous product design (see Fan (2013) as well as Crawford (2012) for a review).² The paper is organized as follows. Section 2 describes the market, and Section 3 presents the reduced form estimations. Section 4 describes and estimates our model of firm behavior. Section 5 presents counterfactual exercises that put the magnitude of our effect in context, and Section 6 concludes.

2 Empirical Setting and Data

In this section, we describe the role of private insurers within Medicare, which provides health insurance to the elderly in the United States.³ Private insurers play an important role in administering benefits; as a result, our setting is a data rich – if institutionally complex – laboratory in which to explore endogenous product design. Medicare Parts A and B cover inpatient and outpatient services, respectively. Medicare Advantage (Part C) and Part D are administered

²Fan (2013) is the closest to our setting, as she explores continuous quality attributes. See also Draganska, Mazzeo and Seim (2009); Eizenberg (2014); Sweeting (2010); Wollman (2016).

³Medicare also provides health insurance coverage for the disabled and those with End Stage Renal Disease. We do not focus on those populations in this paper.

by private insurers. Medicare Advantage is an alternative to traditional Medicare under Parts A and B, and Medicare Part D covers prescription drugs. Under both Medicare Advantage and Part D, Medicare beneficiaries are given information on the plan's premiums and benefit design and can select into any of the available plans in the area; competitive pressures should motivate insurers to offer low premium and cost-efficient products.

2.1 Private Plans and Medicare

Medicare Part C, the first broad private insurance option available to Medicare beneficiaries, was created under the Tax Equity and Fiscal Responsibility Act in 1982. Over its history, the program has gone by a variety of names (see McGuire, Newhouse and Sinaiko (2011) for a comprehensive history), and is currently known as Medicare Advantage. Medicare Advantage plans give Medicare beneficiaries the option to forego traditional Medicare and enroll in a private insurance plan for their health care benefits. MA plans are attractive because they typically offer more generous coverage. For each beneficiary that it enrolls, the plan receives a risk-adjusted, per-capita payment from the Centers for Medicare and Medicaid Services (CMS). The plans also earn revenue from premiums paid directly by enrollees.⁴ The program's popularity has waxed and waned over time, coinciding with the generosity of federal subsidy. As of 2009, the last year of our sample, 24% of all Medicare beneficiaries and 23% of Part D beneficiaries were enrolled in a Medicare Advantage plan.⁵ There is significant geographic and demographic heterogeneity in the popularity of MA-PD plans: MA-PD plans are typically more attractive to middle and lower income as well as healthier beneficiaries within a market. Finally, the typical MA market is concentrated. In 2008, the largest four carriers had 45% of total MA enrollment.⁶

Premiums and benefit generosity in MA are determined through the plan's "bid" (the dollar amount the plan estimates will cover Part A and B benefits for a beneficiary in average health) and the county-level benchmark. If the plan's bid is above the benchmark, the plan's payment from the government is the benchmark plus the premium, which is the difference between the bid and the benchmark. If the plan's bid is below the benchmark, the plan's payment is their bid plus 75% of the difference between the bid and the benchmark. The plan must use the payment between the bid and the benchmark (the "rebate") to fund benefit enhancement for enrollees. Benefit enhancements include reductions in medical care cost-sharing, provision of added, non-Medicare benefits such as dental coverage, increased

⁴During our sample period, there were two types of MA plans – those with Part D coverage bundled in (MA-PD) and those that do not cover pharmaceuticals. MA-PD plans receive an additional Part D subsidy from the government and a premium payment from the enrollee. In 2009, the vast majority of MA plan beneficiaries (82%) were enrollees in a MA-PD plan.

⁵During our entire sample period, from 2007-2009, approximately 1 in 4 beneficiaries was enrolled in a MA-PD plan. Enrollment rates have continued to grow post-Affordable Care Act (ACA).

⁶The MA program is important from a policy perspective due to its sheer size in terms of enrollees and budget impact, but despite its popularity among beneficiaries, the MA program has always been controversial. There is substantial debate about the level of spending in MA as compared to traditional Medicare; cherry-picking by MA plans could lead to over payment by the federal government or skew benefit design to attract favorable risks (Brown et al. (2014); Carey (2017)). Furthermore, a more recent literature argues that a substantial portion of the private gains from the MA program accrue to insurers, though the exact magnitude is a matter of debate (see Cabral, Geruso and Mahoney (2014); Curto et al. (2015); Duggan, Starc and Vabson (2016)). By contrast, a number of papers highlight the potential for better medical management under MA (Afendulis et al. (2011)). There is also evidence that the benefits of Medicare Advantage may spill over to traditional Medicare beneficiaries (Baicker, Chernerw and Robbins (2013)).

generosity of the drug benefit, and reduction of additional premiums. The payments made to the plan are ultimately risk-adjusted based on the expected average cost of the plan's enrollment. MA-PD plans also submit a separate bid for the Part D component and the payments that flow from that bid follow the Part D rules discussed below.

The Medicare Part D program, enacted under the Medicare Modernization Act in 2003, was introduced in 2006. Medicare beneficiaries can enroll in a private insurance plan that provides prescription drug coverage. For most Medicare beneficiaries, there are two ways to obtain Part D coverage. They can enroll in a stand-alone plan (PDP) that only covers prescription drugs or they can enroll in a MA-PD plan. Typically, enrollees in PDPs receive their medical coverage from traditional Medicare. Outside of the direct impact on plan enrollment, the PDP plans have little incentive to consider the influence of their benefit design decisions on enrollee medical care utilization. Part D is also heavily subsidized and because of this subsidy it is financially beneficial for most Medicare beneficiaries to enroll in some form of drug coverage.

The program requires insurers to provide coverage at least as generous as the standard benefit, which has a nonlinear structure in which the beneficiary pays differing out-of-pocket costs depending on the phase of the benefit design. The deductible in 2008 was \$275, followed by 25% cost-sharing in the initial coverage region (ICR) up to \$2510 of expenditure, followed by the infamous donut hole phase where the enrollee incurs the entire cost of drug expenditures and, finally, catastrophic coverage where the enrollee faces a 5% coinsurance rate. Despite the large number of plan offerings typically available, markets are typically concentrated. Over 50% of Part D beneficiaries enroll in plans offered by three carriers.

While the strict regulation of Part D plans creates a minimum standard for plans, PDP and MA-PD plans can provide more generous coverage than the minimum. In fact, the majority of plans in our sample offer coverage more generous than the standard benefit. The majority of these plans eliminate the deductible, and nearly one quarter of MA-PD plans had some form of donut hole coverage in 2006.⁷ In addition to providing coverage that is at least actuarially equivalent to the standard benefit, plans must cover all or substantially all drugs within six protected drugs classes and two or more drugs in another 150 categories. The set of PDP plans available depends on which of the 34 regions an enrollee lives in, while the set of MA-PD plans available depends on the county of residence.

Like MA plans, Part D plan premiums and government payments are determined through plan bids. The premium subsidy, paid by CMS, is also calculated using a formula that averages over plan bids. Premiums are calculated as the difference between the bid and the subsidy paid by CMS. To mitigate adverse selection, CMS employs a three-pillar risk equalization system within Medicare Part D. First, the government provides individual reinsurance during the catastrophic phase of the standard benefit, covering 80% of drug expenditure after an individual has incurred substantial drug costs. Second, risk adjustment attempts to equalize insurer profitability across beneficiaries by increasing

⁷By contrast, only 6% of PDP plans had donut coverage in 2006. The donut hole is being phased out as a part of the ACA. See Hoadley et al. (2014) for additional details.

subsidies for sicker enrollees. Despite this, there may still be selection conditional on the risk adjustment (Brown et al. (2014); Carey (2017)). Third, risk corridors provide down-side protection against plan-level losses and cap plan-level profit margins. Finally, CMS provides additional subsidies to a subset of beneficiaries through the low-income subsidy (LIS) program.⁸

To summarize, during our sample period, a senior eligible for Medicare had multiple private insurance choices. They could opt out of traditional Medicare and into a Medicare Advantage plan and the private Medicare Advantage insurer would be responsible for all medical spending. The federal government would pay the insurer a fixed subsidy payment per month for both the medical and prescription drug portion of the plan benefit; enrollees may be required to pay a premium as well. By contrast, the beneficiary could instead remain in traditional fee-for-service (FFS) Medicare and choose to augment Medicare Parts A and B with a Part D plan. The private Part D insurer would cover drug expenditure, while the Medicare program would cover non-drug medical spending directly, including hospitalizations and physician services. The federal government would pay the insurer a fixed subsidy payment per month for the prescription drug benefit; enrollees are required to pay a premium as well. Private insurers in Medicare Advantage have an incentive to take any offsets into account; in this paper, we focus on the behavior of MA-PD plans relative to stand-alone PDPs. As our discussion highlights, the institutional setting in which plans compete for both Medicare Advantage and Part D are complex. Our empirical analysis, in particular our structural demand and supply framework, accounts for this complexity.

2.2 Pharmaceutical Benefits Design and Medicare Care Offsets

An underlying premise of our analysis is that increased pharmaceutical cost-sharing leads to reductions in prescription drug consumption and that decrease in drug consumption leads to an increase in medical care utilization. In this subsection, we review the existing evidence.⁹ Numerous studies have documented the presence of medical care offsets as related to changes in drug benefit design and the importance of considering these offsets in optimal insurance design (Goldman and Philipson (2007); McGuire (2011); Goldman, Joyce and Zheng (2007)). The evidence for meaningful offsets spans a variety of settings including employer-sponsored insurance (Gaynor, Li and Vogt (2007)), the Medicare

⁸LIS eligibles comprise 28% of the total Part D population. They receive a subsidy equivalent to the region specific LIS benchmark and can enroll in any plan. If they enroll in a plan with a premium below the benchmark, they must pay the difference between that benchmark and premium but they still receive the benefit of the subsidized cost-sharing. Importantly, plans that offer premiums below the LIS threshold are eligible for randomized auto-enrollment of LIS beneficiaries. Previous research has highlighted that the presence of the LIS subsidy can distort plan bidding incentives (Decarolis (2015); Decarolis, Polyakova and Ryan (2015)).

⁹A long literature, including the RAND health insurance experiment (Manning et al. (1987)), has shown that increased cost-sharing causally leads to a reduction in the consumption of pharmaceuticals. More recent evidence indicates that these reductions in consumption affect both high- and low-value services (Brot-Goldberg et al. (2017); Baicker and Goldman (2011); Maciejewski, Farley, Parker and Wansink (2010); Maciejewski, Bryson, Perkins, Blough, Cunningham, Fortney, Krein, Stroupe, Sharp and Liu (2010)). Within the Medicare Part D setting, multiple papers (including this one) have exploited the non-linear benefit structure to measure the behavioral response to cost-sharing (Abaluck, Gruber and Swanson (2015); Einav, Finkelstein and Schrimpf (2015); Finkelstein, Einav and Polyakova (2016); Dalton, Gowrisankaran and Town (2015)). This literature finds that increased cost-sharing reduces drug consumption and that cost-sharing in the donut hole is especially salient to consumers. Dalton, Gowrisankaran and Town (2015) find enrollees reduce the number of prescriptions filled by 21% upon entering the coverage gap. Furthermore, there is evidence that the introduction of the Part D program is associated with reduced hospital admissions in the Medicare population (Afendulis et al. (2011)).

population (Chandra, Gruber and McKnight (2010) and specifically in the Medicare Part D program (McWilliams, Zaslavsky and Huskamp (2011)). The Congressional Budget Office, based on a survey of the literature, assumes that a 1% increase in drug consumption reduces non-drug medical consumption by 0.2% (CBO (2012)). Cost-sharing can lead to sub-optimal consumption because of discrepancies between private willingness to pay and social marginal cost for a variety of reasons. There may be asymmetric information about the value of treatment (Manning et al. (1987)) or misalignment across multiple technologies (Ellis, Jiang and Manning (2015); Goldman and Philipson (2007)) or payers (see Cabral and Mahoney (2014)). Underutilization of drugs may also be “due to mistakes or behavior biases,” referred to in the literature as behavioral hazard (Baicker, Mullainathan and Schwartzstein (2015)).¹⁰ In sum, there is a large, robust literature documenting that among health care consumers in general and Medicare enrollees specifically, drug adherence is negatively impacted by increased cost-sharing. Furthermore, this reduction in adherence leads to an increased likelihood of utilization of non-pharmaceutical medical care.

2.3 Insurer Incentives and Benefit Design

In this subsection, we provide the conditions under which MA-PD plans have an incentive to offer more generous plan benefits than their stand-alone PDP counterparts. Our framework here is similar to Goldman and Philipson (2007) and McGuire (2011). Consider a model in which insurers choose the premium, p , and enrollee pharmaceutical cost-sharing, x^D . In order to focus attention on drug benefits, we hold constant MA-PD medical care cost-sharing. Premiums and drug plan generosity (which affects plan costs, enrollment and selection) are chosen in order to maximize static profits, taking the subsidy setting rules and, for the MA-PD plans, medical care costs, as given. We first describe endogenous benefit design for a stand-alone prescription drug plan, then expand the analysis to take into account differential incentives faced by MA-PD plans. Variable profits for a stand-alone PDP are given by:

$$\Pi^{PDP} = (p + z - c^D(x^D)) s(\mathbf{p}, \mathbf{x}^D) B, \quad (1)$$

where p is the premium, z is the federal subsidy, $c^D(x^D)$ is the average insurer cost of providing pharmaceuticals to consumers given benefit design x^D , B is the size of the relevant Medicare population, and $s(\mathbf{p}, \mathbf{x}^D)$ is the plan’s market share, which is a function of all plans premiums and cost-sharing decisions. Insurers collect premiums and subsidies and pay out claims for each Medicare beneficiary they enroll (“enrollee”). Insurer costs are a function of endogenous plan characteristics. We assume that $\frac{\partial c^D}{\partial x^D} < 0$ and $\frac{\partial s}{\partial x^D} < 0$. As enrollee cost-sharing increases, the plan’s cost decreases because there is a mechanical relationship between cost-sharing and plan cost. Costs can also decrease in x^D because

¹⁰Within the context of the Part D program, the behavioral bias most frequently explored is myopia (Abaluck, Gruber and Swanson 2015, Dalton, Gowrisankaran and Town 2015). Ex ante, consumers may be naive or sophisticated about the potential for underutilization due to information issues, behavioral biases, or both. A sophisticated consumer will demand an insurance contract that corrects for this underutilization of high-value services to the extent that they value reduced spending or improved health, creating a market for value-based insurance designs (Ellison (2006); Chernen, Rosen and Fendrick (2007)).

of both moral hazard and adverse selection. Share is decreasing in out of pocket costs as consumers prefer plans with less cost-sharing. The first order conditions for any product characteristic is given by:

$$(p + z - c^D) \frac{\partial s}{\partial x^D} - \frac{\partial c^D}{\partial x^D} s(\mathbf{p}, \mathbf{x}^D) = 0. \quad (2)$$

As with any consumer good, insurers must trade off margins and market share. In addition, the presence of asymmetric information in the form of adverse selection and moral hazard implies that all insurers must consider how product characteristics affect their costs.

Medicare Advantage plans face a different set of incentives than stand-alone PDP plans. Consider the choice to increase the generosity of a prescription drug plan. The PDP knows that this will directly increase costs, as they bear a higher percentage of a fixed drug expenditure. In addition, higher generosity plans may attract sicker consumers (adverse selection) and induce enrollees to spend more (moral hazard).¹¹ MA-PD plans will also take these factors into account; in addition, a MA-PD plan must consider the impact that drug expenditure has on overall medical expenditure. If there are offsets, increased drug expenditures lead to reduced (non-drug) medical expenditures. More formally, we can write the profit for the MA-PD as

$$\Pi^{MA} = (p + z - c^D(x^D) - c^M(x^D)) s(\mathbf{p}, \mathbf{x}^D) B. \quad (3)$$

Holding the medical benefit structure fixed, average total costs for a MA-PD are the sum of drug and (non-drug) medical costs: $c^D(x^D) + c^M(x^D)$. MA-PD plans may differ from PDPs in their cost-sharing arrangements if there are offsets between drug benefit generosity and medical costs. This is easily seen by examining the first-order conditions with respect to x^D :

$$(p + z - c^D(x^D) - c^M(x^D)) \frac{\partial s}{\partial x^D} - \left(\frac{\partial c^D}{\partial x^D} + \frac{\partial c^M}{\partial x^D} \right) s(\mathbf{p}, \mathbf{x}^D) = 0. \quad (4)$$

In the presence of offsets $\frac{\partial c^M}{\partial x^D} > 0$; $\frac{\partial c^M}{\partial x^D}$ can be written as the offset (how a dollar of drug expenditure reduces medical expenditure) multiplied by the demand response (how increased plan generosity changes drug expenditure). Comparing the first-order conditions in equations 2 and 4, it is clear that the MA plan faces different incentives when designing their drug benefits. Specifically, they have an incentive to internalize the impact of changes in drug plan generosity on medical care utilization and, all else equal, offer a more generous benefit design; estimating the magnitude of this incentive and the subsequent welfare implications is the focus of the remainder of the paper.

¹¹In this simple framework, we do not allow for premiums to affect plan selection.

2.4 Data

Our primary data are the rich Medicare Part D prescription drug event data. We observe every prescription fill for the years 2006-2009 for a random 10% sample of all Medicare eligibles. For much of our analysis, we aggregate these data to the enrollee-year level. We supplement these data with information on beneficiary and plan characteristics and merge in MA subsidy payment levels and county and metropolitan demographic information.

We begin the construction of our analytic sample by capturing all beneficiaries that were enrolled in a PDP or MA-PD plan between 2007 and 2009. This gives us 7,597,476 enrollee/year observations. We exclude any enrollees who receive low-income subsidies and are subject to lower cost-sharing.¹² This restriction leaves us with 4,802,000 enrollee-year observations. We then drop any enrollees for whom we do not have claims in 2006 to control for previous consumption, leaving us with 3,534,965 enrollee/year observations for our analytical data set.

Summary statistics are presented in Table 1. In the full sample, the average enrollee is 77 years old, 62% are female and 90% are white. Average total annual expenditure is \$1763. There is substantial heterogeneity in enrollee annual drug spending, as highlighted in Figure A.2, which plots a histogram of total spending in both MA-PD and standalone PDP plans in 2008.¹³ While we observe rich data on drug spending, we do not observe non-prescription medical claims for MA-PD enrollees; an important goal of the structural analysis is to compensate for this data limitation by inferring the level of medical expenditures and, importantly, the offset from insurer benefit design decisions.¹⁴

3 Reduced Form Evidence

We begin our empirical analysis by examining whether the differential benefit design incentives described above translate into realized differential benefit designs between PDP and MA-PD plans. We estimate the parameters of the claims-level OLS regression equation given by:

$$\text{Log}(OOPC_{cdjt}) = \alpha_d + \tau_t + \beta 1(MA_{jt}) + \varepsilon_{cdjt}, \quad (5)$$

where $OOPC_{cdjt}$ is out-of-pocket price per day supplied in prescription claim c for drug d in plan j and year t . The parameters α_d and τ_t are drug and year fixed effects. The drug fixed effects are at the National Drug Code (NDC) level, which capture all of the variation related to the detailed product and package (i.e. 20mg of Lipitor).

Table 2 presents the results. In the first specification, we include drug and year fixed effects. In the second, we

¹²While we drop LIS enrollees for our main analysis, we run numerous robustness analyses to test the sensitivity of our findings to supply-side responses to the presence of the LIS population.

¹³There are a couple observations to highlight: first, as expected, there is excess mass at the initial coverage limit, as highlighted by Einav, Finkelstein and Schrimpf (2015). Second, enrollees in MA-PD plans spend substantially less on prescription drugs than PDP enrollees. We control for this observed heterogeneity by controlling for lagged consumption in both our reduced form results and consumer demand system. To do this, we create five consumer types corresponding to the five quintiles of their 2006 spending. Total consumption in the first group averages \$895 per year for 2007-2009, while yearly spending in the top quintile averages \$3503.

¹⁴Given that CMS encrypts the beneficiary identification variable, linking the CMS pharmacy claims to MA claims is not currently feasible.

Table 1: Enrollee Summary Statistics (Means and Standard Deviations)

Total Drug Expenditure	1762.78 [2620.15]
Total Insurer Drug Expenditure	1114.65 [2068.99]
Enrollee OOPC	648.13 [879.49]
Total Rx Days Supply	1302.61 [875.93]
% in MA	0.40 [0.49]
Age	76.87 [7.25]
% Female	0.62 [0.49]
% White	0.90 [0.30]
Observations	3,465,139

Notes: Table presents summary statistics describing mean enrollee demographics, coverage, and consumption. The unit of observation is the enrollee-year. Standard deviations are in brackets.

control for drug, year and the phase of the prescription drug benefit, as insurers can alter enrollee cost-sharing given the benefit structure or the benefit structure itself. The results show a consistent pattern. For MA-PD plans, enrollees face an OOPC per days supply that is 4-7% lower, holding the drug (NDC) constant.¹⁵ Table A.7 shows that the plan's cost per day supplied for a given drug is equal across plan types; negotiated prices are not systematically higher or lower for MA-PD plans and do not explain our empirical results.

We also estimate equation 5 for different samples based on the type of drug. In Figure A.3, we show that the differential cost-sharing results are larger for specific drug classes likely to have large medical care offsets.¹⁶ We find statistically larger effects among drugs used to treat diabetes, asthma, and hyperlipidemia (high cholesterol). In Panel B of Table 2, we present the equivalent regression results. Enrollee OOPC per days supply in MA-PD plans is approximately 10% lower for asthma drugs, diabetes drugs, and cholesterol drugs. The results for hypertension (high blood pressure) are more mixed. In Figure A.4, we show that this is due to heterogeneity across types of hypertensives. For the most cost-effective, recommended initial therapy (non-beta blockers, NICE (2011)), the effect is in the expected direction. Finally, to examine how these OOPC differences manifest across the benefit phases, we estimate Equation 5 restricting the sample to deductible and ICR claims and donut hole claims. Panel C of Table 2 shows that the OOPC differences are most pronounced in the donut hole. Going forward, we will focus on enrollee

¹⁵In the appendix, we show some evidence that enrollees in MA-PD plans are more likely to fill 90-day prescriptions, which likely contributes to increased adherence; the estimates imply that 1.4% more prescriptions are 90-day fills under MA-PD plans, making the effect small, but still indicative of differential strategies by plan type. In unreported regressions, we find that the OOPC cost for hyperlipidemia drugs in MA-PD plans is 12-15% lower than in PDP plans, consistent with lower out-of-pocket costs for drugs for chronic conditions.

¹⁶Specifically, we targeted by value-based insurance designs in the commercial insurance market (Chernew, Rosen and Fendrick (2007); Gowrisankaran et al. (2013)).

costs in the initial coverage range – which composes the bulk of fills and represents the effective marginal price for most enrollees – and in the donut hole – where we observe substantial variation and which is especially salient to enrollees.

Table 3 aggregates to plan level characteristics. The first row shows that PDP plans are slightly more likely to have a deductible than MA plans (though the difference is not statistically significant). The next two rows show that the average costs – per prescription or per day supplied – are significantly lower in MA plans. However, the substantial difference in costs per day conflates lower costs for identical drugs and a different mix of drugs among MA and PDP enrollees. The latter is especially likely to be important, as MA enrollees tend to be healthier on average and may take lower cost drugs. Therefore, we characterize plans by a drug premium p_{jmt}^D and a cost-sharing schedule x_{jmt}^D . Each element of the vector is defined as a weighted average of enrollee OOPC per days supply using national consumption weights, as defined below.

To create each plan design variable, we construct an average price per days supply for each product d in each phase-plan j specific combination in year t denoted by P_{jmt}^{Phase} . Plan- and phase-specific enrollee price measures, $\pi_{d,jmt}^{Phase}$, do not depend on the composition of enrollee consumption within that plan. To capture average, national levels of consumption, we average the days supply by drug-year combination to create q_{dt} . The weighting allows us to construct a measure of enrollee OOPC that does not depend on enrollee behavior:

$$P_{jmt}^{Phase} = \sum_d \pi_{d,jmt}^{Phase} q_{dt}. \quad (6)$$

Our measure of plan generosity captures the average OOPC per day for the average Medicare beneficiary. We focus on enrollee costs in the initial coverage range (ICR) and donut hole: $x_{jmt}^D = \left[P_{jmt}^{ICR}, P_{jmt}^{Donut} \right]'$.¹⁷ Our construction nests formulary inclusion, tiering, and coinsurance levels, but does not allow substitution if, for example, a particular drug was excluded from an alternative formulary. While formularies are discrete, we create two continuous choice variables – P_{jmt}^{ICR} and P_{jmt}^{Donut} – for tractability and explore alternative constructions in robustness checks.

Table 3 also describes summary statistics for each of these variables. Cost-sharing is lower in MA-PD plans, especially in the donut hole, where the average enrollee OOPC per days supply is 11% lower (\$1.71 versus \$1.93 for PDP plans). Figure 1 shows the vast majority of stand-alone enrollees do not have any gap coverage, while over half of MA-PD enrollees have at least some gap coverage by the end of our sample. Furthermore, we show in Appendix Table A.1 that it is costly for MA-PD firms to increase the generosity of their drug benefit using our enrollee sample. Once we account for the average consumption bundle, MA-PD plans also have lower cost-sharing in the initial coverage phase (46 cents versus 50 cents) and lower premiums. These differences are smaller than those that do not correct for the composition of drugs consumed, but also indicate that MA-PD plans are likely to be more generous and have flatter

¹⁷We will refer to the vector of product characteristics in the market as x_{mt} .

Table 2: Estimates of the Relationship between Plan Enrollment and Out-of-Pocket Costs

	Dependent Variable: Logged OOPC/Day	
Panel A: Main Results	(1)	(2)
1(MA)	-0.0688*** (0.00323)	-0.0419*** (0.00344)
Constant	-1.027*** (0.00235)	-2.221*** (0.00585)
Observations	123,035,098	123,035,098
Adjusted R-Squared	0.610	0.678
Panel B: By High Offset Class		
1(MA)	-0.0574*** (0.00415)	-0.0321*** (0.00438)
1(MA)*Asthma	-0.0747*** (0.0173)	-0.0833*** (0.0178)
1(MA)*Hypertension	0.0262** (0.00813)	0.0351*** (0.00858)
1(MA)*Diabetes	-0.0758*** (0.0121)	-0.0752*** (0.0123)
1(MA)*Cholesterol	-0.0881*** (0.0116)	-0.0851*** (0.0122)
Observations	123,035,098	123,035,098
Adjusted R-Squared	0.610	0.678
Product Fixed Effects	X	X
Phase Fixed Effects		X
	ICR or Deductible	
Panel C: By Benefit Phase	(Ded Amt. = 0)	Donut Hole
1(MA)	0.00944* (0.00367)	-0.296*** (0.00507)
Constant	-1.182*** (0.00291)	-0.287*** (0.00333)
Observations	96,758,755	17,210,240
Adjusted R-Squared	0.680	0.646

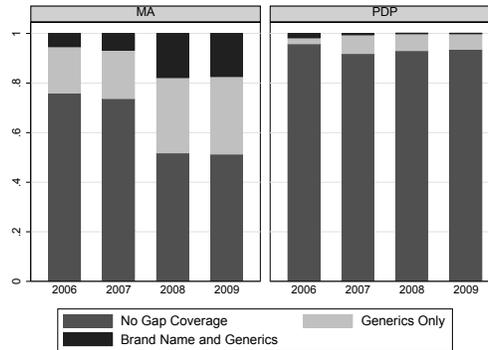
Notes: Table presents linear regression models with logged out-of-pocket costs per day supply as the dependent variable. The unit of observation is at the fill level (weighted by days supply), for the 2007-2009 period. The original data are obtained from a 10% sample of CMS prescription drug event files. We include year-level indicators and product fixed effects in all specifications. In some specifications, we also control the phase of the standard Part D benefit. Standard errors are clustered at the plan-product level. Statistical significance at the 10%, 5%, and 1% levels are denoted by *, **, and *** respectively.

Table 3: Plan Summary Statistics

	PDP	MA
I(Deductible)	.1912	.1655
Mean Enrollee Cost Per Prescription	20.31	15.57
Mean Enrollee Cost Per Days Supply	0.585	0.419
P^{ICR}	.5026	.4608***
	[.0062]	[.0051]
P^{Donut}	1.93	1.71***
	[.0306]	[.0213]
Premium	23.16	12.77***
	[.5507]	[.3182]
Observations	381	1926

Notes: The unit of observation is the year-plan. Mean enrollee cost per prescription and days supply are calculated given observed utilization levels. P^{ICR} and P^{Donut} are calculated for a standardized population using claims data; for stand-alone plans, we aggregate across markets to the contract level. Deductible and premium information is taken from the Part D Plan Characteristics file. In the lower panel, standard deviations are in brackets. Statistically different means at the 1% level denoted by ***.

Figure 1: Percentage of Enrollees with Gap Coverage by Plan Type

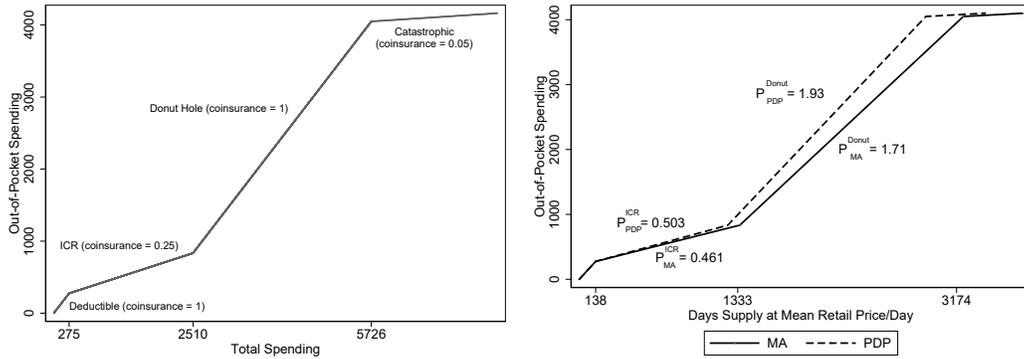


Notes: Figure constructed from plan characteristics data and author calculations.

cost-sharing schedules than their PDP counterparts. Figure 2 summarizes the results of differences between PDP and MA-PD plans in their out-of-pocket prices. The left panel depicts the standard benefit, and the right panel shows the mean structure by plan types. The figure shows that differences in OOPC are larger for sicker enrollees, for whom offsets are likely larger in magnitude. An enrollee in the lowest spending quintile would face OOPC 1.5% lower in the average MA-PD plan; for an enrollee in the highest spending quintile, the difference is 8.6%.

Armed with evidence of meaningful out-of-pocket price differentials between PDP and MA-PD plans, we turn attention to describing the behavioral response by consumers. There are several challenges to estimating the demand response including the potential for enrollee selection and the role of dynamic consumption decisions given the nonlinear structure of the benefit design. We address these challenges using the approaches outlined in Einav, Finkelstein and Schrimpf (2015) and Finkelstein, Einav and Polyakova (2016) and Dalton, Gowrisankaran and Town (2015). Einav, Finkelstein and Schrimpf (2015) use bunching at the kink to measure the behavioral response to cost sharing. Dalton,

Figure 2: Plan Design



Gowrisankaran and Town (2015) show that given the nonlinear benefit structure they face, Part D enrollees' prescription drug filling behavior is not consistent with standard models of dynamic drug consumption; the authors estimate a dynamic model of consumer price salience that explains this behavior.

Two plots summarize the behavioral response of consumers. The first describes bunching at the kink, similar to Einav, Finkelstein and Schrimpf (2015), which focuses on the extensive margin (any fill in the last month of the year). The left panel of Figure 3 shows a consistent pattern across plans, with a similar quantity response in stand-alone and MA-PD plans; the empirical pattern implies a larger elasticity among MA-PD consumers, who face a smaller price increase. To examine the differential response to hitting the donut hole, we focus on a sub-sample of enrollees who start a week with total drug spending between \$2000 and \$2510 following Dalton, Gowrisankaran and Town (2015).¹⁸ We plot weekly spending, which captures both intensive and extensive margin substitution, as a function of total spending relative the location of the donut hole. The right panel of Figure 3 shows that consumers reduce overall consumption as they approach the donut hole; furthermore, there is a discontinuous drop in consumption upon entering the donut hole.

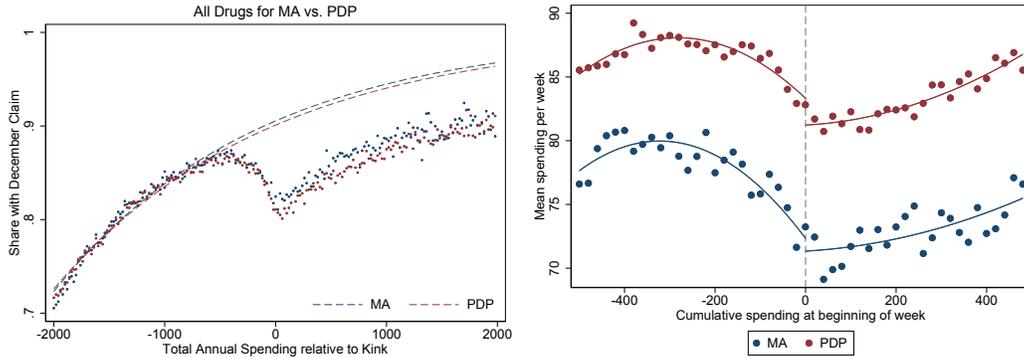
Both approaches show that consumers respond to higher cost sharing by reducing drug consumption. In the appendix, we calculate elasticities using the methods described in Einav, Finkelstein and Schrimpf (2015) and Dalton, Gowrisankaran and Town (2015) and the variation describe above. We estimate elasticities between -0.53 among stand-alone consumers and -0.79 (Dalton, Gowrisankaran and Town (2015) approach), consistent with the Einav, Finkelstein and Schrimpf (2015) approach.¹⁹

Given that consumers respond to OOPC, we expect plan design to affect consumption. The reduced form estimates show that OOPC is 5-7% lower on average in MA-PD plans, with larger differences among drugs for asthma, diabetes, hypertension, and hyperlipidemia. Given the elasticities we measure, our results imply an increase in drug

¹⁸We focus on the period starting on the last Sunday of March and ending on the second to last Sunday of July. The logic behind this sample selection criterion is that the shock that results in the enrollee entering the donut hole should not change the enrollee's expectation about the likelihood of entering into the donut hole during the year. That likelihood is very close to one.

¹⁹We calculate similar differences between predicted and actual probabilities of purchase in December.

Figure 3: Behavioral Response



consumption of 6%, with larger impacts among high offset drugs; our reduced form results imply an additional \$99 of drug spending in MA-PD plans.²⁰ Using the Congressional Budget Office estimates of offsets (a 1% increase in drug consumption reduces non-drug medical consumption by 0.2%), the increase in drug consumption leads to a 1.2% reduction in medical spending. Given that our results are concentrated in high offset drugs, we expect the true effect to be larger in magnitude. Of course, MA-PD insurers have additional tools to increase consumption; for example, Table A.7 shows suggestive evidence of additional use of 90 day fills in MA-PD plans, which may improve medication adherence. In the appendix, we show that the causal effect of MA-PD enrollment is to both decrease overall out-of-pocket expenditure and increase total drug expenditure, as well as the fraction of expenditure paid for by insurers.²¹ The effect is larger in magnitude (13%) but not statistically different from the effect that would be predicted by price differences alone. Taken together, the results in this section show a consistent pattern. MA-PD plans are designed in ways that reduce enrollee cost-sharing and, to a lesser extent, increase drug consumption. The effect is concentrated in drugs likely to generate large offsets. We next examine the incentives MA carriers face to internalize offsets using a model of insurer benefit design.

4 An Oligopoly Model of Premiums and Benefit Design

In this section, we describe our empirical model of equilibrium insurer benefit design and outline our estimation strategy. We estimate the structural parameters of the model in order to (1) decompose demand and cost side rationales for MA-PD plans to offer more generous drug coverage; (2) provide estimates of the implied externality of increased drug coverage and the magnitude of the offset; and (3) perform policy counterfactuals. Our model is simple enough to be tractable yet rich enough to capture the complexity of equilibrium insurer behavior when setting premiums and benefits. We describe the empirical model, then describe our demand specification and parameter estimates, and then

²⁰A 7% decrease in OOPC and an elasticity of -0.79 implies a 5.6% increase in consumption; average consumption (taken from Table 1 is \$1762.78 per year.

²¹This causal effect would not be detected via an ordinary least squares regression, which is likely to be contaminated by selection.

estimate the key cost parameters.

4.1 Empirical Model

Our empirical model builds on the framework described in Section 2.3; however, several institutional complexities require modifying the illustrative model. Following Decarolis, Polyakova and Ryan (2015), we explicitly incorporate plan bidding and the risk stabilization programs. We assume that the expected profits for stand-alone PDPs are a function of the vector of bids for drug coverage, $\mathbf{b}_{\mathbf{m}t}^{\mathbf{D}}$, and plan characteristics, $\mathbf{x}_{\mathbf{m}t}^{\mathbf{D}}$, in year t :

$$\Pi_{jmt}^{PDP}(\mathbf{b}_{\mathbf{m}t}^{\mathbf{D}}, \mathbf{x}_{\mathbf{m}t}^{\mathbf{D}}) = \Gamma \left(\sum_{i \in A_{jmt}} [(p_{jmt}^{\mathbf{D}}(\mathbf{b}_{\mathbf{m}t}^{\mathbf{D}}) + z_{jt}^{\mathbf{D}} - c_{ijmt}^{\mathbf{D}}(\mathbf{x}_{\mathbf{m}t}^{\mathbf{D}}, r_{it}, \eta_{ijmt}))] \right), \quad (7)$$

where Γ is a function that adjusts for ex-post risk corridor transfers, which “directly decreases insurers’ exposure to bottom line risk” but do not affect profits in expectation (Decarolis, Polyakova and Ryan (2015)).²² The drug premium of policy j in county-level market m and year t is given by $p_{jmt}^{\mathbf{D}}$; it depends on the entire vector of bids. In addition, $z_{jt}^{\mathbf{D}}$ is the net of costs risk-adjusted subsidy for plan j in year t , and is not affected by selection. Individual costs $c_{ijmt}^{\mathbf{D}}$ are not fully compensated by risk-adjustment and are a function of the entire vector of product characteristics, $\mathbf{x}_{\mathbf{m}t}^{\mathbf{D}}$, the individual’s risk score, r_{it} , and an idiosyncratic shock, η_{ijmt} . Insurers choose the two by one vector of plan characteristics $x_{jmt}^{\mathbf{D}} = \left[p_{jmt}^{ICR}, p_{jmt}^{Donut} \right]'$ corresponding to the normalized out-of-pocket prices in the initial coverage region and the donut hole.²³ Firm profits depend on the entire vectors of bids and product characteristics in the market: competitor bids affect the benchmark subsidy, while competitor product characteristics affect plan selection. A_{jmt} represents the set of consumers who purchase good j which yields share s_{jmt} . As described in Section 2, the insurer collects premiums and subsidies and pays out claims for each Medicare beneficiary it enrolls. We decompose individual costs into two additive components such that $c_{ijmt}^{\mathbf{D}}(\mathbf{x}_{\mathbf{m}t}^{\mathbf{D}}, r_{it}, \eta_{ijmt}) = c_{ijmt}^{\mathbf{D}}(\mathbf{x}_{\mathbf{m}t}^{\mathbf{D}}, r_{it}) + \eta_{ijmt}$. The first component varies with plan benefits (either because of moral hazard or adverse selection or both) that may not be fully captured in risk-adjustment. The idiosyncratic term, η_{ijmt} , is realized after all the plan’s premium and benefit design decisions.

We write average plan costs as $\frac{1}{s_{jmt}B_{mt}} \sum_{i \in A_{jmt}} c_{ijmt}^{\mathbf{D}}(\mathbf{x}_{\mathbf{m}t}^{\mathbf{D}}, r_{it}) = c_{jmt}^{\mathbf{D}}(\mathbf{x}_{\mathbf{m}t}^{\mathbf{D}}, \bar{r}_{jmt})$ where \bar{r}_{jmt} is the average risk score and B_{mt} is the number of Medicare beneficiaries eligible to enroll in a PDP or MA-PD plan. We separate out the individual component of costs and risk-adjusted subsidy payments from the federal government η_{ijmt} , letting $H_{jmt}^{\mathbf{D}} = \frac{1}{s_{jmt}B_{mt}} \sum_{i \in A_j} \eta_{ijmt}$ be the *ex post* average of the realized cost shocks. The post-enrollment profit function is:

$$\Pi_{jmt}^{PDP}(\mathbf{b}_{\mathbf{m}t}^{\mathbf{D}}, \mathbf{x}_{\mathbf{m}t}^{\mathbf{D}}) = \Gamma \left(p_{jmt}^{\mathbf{D}}(\mathbf{b}_{\mathbf{m}t}^{\mathbf{D}}) + z_{jt}^{\mathbf{D}} - c_{jmt}^{\mathbf{D}}(\mathbf{x}_{\mathbf{m}t}^{\mathbf{D}}, \bar{r}_{jmt}) + H_{jmt}^{\mathbf{D}} \right) s_{jmt} B_{mt}, \quad (8)$$

²²Due to CMS regulation, Γ is piece-wise linear and symmetric about zero.

²³In this analysis we are holding the other characteristics of the insurer as fixed (e.g., marketing).

which is a function of PDP and MA-PD premiums and characteristics. Equation 8 differs from equation 7 only in that we have rearranged the sum over enrollees to be equal to the plan average and the sum of individual specific deviations from that average.

The stand-alone PDP plan chooses their bid and cost-sharing parameters x_{jmt}^D to maximize profit subject to an actuarial equivalence (minimum generosity) requirement \underline{x} :

$$\begin{aligned} \max_{b_{jmt}^D, x_{jmt}^D} \quad & \Pi_{jmt}^{PDP}(\mathbf{b}_{mt}^D, \mathbf{x}_{mt}^D) \\ \text{s.t.} \quad & x_{jmt}^D \geq \underline{x}. \end{aligned} \quad (9)$$

The profit function for Medicare Advantage plans is similar, though plans submit a bid for non-drug medical coverage, b_{jmt}^M , and there is an additional (risk-adjusted) subsidy payment for non-drug medical costs, z_{mt}^M .²⁴ We write:

$$\begin{aligned} \Pi_{jmt}^{MA}(\mathbf{b}_{mt}^D, \mathbf{b}_{mt}^M, \mathbf{x}_{mt}^D) = & (\Gamma(p_{jmt}^D(\mathbf{b}_{mt}^D) + z_{jt}^D - c_{jmt}^D(\mathbf{x}_{mt}^D, \bar{r}_{jmt}) + H_{jmt}^D) + \dots \\ & b_{jmt}^M + z_{mt}^M - c_{jmt}^M(\mathbf{x}_{mt}^D, \bar{r}_{jmt}) + H_{jmt}^M) s_{jmt} B_{mt}, \end{aligned} \quad (10)$$

where the M superscripts reflect medical (“Part C”) bids and costs, b_{jmt}^M is equal to the premium plus any rebates, $c_{jmt}^M(\mathbf{x}_{mt}^D, \bar{r}_{jmt})$ is the plans expected medical cost as a function of prescription drug benefits and mean Part D risk score. Similar to stand-alone PDPs, MA-PD plans must submit bids for medical coverage, incur costs that depend on individual and plan characteristics, and receive risk-adjusted subsidies. Unlike stand-alone plans, we observe separate bids for the drug and medical portions of MA benefits. Moreover, product characteristics affect MA-PD costs in two additional ways. First, drug plan characteristics could lead to selection that increases or decreases medical costs. Second, prescription offsets imply that drug plan characteristics could increase or decrease prescription drug expenditure, which could, in turn, increase or decrease medical expenditure. Because MA subsidies were generous during our time period, many plans have zero premiums; these plans may include “rebates” to enrollees, which can be used to provide additional services or reduce Part B premiums (which are required even for consumers in MA plans).²⁵

Our formulation of firm profits captures the bidding mechanism, risk corridors, and risk adjustment. However, four assumptions deserve additional attention. First, we focus on the coinsurance schedule, as directly modeling formulary placement of every drug is computationally unfeasible. Second, we assume that plans do not account for

²⁴There are separate subsidies for the non-drug component of MA-PD plans that vary at the market level; we incorporate these explicitly. Therefore, higher generosity due to more generous subsidies will not imply offsets. We abstract from MA rebates, which should increase the value of plans, but relax this assumption in Table A.12. We also estimate similar parameters when restricting to subsamples of plans for which the minimum standard regulation does not bind.

²⁵We include these net premium reductions directly in the definition of b_{jmt}^M ; the consumer pays a composite premium equal to $b_{jmt}^M + p_{jmt}^D$. We can allow the value of the rebate to be reduced by 25% in accordance with CMS bidding rules, which require that the government share in the savings of MA plans that bid below the benchmark. Table A.12 shows that this will not affect our estimates of the key parameters of interest.

their impact on the benchmark when submitting bids. Given the large number of plans, we believe this is a reasonable assumption. Third, we initially assume that insurers set optimal strategies ignoring cross-elasticities between plans in different sectors (for example, increasing the bid of Humana’s PDP plan decreases the enrollment of Humana’s MA plan) and then explore extensions to the model that incorporate multi-product firms.²⁶ Finally, we note that we need not impose an assumption of optimal pricing for PDP plans; we will never use the first-order condition with respect to *drug* premiums, though we will utilize the optimality of MA bids to infer costs (similar to Curto et al. (2015)). We do not observe medical claims for Medicare Advantage plans because the federal government does not collect these data for researchers, and private proprietary data cannot be linked to our administrative data.²⁷

4.2 Plan Choice

The first step in our analysis is to estimate consumer preferences over plans. We flexibly estimate PDP and MA-PD demand using a nested logit model that allows preference parameters to vary with observed health status, which we can measure using our detailed claims data. A consumer’s choice set is defined at the county-level and a product is a county specific insurance contract. While Part D insurers have identical offerings within the 34 large PDP regions, MA-PD plans can choose which counties to enter within a region. Medicare assigns both contract IDs and plan IDs. In MA-PD plans, the contract ID is typically specific to a geographic market; in stand-alone plans, the contract ID is typically national and the plan ID within a contract ID specific to the geographic market. As is standard in the literature, a MA-PD product is defined as a unique Medicare contract ID. Following a number of papers (Lustig (2010); Nosal (2011)), if there is more than one MA plan within a contract-county combination, we use the premium of the lowest numbered plan among MA-PD plan. A stand-alone PDP plan is defined as a Medicare contract ID-county combination; we average cost-sharing parameters within a county and the individual’s OOPC are based on their market. Assuming one Medicare plan ID per county within a contract, this is equivalent to defining the product at the Medicare contract ID-plan ID combination.

While firms select x_{jmt}^D , consumers care about their expected out-of-pocket costs for enrollment in plan j , $OOPC_{ijmt}$. $OOPC_{ijmt}$ is a nonlinear, individual specific function of x_{jmt}^D . We construct $OOPC_{ijmt}$ by developing a function that maps total drug spending into out-of-pocket costs, taking insurer cost-sharing, consumer consumption, and the structure of the standard benefit as given. For example, a consumer ending the year in the initial coverage phase will have $OOPC_{ijmt}$ equal to the deductible (if any) and the number of days supplied in the initial coverage phase multiplied by

²⁶Finally, we note that assumptions about drug premium setting are especially important as previous research has highlighted that the presence of the LIS subsidy can distort plan bidding incentives (Decarolis (2015); Decarolis, Polyakova and Ryan (2015)). In Section 4.3.4, we test the robustness of our estimates to each these assumptions.

²⁷The first order condition with respect to bids provides no information about the relationship between generosity and costs assuming that for small changes in generosity, the optimal drug premium is unchanged. In counterfactual exercises, we will allow firms to reprice and product characteristics to adjust simultaneously.

P_{jmt}^{ICR} .²⁸ The relationship between the structure of Part D benefit, insurer choice variables, and OOPC is described in Figure 2: the first panel describes the standard benefit, the second panel describes how the insurer choice variables captured by x_{jmt}^D map into $OOPC_{ijmt}$ and differences across plan types. This formulation implies that market shares are a function of insurer choice variables because insurer choice variables affect consumer OOPC.

In our baseline specification, we divide the sample into five types of consumers based on quintiles of 2006 total drug spending. In each quintile q , consumer utility for plan j (which can be either a PDP or a MA-PD plan) in market m at time t is given by:

$$u_{qjmt} = \bar{\xi}_{qj} + \alpha_q^p p_{jmt}^D + \alpha_q^x OOPC_{qjmt} + \xi_{qjmt} + \zeta_{ig} + (1 - \sigma_q)\epsilon_{ijmt}, \quad (12)$$

where $\bar{\xi}_{qj}$ are time-invariant, vertical plan characteristics (i.e., plan fixed effects) that vary across consumer types, p_{jmt}^D is the drug premium, $OOPC_{qjmt} = f(x_{jmt}^D)$ is a function of the average OOPC per days supply as is defined above, ξ_{qjmt} represents time varying shocks to unobservable vertical plan characteristics, ζ_{qg} is common to all products in nest g and has a distribution function that depends on σ_q with $0 \leq \sigma_q < 1$. Beneficiaries are assumed to choose the plan that yields the highest utility. We assume three nests in this specification: PDP plans, MA-PD plans and the outside good. We do not directly model the impact of MA-PD non-drug premiums on consumer choice in MA-PD plans. The vast majority of plans have zero (non-drug) premiums, and some rebate a portion of the Part B premium, reducing salience to consumers and making measurement difficult.²⁹ As is standard in differentiated product demand systems, unobserved plan characteristics are captured by ξ_{qjmt} . The outside good consists of both no drug coverage and any coverage not associated with a stand-alone PDP or MA plan (such as an employer-sponsored plan). We assume that ϵ_{ijmt} is distributed extreme value and thus $\zeta_{qg} + (1 - \sigma_q)\epsilon_{ijmt}$ is also distributed extreme value. We allow plan unobserved quality, premium and OOPC coefficients, and the dissimilarity term to vary by consumer type. In mathematical terms, $\bar{\xi}_{qj}$, α_q^p , and α_q^x are all indexed by q , implying that preferences over the time-invariant features of plans (plan fixed effects), premiums, and generosity are all allowed to vary by consumer type.³⁰

In order to capture firm incentives, we must identify the causal impact of premiums and OOPC on plan enrollment. Our estimates will be biased if ξ_{qjmt} is correlated with premiums or product characteristics that influence OOPC. We

²⁸For a consumer of type q in market m with consumption (days supplied) d , this can be written as:

$$OOPC_{qjmt} = \begin{cases} R_{jt}d \text{ if } R_{jt}d < DED \\ P_{jmt}^{ICR} \left(d - \frac{DED}{R_{jt}} \right) + DED \text{ if } R_{jt}d \geq DED \text{ and } R_{jt}d < ICL \\ P_{jmt}^{Donut} \left(d - \frac{ICL}{R_{jt}} \right) + DED + \gamma_{ICR}(ICL - DED) \text{ if } R_{jt}d \geq ICL \text{ and } R_{jt}d < CAT \\ .05R_{jt} \left(d - \frac{CAT}{R_{jt}} \right) + DED + \gamma_{ICR}(ICL - DED) + \gamma_{Donut}(CAT - ICL) \text{ if } R_{jt}d \geq CAT \end{cases}, \quad (11)$$

where d is the days supplied, γ represents the average coinsurance in each phase, R_{jt} is the mean retail price for plan j , and DED , ICL , and CAT represent the statutory deductible, initial coverage limit, and catastrophic cap, respectively. For any enrollee and level of consumption, there is a mechanical and monotonic relationship between the insurer's choice variables P_{jmt}^{ICR} and P_{jmt}^{Donut} and enrollee costs $OOPC_{qjmt}$.

²⁹The supply-side inversion will assume – consistent with a neoclassical model – that the elasticity with respect to drug premiums, non-drug premiums, and subsidies is the same.

³⁰In unreported specifications, we find that estimated sensitivities to price and OOPC are not different by plan type.

address this issue via a two-pronged approach. First, we include product fixed effects, $\overline{\xi_{qj}}$, that are allowed to vary with consumer type: the unobserved product characteristic is the deviation from the plan mean for the quintile in question. Second, we instrument for the premium, OOPC, and the inside share. Following the logic of the reduced form identification strategy described in the appendix, the instrument for the inside share is the urban dummy interacted with an MA-PD dummy, which captures the fact that MA-PD plans are more popular in urban counties. As is common in this setting, we also use Hausman-style instruments: the average premiums and OOPC in all other markets.

The demand specifications reflect several additional modeling choices. Critically, a number of recent papers (Ericson (2014); Heiss et al. (2016); Ho, Hogan and Scott Morton (2015); Miller and Yeo (2014); Polyakova (2016); Wu (2016)) have documented “stickiness” in plan choice over time. Following Decarolis, Polyakova and Ryan (2015), we take a reduced form approach to modeling inertia and include plan vintage in the utility function in our preferred specification; they show that their approach corresponds to “an explicit structural model of inattention and choice” (Hortascu, Madanizadeh and Puller (2015)). Because we are not explicitly interested in the effect of switching costs on premiums or benefit design, we do not develop a dynamic model of premium setting or benefit design.³¹ We allow for nests containing MA-PD plans, stand-alone PDPs, and the outside option but do not allow for behavioral responses to average OOPC per days supply to factor into the calculation.³² We do not directly model the impact of the LIS policy on plan behavior, though we also examine the magnitude of this potential distortions in the robustness checks subsection and find that our conclusions are robust to the LIS incentives. Finally, while we allow the dissimilarity term, time-invariant plan quality, and premium and OOPC coefficients to vary by consumer type, the nested logit error term, $\zeta_{ig} + (1 - \sigma_q)\varepsilon_{ijt}$, is the only source of unobserved consumer heterogeneity in the model. However, in robustness analyses, we allow for unobserved consumer heterogeneity; our main findings are not sensitive to allowing for more flexible patterns of substitution.

4.2.1 Selection

Our model explicitly accounts for selection of enrollees into plans. Advantageous selection into Medicare Advantage on the extensive margin is accounted for by the fixed effects, which vary by quintile. More importantly, when setting plan characteristics, the insurer may also be concerned that a more generous drug plan will attract sicker enrollees on the margin. In order to address potential selection with respect to plan characteristics, we allow for preferences and drug costs to vary flexibly by consumer type (defined as quintiles of the 2006 drug spending distribution). As premiums

³¹Table A.11 replicates the reduced form results in Ho, Hogan and Scott Morton (2015) that suggest an “invest then harvest” strategy of premium setting. However, we also show that plan generosity, measured by insurer drug spending, does not exhibit the same pattern. As a result, we do not believe modeling dynamic insurer behavior in our setting is critical for understanding endogenous plan design.

³²Put differently, there is no selection on moral hazard and the relationship between P_{jmt}^{ICR} , P_{jmt}^{Donut} , and $OOPC_{qjtm}$ is purely mechanical; therefore, the derivative of shares with respect to any element of the cost-sharing schedule is given by $\sum_q \frac{\partial s_{qjtm}}{\partial OOPC_{qjtm}} \frac{\partial OOPC_{qjtm}}{\partial P_{jmt}}$. Differences in drug demand across plan types is consistent with plans attracting different types of consumers on average, but not necessarily selection on moral hazard on the margin. This is an interesting avenue for future research.

or generosity changes, plans attract a different mix of enrollees (across quintiles), which accommodates the potential for selection. If a plan becomes more generous and that disproportionately attracts higher cost enrollees because of demand heterogeneity, increased costs will be reflected (through a greater weighting of higher cost quintiles) in H_{jmt}^D .³³

Given the literature and reduced form estimates in Table A.3, which show that a MA-PD and consumer type dummies account for a substantial amount of variation in insurer expenditure, our approach accounts for nearly all of the heterogeneity in drug costs that could be observed or predicted by the insurer. Einav, Finkelstein and Schrimpf (2015) find a raw monthly correlation of drug consumption of 0.5; Hsu et al. (2009) argue that “approaches that include information on prior-year drug use or costs perform markedly better than the current Medicare risk-adjustment approaches.”³⁴ Furthermore, our flexible demand estimates explicitly allow heterogeneity in preferences to be correlated with heterogeneity in costs. Our model does not explicitly accommodate selection with respect to formulary design (Carey (2017); Lavetti and Simon (2014)), but we will show that our results are robust to allowing for more heterogeneity by defining finer consumer types and allowing for unobserved consumer heterogeneity.³⁵

4.2.2 Demand Parameter Estimates

The results of the demand specifications for each of the five consumer groups are in Table 4. Panel A describes baseline results that do not account for plan vintage. While we allow the utility parameters to vary across drug spending quintiles, we do not find meaningful heterogeneity in the premium or OOPC coefficients. The premium coefficient is negative and significant in all specifications, sicker consumers are slightly less price sensitive than healthier consumers, consistent with adverse selection with respect to generosity. The OOPC coefficient is much smaller in magnitude than the premium coefficient, consistent with the results in Abaluck and Gruber (2011) and becomes attenuated among sicker consumers. We note OOPC is observed with error and its coefficient estimate may be attenuated; while differences across consumer groups may reflect differential preferences, they could also reflect larger measurement error among higher spending enrollees.³⁶

Own premium elasticities are quite sensible and range from -4.6 to -5.7, depending on (observed) consumer type. This is consistent with the results in Decarolis, Polyakova and Ryan (2015); our estimates for stand-alone plans range from -5 to -6.3. Finally, across all groups, the nested logit dissimilarity parameter indicates that MA-PD (stand-alone) plans are much closer substitutes for other MA-PD (stand-alone) plans than stand-alone (MA-PD) plans.

Panel B describes our preferred estimates, which account for the impact of plan vintage. While the premium

³³We also implicitly allow changes in drug premiums to alter the risk pool the firm attracts through deviations from the average claim within a consumer type, though this will not affect our empirical implementation.

³⁴Furthermore, the degree of selection conditional on risk adjustment in these markets is a matter of debate; see Newhouse et al. (2015); Brown et al. (2014). Our specification allows for consumer heterogeneity in preferences by including flexible plan fixed effects that can vary by consumer type, which implicitly allows for differential selection into plans based on consumer type.

³⁵For example, we explore deciles of 2006 total drug spending, conditioning on demographics, or considering only high offset drugs.

³⁶The estimated difference in the parameters suggests that consumers are naive about under-consumption that may lead to future adverse medical events or financial costs. They do not demand more generous plans as a commitment device.

Table 4: IV Nested Logit Parameter Estimates

Spending Quintile	(1)	(2)	(3)	(4)	(5)
Panel A: Baseline Estimates					
Premium	-0.191*** (0.0227)	-0.187*** (0.0207)	-0.246*** (0.0179)	-0.230*** (0.0161)	-0.208*** (0.0145)
OOPC	-0.138*** (0.0132)	-0.0974*** (0.00920)	-0.0491*** (0.00642)	-0.0327*** (0.00437)	-0.0157*** (0.00263)
$1 - \sigma$	0.506*** (0.0142)	0.508*** (0.0144)	0.508*** (0.0144)	0.532*** (0.0135)	0.521*** (0.0127)
Adjusted R ²	0.295	0.279	0.272	0.266	0.244
Panel B: Accounting for Inertia					
Premium	-0.198*** (0.0211)	-0.189*** (0.0193)	-0.237*** (0.0172)	-0.224*** (0.0154)	-0.198*** (0.0140)
OOPC	-0.136*** (0.0120)	-0.0939*** (0.00844)	-0.0516*** (0.00605)	-0.0336*** (0.00419)	-0.0151*** (0.00255)
$1 - \sigma$	0.439*** (0.0132)	0.441*** (0.0134)	0.474*** (0.0130)	0.472*** (0.0128)	0.472*** (0.0120)
Plan Vintage	1.136* (0.631)	1.653*** (0.632)	1.018 (0.728)	1.088* (0.634)	2.412*** (0.640)
Adjusted R ²	0.307	0.294	0.288	0.281	0.262
Observations	58,189	58,626	59,885	60,463	61,317

Notes: Table presents instrumental variable regression models as described in Berry (1994). The outside share is constructed as all Medicare eligibles not enrolled in a stand-alone Medicare Part D plan or MA-PD plan. In all specifications, we include plan fixed effects. Excluded instruments are an urban county dummy, and premiums and out-of-pocket costs in other markets, where a market is defined as a county-year combination. Standard errors are presented in parentheses. Statistical significance at the 10%, 5%, and 1% levels are denoted by *, **, and *** respectively.

elasticities and the relative preference for lower OOPC are similar in magnitude, there are two key differences to highlight. First, we estimate a smaller nesting parameter, indicating greater substitution within nests. Second, the effect of plan vintage differs with consumer characteristics. Consistent with Ho, Hogan and Scott Morton (2015), we find that consumers with relatively low claims are unlikely to switch. In addition, we find that the sickest consumers are even less likely to switch, perhaps due to strong brand preferences. This enrollee inertia implies that, if lower drug prices is an investment in reducing future enrollee hospitalizations, plans are likely to earn a return from that investment. Heterogeneity in the degree of inertia is critical for understanding differences in plan incentives, as we describe below.

4.3 Endogenous Plan Design

To understand the role of medical cost offsets in plan design and to perform counterfactuals, we model insurer behavior, taking our demand estimates as given. In order to estimate the key parameters of interest, we must first construct plan level measures of expected medical and drug costs. We first estimate c_{jmt}^M by utilizing the fact that we observe separate bids (and, therefore premiums) and subsidies for the medical and drug spending components of MA plans; we calculate these medical costs “offline” separately for each plan and assume that they are known to the firm in the rest of the supply side estimation. We impute medical costs for MA-PD enrollees (for whom medical claims are not observed) by inverting the plan’s first-order condition with respect to the MA bid, assuming firms bid optimally given premium elasticities, bids, subsidies, and observed market shares, as in Curto et al. (2015). Formally, for MA-PD plans,

$$c_{qjmt}^M = (p_{jmt}^D + z_{qt}^D - c_{qjmt}^D + b_{jmt}^M + z_{mt}^M) + s_{qjmt} \frac{\partial s_{qjmt}}{\partial b_{jmt}^M}^{-1}, \quad (13)$$

where costs c_{qjmt}^D , subsidies z_{qt}^D , shares s_{qjmt} are allowed to vary with consumer type.³⁷ For this calculation, we average costs across quintiles and assume consumers view drug and medical premiums equivalently. We further assume that the expectation of the deviations of medical costs from the average is zero: medical bids do not affect selection among MA-PD plans and we do not explicitly model non-drug benefit design except through the determination of the rebates. We calculate subsidies using the formula provided by CMS, averaging 74.5% of bids. We estimate average margins of 14%, similar in magnitude to those obtained by Curto et al. (2015).

Table 5 lists the variables used in our estimating equations and categorizes them as either data, implied objects from the demand estimation, or parameters to be estimated. Prices, subsidies, market shares, and realized costs are

³⁷For the supply side model, we assume that firms optimize each plan’s characteristics, rather than optimizing over their entire portfolio. This is a simplification due to data availability: we only observe encrypted Medicare contract and plan identifiers in the drug event data. To minimize the potential bias, we treat the stand-alone PDP contract IDs as the unit of analysis in the supply side estimation. Given high correlation within nests in the demand system (e.g., a consumer substituting away from a stand-alone PDP is most likely to choose another stand-alone PDP offered by a different firm), we believe it is unlikely that a consumer will substitute between the MA-PD and PDP plans within a single firm; therefore, this assumption seems fairly reasonable. The above expression assumes that the minimum benefits constraint is not binding.

all observed in the data. We allow the realization of drug costs to differ from firm expectations. In order to construct expected drug costs, denoted in the model by c_{jmt}^D , we estimate the regression described in the sixth and final column of Table A.3 on the entire sample.³⁸ The derivative of shares with respect to drug premiums, $\frac{\partial s_{qjmt}}{\partial p_{jmt}^D}$, is derived from the demand estimates by consumer type. The derivative of shares with respect to phase-specific cost-sharing has two components. The derivative of shares with respect to OOPC, $\frac{\partial s_{qjmt}}{\partial OOPC_{qjmt}}$, is derived from the demand estimates by consumer type, while the derivative of OOPC with respect to cost-sharing parameters is mechanical.³⁹ Identification of the key parameters of interest – $\frac{\partial c_{jmt}^D}{\partial OOPC_{jmt}}$ and $\frac{\partial c_{jmt}^M}{\partial OOPC_{jmt}}$ – relies on the first order conditions with respect to the cost-sharing schedule, along with variation across plan types. There is a direct mapping from different values of the key parameters of interest to levels of insurer drug spending. Intuitively, the more “expensive” it is to make plans more generous, because of asymmetric information or the absence of offsets, the less willing the firm is to increase generosity. Figure A.5 illustrates the basic logic of our identification argument using average values of the derivatives of shares with respect to premiums and OOPC.⁴⁰

4.3.1 Estimation and Selection

We parametrize insurer drug costs as a linear function of $OOPC_{jmt}$ scaled by an average effect across plans θ_1 plus a plan-specific deviation ψ_{jmt} : $c_{jmt}^D = (\theta_1 + \psi_{jmt})OOPC_{jmt} + \omega_{jmt}^D$, where ω_{jmt}^D captures all the exogenous features of the plan that affect its drug costs. Medical costs are similarly scaled by a constant effect across plans: $c_{jmt}^M = \theta_2 OOPC_{jmt} + \omega_{jmt}^M$. This parameterization yields $\frac{\partial c_{jmt}}{\partial OOPC_{jmt}} = \theta_1 + \theta_2 1(MA_{jmt}) + \psi_{jmt}$ and the profit function can be written as:

$$\Pi_{jmt} = \Gamma \left(\sum_q [p_{jmt}^D + z_{qt}^D - c_{qjmt}^D + 1(MA_{jmt}) * (b_{jmt}^M + z_{mt}^M - c_{jmt}^M)] s_{jqmt} B_{qmt} \right). \quad (14)$$

H_{jmt}^D and H_{jmt}^M do not enter this formulation because the costs and subsidies are allowed to vary by quintile directly and observable types summarize everything that the firm observes about individual consumers when setting plan characteristics; selection will be captured in our estimate of $\frac{\partial c_{jmt}}{\partial OOPC_{jmt}}$. Therefore, the plan’s first-order conditions with respect to x_{jmt}^D for Nash equilibrium imply that:⁴¹

$$\sum_q \left[\left(p_{jmt}^D + z_{qt}^D - c_{qjmt}^D + 1(MA_{jmt}) \right) \left(b_{jmt}^M + z_{mt}^M - c_{jmt}^M \right) \right] \frac{\partial s_{qjmt}}{\partial x_{jmt}^D} - \frac{\partial c_{jmt}}{\partial OOPC_{jmt}} \sum_q \frac{\partial OOPC_{qjmt}}{\partial x_{jmt}^D} s_{qjmt} = 0. \quad (15)$$

³⁸This allows us to abstract from plan selection and allow for “medical management” on the part of MA-PD plans. While we observe point-of-sale drug costs, we do not observe any rebates negotiated by payers, which are unlikely to affect generosity on the margin. A full model of bargaining between manufacturers and pharmacy benefit managers is outside the scope of this paper.

³⁹For example, consider a \$1 increase in the ICR cost. For an enrollee with total drug spending below the deductible, the derivative is zero. For an enrollee above the initial coverage limit, the derivative is also zero (though this enrollee will reach the limit earlier in the year). In the ICR, the derivative is equal to the days supply less the days supply required to hit the deductible (the deductible divided by the average retail price). We do not consider cases in which a (small) change in the phase-specific costs per day supplied would push an enrollee into the next phase of the benefit design as this effect complicates the analysis substantially and is of second order relevance to the analysis. Furthermore, we do not allow the enrollee to forecast a behavior response to changes in the phase-level average enrollee costs per days supply.

⁴⁰On the x-axis, we plot different values of the derivative of interest, denoted by θ , while on the y-axis, we plot insurer spending. The results are slightly different across different phases, but illustrate that the smaller in absolute value θ is, the more the plan spends on drugs.

⁴¹Recall that $x_{jmt}^D = [p_{jmt}^{ICR}, p_{jmt}^{Donut}]$.

Table 5: Parameters and Identification

Object	Description	Inference
P_{jmt}^D	drug premium	data
b_{jmt}^M	MA bid	data
z_{jmt}^M, z_{jmt}^D	Part C (medical) and D subsidies	data
s_{qjmt}	market share	data
c_{qjmt}^D	drug costs	data, with expectation formed using specification in Table A.3 by consumer type to allow for selection
$\frac{\partial s_{qjmt}}{\partial P_{jmt}^D}$	derivative of market share w/r/t drug premium	calculated from demand estimates, by consumer type to allow for selection
$\frac{\partial s_{qjmt}}{\partial x_{jmt}^D}$	derivative of market share w/r/t drug cost-sharing	calculated from demand estimates, by consumer type to allow for selection
c_{jmt}^M	medical cost	inferred from bidding decision using Part C bids and subsidies, MA plans only
$\frac{\partial OOPC_{jmt}}{\partial x_{jmt}^D}$	derivative of out-of-pocket costs w/r/t drug cost-sharing	calculated from data
$\frac{\partial c_{jmt}^D}{\partial OOPC_{jmt}}$	mean derivative of insurer drug costs w/r/t consumer out-of-pocket costs	object of interest to estimate
$\frac{\partial c_{jmt}^M}{\partial OOPC_{jmt}}$	mean derivative of insurer medical costs w/r/t consumer out-of-pocket costs	object of interest to estimate

Given our parameterization of $\frac{\partial c_{jmt}}{\partial OOPC_{jmt}}$, estimates of θ_1 and θ_2 represent the percentage increase (decrease) in OOPC that is passed on to the insurer in the form of total savings (costs). If there is no adverse selection, moral hazard or offsets, then $\theta_1 + \theta_2 1(MA) = -1$. In our specification, we implicitly assume that profits are additive across consumer types and estimate the relationship between OOPC and insurer costs on average: $\frac{\partial c_{jmt}}{\partial OOPC_{jmt}}$ is not allowed to vary by consumer type. If more generous plans attract sicker enrollees, increasing OOPC will attract healthier enrollees, lowering insurer costs more than one-for-one. If there are offsets and drug demand slopes down, higher OOPC will increase non-drug medical costs. That is, with adverse selection and/or moral hazard, $\theta_1 < -1$. The presence of offsets implies a larger $\theta_2 > 0$. We infer $\frac{\partial c_{jmt}}{\partial OOPC_{jmt}}$ by inverting (15) and then use OLS to estimate $\frac{\partial \hat{c}_{jmt}}{\partial OOPC_{jmt}} = \theta_1 + \theta_2 1(MA_{jmt}) + \psi_{jmt}$ where $\frac{\partial \hat{c}_{jmt}}{\partial OOPC_{jmt}}$ is the inferred value from the inversion.

4.3.2 Supply Side Parameter Estimates

Before presenting our preferred model estimates, we estimate θ assuming that is is constant across all plans. The estimates are in Table 6. We estimate two sets of parameters: the first set is based on the demand estimates without vintage effects and the second set uses the demand parameters with vintage effects. The estimates in column 1 imply that the average Part D plan would save \$99.73 per member by increasing OOPC by \$100. The estimates using the vintage effects specification imply a smaller impact of an increase in OOPC. In columns 2 and 4 we allow θ to vary

Table 6: Supply Results

$\partial c/\partial OOPC$	Baseline		Including Inertia & High Spenders	
	(1)	(2)	(3)	(4)
Constant	-0.9973 (0.0098)	-1.0323 (0.0109)	-0.9700 (0.0078)	-1.3427 (0.0078)
MA		0.3488 (0.0341)		0.4227 (0.0315)
Plan-Market-Year Obs.	34,431	34,431	33,538	33,538

Notes: Parameters are estimated using generalized method of moments as described in Section 4. Standard errors are calculated using a bootstrap that re-samples plans with replacement and presented in parentheses.

between PDP and MA-PD plans. Specifically, we parameterize $\theta_j = \theta_1 + \theta_2 \times 1(MA_j)$. The impact of increasing generosity in a MA-PD plan is simply the sum of the two coefficients.⁴²

The results in column 2 and 4 show that the relationship between plan generosity and insurer costs is economically statistically different across different types of plans. The impact of increasing generosity is smaller in magnitude for MA-PD plans (by \$35-\$42). The parameter θ_2 captures differences across plan types in both $\frac{\partial c_{jmt}^D}{\partial x_{jmt}^D}$ and $\frac{\partial c_{jmt}^M}{\partial x_{jmt}^D}$. Under the assumption that asymmetric information – including selection along the intensive margin – is constant across plan types, θ_2 measures the cost offset effect. Absent cost offsets, our results would imply substantial intensive margin advantageous selection among MA-PD plans (increased drug benefit generosity leads to lower average costs, or the marginal enrollee is much healthier than the average). We find this interpretation implausible and argue that our results indicate a fraction of the MA-PD insurer’s drug cost is offset by reductions in insurer spending in other areas.

4.3.3 Implied Offsets

Our estimates quantify how changes in cost-sharing arrangements affect insurers’ cost structure. By combining our estimates with estimates of the observed behavioral response to enrollee cost-sharing, we can calculate the implied offsets and the impact of counterfactual policies. We take estimates of the behavioral response to cost-sharing from Einav, Finkelstein and Schrimpf (2015), who estimate a dynamic model of drug consumption.⁴³ Consider a small enrollee out-of-pocket cost decrease. The price decrease will increase insurer drug costs by an amount equal to the drug point-of-sale price less enrollee cost for the marginal units, and the price difference times all of the infra-marginal units. We know from the supply side estimation that the increase in insurer costs associated with lowering enrollee OOPC is smaller for MA-PD plans than stand-alone PDP plans. That is, $\frac{\partial c_{jmt}^M}{\partial x_{jmt}^D} > 0$. We can infer the increase in quantity from the *observed* behavioral elasticity and calculate the magnitude of implied offsets that rationalizes firm

⁴²MA variable costs are given as $c_{jmt} = c_{jmt}^D + c_{jmt}^M$, thus the derivative of MA-PD costs with respect to the cost-sharing schedule can be written as $\frac{\partial c_{jmt}}{\partial x_{jmt}^D} = \frac{\partial c_{jmt}^D}{\partial x_{jmt}^D} + \frac{\partial c_{jmt}^M}{\partial x_{jmt}^D}$. We note that c_{jmt} at the plan level is just the enrollment-weighted average over quintiles.

⁴³The elasticity is identified by exploiting the kink in individuals’ budget sets created by the donut hole; we reproduce their elasticity estimates in Table A.5 and use the elasticity for a 1% uniform out-of-pocket cost reduction of -0.54 in our calculation.

behavior.⁴⁴

For a 1% uniform decrease in cost-sharing in all benefit phases, we calculate the implied difference in insurer costs for MA-PD plans as (omitting subscripts for simplicity):

$$\frac{\partial c^M}{\partial x^D} = \theta_2 \frac{\partial OOPC}{\partial x^D}, \quad (16)$$

which gives the difference in the change in insurer costs between MA-PD plans and stand-alone PDP plans. The implied drug cost is simply the point-of-sale cost less implied offsets. This quantity must be equal to the change in quantity times the difference in the drug cost per day, denoted by c_{PDP} , and a shadow drug cost that incorporates offsets, c_{MA} :

$$\frac{\partial q}{\partial x^D} (c_{PDP} - c_{MA}).$$

We use the value of $\theta_2 = 0.42$ as estimated in Table 6 and the mechanical value of $\frac{\partial OOPC}{\partial x^D}$, \$10.15.⁴⁵ We take $\frac{\partial q}{\partial x^D}$ directly from Einav, Finkelstein and Schrimpf (2015), and use the average empirical value of c_{PDP} (\$2.20). We find $c_{MA} = \$1.61$. Put differently, the “discount” implied by offsets is 59 cents per day supplied, or 27%. The offset we calculate is very close to previous estimates (see CBO (2012) for a synthesis of the literature using demand side variation) and obtained using supply side variation.

4.3.4 Robustness

In this section, we test the robustness of the model to assumptions along two key dimensions. First, we explore the impact and modeling of regulatory constraints. Second, we explore the robustness of our results to alternative demand specifications.

Section 2 highlighted a number of regulatory features of the Part D market – including distortions created by the LIS subsidy – that may affect firm incentives. Decarolis, Polyakova and Ryan (2015) account for this possibility in their analysis and we following their approach by restricting attention to those contracts that bid above the LIS benchmark amount in the majority of regions in which they operate. We re-estimate the supply side parameters, assuming that these remaining contracts will not be distorted by the structure of the LIS subsidy. This restriction leaves fewer contracts available to estimate the parameters. Nevertheless, the results (presented in Table A.12) are qualitatively similar to our main specification, indicating that our basic results hold even allowing for the LIS distortion. Similarly, in the bottom panel of Table A.12, we restrict attention to “enhanced” plans that are unlikely to be constrained by minimum plan generosity requirements; again, the results are quantitatively similar.⁴⁶

⁴⁴The shadow drug cost only takes insurer medical costs c_{int}^M into account and is therefore strictly higher than the “true” shadow cost. This has important implications for welfare, as discussed in Glazer and McGuire (2013).

⁴⁵For this calculation, we use the average value of enrollee out-of-pocket cost per day supplied and the formula described above, evaluated at average consumption.

⁴⁶We define a plan as “enhanced” if it has a supplemental premium in at least half of all markets in which it operates. This accounts for 7% of

Our current approach relies on plan fixed effects and identifies sensitivity to price and OOPC using cross market variation. In order to address the role of selection in the determination of contract characteristics more directly, we model plan choice as a function of contract characteristics. In these specifications, we include premiums, OOPC, and dummies for deductible and donut hole coverage as the observable characteristics.⁴⁷ Consistent with previous results, we find evidence of selection with respect to donut hole coverage in Table A.9: consumers with higher 2006 drug consumption (our measure of severity) have a higher preference for donut hole coverage. We re-estimate our supply side parameters using variation from only the ICR price in Table A.12, which eliminates any concerns about selection with respect to donut hole coverage. Consistent with selection being driven by coverage in the donut hole, we do not see evidence of adverse selection when we restrict attention to the ICR price. However, the differential effect for MA-PD plans (θ_2) is quite similar to our main specifications.⁴⁸ Because of the flexibility of our demand system, we believe our baseline approach removes the confounding unobservables that would bias parameter estimates and that we accurately capture plan incentives for choosing plan characteristics in our setting.⁴⁹

Our current approach also models heterogeneity in demand across observable, but not unobservable consumer characteristics. We argue that, given both the empirical literature and reduced form results, we include most of the variation across consumers that is predictable to insurers. However, a more flexible demand system may better capture the patterns of substitution in the data. In Table A.10, we report estimates that allow for a mixture of consumer types, similar to Berry and Jia (2010). The results are qualitatively similar to our main estimates, though they imply slightly less price sensitivity among high spenders; we also find evidence that a subset of these consumers are fairly insensitive to both premiums and OOPC. In Table A.12, we show that our supply side estimates are not sensitive the inclusion of additional consumer heterogeneity.

4.3.5 Extensions

In this section, we show how our framework can be extended to incorporate additional contractual richness and answer other key questions of economic and policy importance. In the main empirical exercise, we model insurer choice of composite measures of plan generosity. We further show that our estimates of consumer preferences are not sensitive to the exact definition of product characteristics and are in line with previous estimates. Other product characteristics

all stand-alone plans. The analysis omitting LIS plans serves as an additional robustness check along this dimension as plans not eligible for LIS enrollment are also likely to be more generous than the minimum actuarial standard. These two subsets overlap substantially.

⁴⁷The results in Table 4 imply that sicker consumers are slightly less sensitive to both premiums and OOPC. One concern is that the last result could be driven by increased OOPC measurement error among sicker consumers and that could have broader implications for our base estimates. In order to address this, we examine the sensitivity of our supply side estimates to differences in preferences over plan generosity of sicker consumers. In unreported specifications, we find that increasing the plan generosity elasticity of sicker consumers affects our mean estimate of θ_1 , but not the principal parameter of interest, the interaction with the MA-PD plan dummy (θ_2). This finding is consistent with a greater degree of adverse selection but no difference in offsets across severity levels. Finally, we can also estimate our results assuming that firms return the entire rebate to enrollees and find qualitatively similar results. These results are available from the authors upon request.

⁴⁸In unreported regressions, we also estimate our model with the two cost-sharing parameters directly. While the results are qualitatively similar, we note that the tariffs are highly correlated, leading to noisier coefficients. Finally, while differential reinsurance and risk corridors may have a small effect on our results, we believe their impact is directionally ambiguous (Medpac (2015b)).

⁴⁹Finally, we can also estimate our results assuming that firms return the entire rebate to enrollees and find qualitatively similar results.

may be important to both consumers and policymakers; however, our framework is flexible enough to model additional plan features.

For example, we are interested in modeling different generosity among high offset drugs directly. We proceed in three steps. First, we define \mathbf{x}_{mtd}^D , the vector of plan generosity specific to a given drug class d . Second, we assume that consumption of the drug of interest is evenly timed throughout the year in proportion to overall spending, which implies that $\frac{\partial OOPC_{qjmt}}{\partial x_{jmt}^D}$ will be proportional to $\frac{\partial OOPC_{qjmt}}{\partial x_{jmt}^D}$, scaled by the difference in phase-specific prices for the focal drug and all drugs. Finally, we assume that consumer preferences over generosity are not drug class specific; if one wanted to relax this assumption, they would need to separately estimate demand taking the new product characteristic into account.

Table A.13 displays our estimates of θ_2 by drug class. Across drug classes with large offsets, we estimate higher values of θ_2 , ranging from 0.36 for hypertension (consistent with MA-PD offering less generous coverage on the whole in the reduced form) to 0.58 for cholesterol and diabetes drugs. While the estimates are somewhat noisy, the estimates for asthma, cholesterol, and diabetes are all larger in magnitude than our main estimate of 0.39, indicating that heterogeneity across drug classes is important in our setting. Our estimates are also closely linked to the reduced form estimates in Figure A.3, highlighting the variation in the data that allows us to estimate the key parameters of interest.

Our framework can be used to further explore endogenous product design in the context of selection and can help researchers model firm behavior and measure welfare in an “endogenous contracts framework” (Geruso and Layton (2018)). Thus, we build on an existing literature that considers premium setting conditional on fixed plan characteristics (Handel (2013); Lustig (2010); Starc (2015); Town and Liu (2003); Tebaldi (2017); Ericson and Starc (2015)). Future work should explore the impact of supply side regulation, including premium and contract regulation, within the context of a rich model that does not require fixed insurance contracts.

5 Counterfactuals

In this section, we aim to answer three sets of questions with our model and parameter estimates. First, how would stand-alone plans design drug benefits if they took externalities into account? Second, how large is the effect relative to distortions due to adverse selection? Third, how would different subsidy policies affect equilibrium premiums and benefit design?

5.1 The Effect of Internalizing Externalities

To quantify the importance of the offset, we first consider how PDP plans would adjust plan benefit design if they were forced to account for non-drug medical costs in the same way as their counterparts in the MA program. Mechanically,

we set $\theta_{PDP} = \hat{\theta}_{MA}$ and then resolve for a new Nash equilibrium conditional on θ . In the counterfactual equilibrium, we allow drug costs to the insurer, c_{jmt}^D , and premiums to adjust to account for the new incentives.⁵⁰ In our baseline scenario, the average MA-PD plan has lower premiums (because of generous subsidies) than the average PDP plan, which have average premiums of \$407 per year. By contrast, the average MA-PD plan spends almost \$75 dollars more per year on drugs (\$1285 versus \$1211).

In the top panel of Table 7, we report the results of a simulation in which premiums are not allowed to adjust, but PDP plans internalize the offset. In the counterfactual, we see that the average PDP plan would spend 7% more on prescription drugs if they took the entire medical offset into account. In addition, we note that MA-PD plans increase drug spending: plan generosity is a strategic complement, and there is no implicit trade-off between higher generosity and higher premiums. Next we report the results of a counterfactual exercise in which insurers are allowed to adjust both drug spending and premiums. Here, PDP plans increase drug spending by roughly the same amount, but also increase their premiums by 10% given the additional drug costs.

The counterfactual benefit design choices have important implications for both enrollees and government budgets. Critically, the increase in spending on prescription drugs leads to a fiscal savings for the Medicare program through lower medical expenditures. Our counterfactuals show that stand-alone PDP plans would spend \$80 more per year if they took medical expenditure into account, implying an offset of \$21.60 per enrollee per year. Multiplying \$21.60 by 17.5 million, the number of stand-alone PDP enrollees in 2008, we find that PDP plans impose an externality of \$378 million per year; based on our bootstrapped estimates, the confidence interval for our estimate ranges from \$325 million per year to \$431 million per year.⁵¹

We also calculate a measure of consumer surplus (CS) that assumes the parameters of the plan utility function we

⁵⁰We solve for drug costs, rather than phase-specific average out-of-pocket costs, for three reasons. First, drug costs enter into the plan's first order condition directly. Second, from a policy perspective, we are primarily interested in impact of different incentives on drug spending. Mechanically, costs change to fit the model in which the elasticity is taken as given and consumption increases as the cost-sharing schedule changes. Finally, given the monotonic relationship between the insurer's choice variables and both OOPC and drug spending, it suffices to solve for the equilibrium outcome. The first panel of Table 7 presents the baseline results.

⁵¹We would obtain a similar estimate if we used the implied additional spending by PDP plans in Table 7. By contrast, MedPAC (Medpac (2015a)) estimates that spending on an equivalent enrollee in a MA-PD plan is approximately 2% higher than traditional Medicare. The average total Medicare spending was approximately \$10,000 per enrollee in 2008 (\$200 of additional spending in MA); the externality due to offsets does not, on its own, imply greater efficiency in MA-PD plans. However, the externality provides evidence of a potential channel through which MA-PD plans can obtain efficiency gains.

By contrast, McWilliams et al. (2016) find that the much-discussed Medicare Shared Savings Program led to an aggregate \$238 million spending reduction in the early years; our larger results indicate the potential power of incentive alignment in equilibrium. Our results illustrate the value of insurer incentive alignment in the Medicare Advantage program and the potential of more efficient benefit design by private insurers, who may align benefit design more efficiently than the federal government (see also Finkelstein, Einav and Polyakova (2016)). To give additional context, the changes in MA over-payments due to risk adjustment policies measured in Brown et al. (2014) are approximately \$317, or about twice as large as our causal increase in drug spending. Furthermore, Table A.5 shows that it would be costly for the government to achieve the increased drug consumption generated by MA-PD plans using a flat cost-sharing subsidy alone; this calculation highlights the advantage of nimbler private insurers. However, there is no reason that the profit maximization incentives of MA-PD plans necessarily align with any social welfare criterion. Therefore, another natural policy intervention would be to better align consumer plan choices with value (from a societal perspective, including any externalities on the traditional Medicare program). Mechanically, we implement this by setting the coefficient on OOPC in the demand system equal to the coefficient on premiums, such that consumers treat a \$1 increase in premiums equal to a \$1 increase in OOPC. The results are in the final two columns of the top panel of Table A.6. If consumers were more "sophisticated" about the potential for underutilization, plans would increase their generosity. MA-PD plans would spend 5.6% more on prescription drugs, while PDP plans would spend 10.4% more. This increased spending by stand-alone PDP plans is less than the amount that fully internalizes the fiscal externality, yet shows that public policies that align consumer demand or the structure of subsidies with providing value will incentivize insurers to offer contracts that reduce costs or increase health.

estimate represent primitive preferences. In the counterfactual in which insurers cannot adjust their prices, consumers like more generous plans and are made better off; the effect is slightly larger than the fiscal savings to the Medicare program. However, when plans are allowed to adjust their premiums, they raise prices (though by less than the increase in benefit generosity). Given that consumers dislike premium increases more than they like more generous plans, enrollees are made slightly worse off.

5.2 The Relative Importance of Strategic Incentives and Adverse Selection

We find that strategic incentives created by offsets drive benefit design in the prescription drug insurance market. Yet a long literature, dating back to the theoretical contributions of Rothschild and Stiglitz (1976), argues that adverse selection should have a significant impact on benefit design. To better understand the magnitude of strategic incentives and selection, we estimate a counterfactual in which we eliminate selection on the intensive margin. We set $\theta_1 = -1$ in addition to internalizing the externality as in the previous counterfactual. As discussed in the previous section, absent asymmetric information or offsets, we expect a one-for-one relationship between insurer expenditure and enrollee costs. In the counterfactual, all plans face incentives to internalize offsets, but we have eliminated baseline adverse selection (along with any moral hazard) among stand-alone PDPs.

The results are presented in the second panel of Table 7 and show that MA-PD plans would become more generous absent selection. The magnitude of the effect is the same order of magnitude as the effect of strategic incentives in the previous counterfactual, though the effects are smaller if premiums are allowed to adjust. Stand-alone PDPs also become more generous: they now spend 13% more on prescription drugs. Given the counterfactual estimates, we conclude that strategic incentives due to benefit integration are as important as selection in driving market outcomes.⁵²

We perform two additional counterfactuals to explore the extent to which other features of the institutional environment drive benefit design. First, we estimate our main counterfactual using our baseline demand estimates, which do not account for plan vintage and imply substantially less adverse selection. In this counterfactual, stand-alone PDPs are much more generous than in our preferred estimates. This implies that accounting for inertia is critical and that selection – in addition to strategic incentives – plays an important role in benefit design.

Finally, we explore the effect of imperfect competition on benefit design and premium setting. We estimate a counterfactual in which plans are constrained to pay out 80% of the premiums they collect in (drug) claims, known as a (binding) minimum loss ratio regulation. The regulation (enacted as part of the Affordable Care Act) is essentially a blunt tool that caps insurer profits and reduces incentives to price strategically. If stand-alone PDPs account for medical expenditure under this pricing rule, they become more generous, consistent with previous results. However, MA-PD plans actually become substantially less generous – largely because they cannot cross-subsidize prescription

⁵²This finding is consistent with Starc (2015), which finds that strategic firm incentives, rather than adverse selection, drive firm behavior (in that instance, premiums) and margins in a similar setting.

Table 7: Counterfactuals

	Counterfactual Plan Characteristics		Counterfactual Plan Characteristics	
	MA	PDP	MA	PDP
Internalize Externality: $\theta_{PDP} = \hat{\theta}_{MA} = \theta_1 + \theta_2$				
Premium	206.21	406.58	214.02	439.41
% Change	-	-	2%	10%
Insurer Drug Spend	1332.91	1298.72	1253.89	1295.56
% Change	4%	7%	-2%	7%
Change in CS (\$/Enrollee)	10.98	25.89	-13.29	-5.40
Fiscal Savings (\$/Enrollee)		21.60		22.88
Premium Adjustment	no		yes	
No Selection, No Externality: $\theta_{PDP} = \theta_{MA} = -1 + \theta_2$				
Premium	206.21	406.58	159.48	392.43
% Change	-	-	-0.14	0.08
Insurer Drug Spend	1373.46	1369.69	1299.33	1390.95
% Change	7%	13%	1%	15%
Change in CS (\$/Enrollee)	18.35	50.99	-3.94	24.62
Fiscal Savings (\$/Enrollee)		37.62		44.32
Premium Adjustment	no		yes	
No Inertia, Internalize Externality:			Fixed Loss Ratio, Internalize Externality:	
$\theta_{PDP} = \hat{\theta}_{MA} = \theta_1 + \theta_2$				
Premium	231.42	459.22	194.86	358.55
% Change	4%	15%	-6%	-12%
Insurer Drug Spend	1337.38	1394.30	1121.04	1288.45
% Change	12%	13%	-13%	6%
Change in CS (\$/Enrollee)	5.49	29.27	-23.53	72.89
Fiscal Savings (\$/Enrollee)		32.83		20.74
Premium Adjustment	no		yes	
Policy Counterfactual: Reduction in Reinsurance Generosity				
Premium	206.21	406.58	159.96	276.37
% Change	-	-	-0.22	-0.32
Insurer Drug Spend	1202.86	1071.14	1166.28	1029.81
% Change	-6%	-12%	-9%	-15%
Change in CS (\$/Enrollee)	-16.45	-47.27	-24.71	-58.81
Fiscal Savings (\$/Enrollee)*		-36.52		-45.43
Premium Adjustment	no		yes	

Notes: Results are calculated as described in Section 5. Means across markets are reported, as well as the % change from baseline.

drug expenditure. As a result, enrollees in MA-PD plans are actually made slightly worse off despite lower premiums, whereas enrollees in stand-alone PDPs are made substantially better off.

5.3 Alternative Subsidy Policies

The Medicare Payment Advisory Commission has recommended that Medicare significantly lower the amount of reinsurance paid in the catastrophic phase from 80% to 20% while increasing the premium subsidy to maintain the overall generosity level of government subsidies (Medpac (2016)). We use our model to simulate the impact of this policy in equilibrium. This counterfactual policy also illustrates the key mechanisms of selection and strategic incentives. The commission notes that while the change could exacerbate the incentives created by selection, it also creates incentives for insurers to reduce costs. We expect policies that exacerbate adverse selection to lead to less generous insurance plans in equilibrium; mechanically, we re-solve for firm incentives and endogenous plan characteristics under the alternative subsidy structure.

The results are in Table 7. Consistent with selection driving plan design, plans are much less generous absent reinsurance: stand-alone PDPs spend 12-15% less on prescription drugs under the alternative policy. Consistent with strategic incentives driving plan design, the effect is muted for MA-PD plans, which spend 6-9% less on prescription drugs. Consistent with previous counterfactual results, selection and strategic incentives are equally important determinants of plan characteristics. The policy harms consumers and is costly to the federal government, as consumers consume fewer drugs, leading to additional medical expenditure.⁵³

In the appendix, we explore additional counterfactual policies, including those aimed at increasing generosity by covering a fraction of the plan's cost. These results highlight the importance of taking strategic insurer incentives into account when designing policy; insurers will optimize in the face of consumer demand and government regulation.

6 Conclusion

This paper examines how health insurers set premiums and design benefits in equilibrium. We build on empirical literatures that estimate structural models of insurer decisions and model endogenous product characteristics to show how cost-side incentives affect insurance benefit design. We examine these issues in the Medicare Advantage and Part D markets and show that differences in incentives across plan types drive the generosity of the benefits.

MA-PD plans offer more generous prescription drug plans than their stand-alone counterparts; this increased generosity is concentrated in those drug categories with large offsets. Our model of firm behavior highlights the mechanisms that drive this differential: MA-PD plans have an incentive to internalize the effect of medical care

⁵³We hold upstream prices fixed in this simulation; insurers may be incentivized to negotiate additional discounts from manufacturers under the alternative subsidy structure.

offsets. By measuring firm incentives, we are able to calculate the size of the implied offset. Our estimate of an approximately 27% offset is similar in magnitude to demand-side estimates in the literature (Chandra, Gruber and McKnight (2010); Gaynor, Li and Vogt (2007)). The counterfactuals show how policy changes can increase plan incentives to help enrollees internalize offsets and explore the impact of alternative subsidies.

Much of the empirical work on health insurance markets has focused on premium setting. Our model extends this logic from premium setting (Handel (2013); Starc (2015); Town and Liu (2003); Tebaldi (2017); Ericson and Starc (2015); Decarolis, Polyakova and Ryan (2015)) to incorporate endogenous plan design. Strategic incentives meaningfully affect plan features, which in turn impact both health and health care costs. As a result, government policy plays an important and understudied role in determining how market forces will affect health care utilization and consumer welfare in the Medicare program.

References

- Abaluck, Jason and Jonathan Gruber. 2011. "Choice Inconsistencies among the Elderly: Evidence from Plan Choice in the Medicare Part D Program." *American Economic Review* 101(4):1180–1210.
- Abaluck, Jason, Jonathan Gruber and Ashley Swanson. 2015. "Prescription Drug Use under Medicare Part D: A Linear Model of Nonlinear Budget Sets." NBER Working Paper No. 20976.
- Afendulis, Christopher C, Michael E Chernew and Daniel P Kessler. 2013. "The Effect of Medicare Advantage on Hospital Admissions and Mortality." NBER Working Paper No. 19101.
- Afendulis, Christopher C., Yulei He, Alan M. Zaslavsky and Michael E. Chernew. 2011. "The Impact of Medicare Part D on Hospitalization Rates." *Health Services Research* 46(4):1022–1038.
- Baicker, Katherine and Dana Goldman. 2011. "Patient Cost-sharing and Healthcare Spending Growth." *Journal of Economic Perspectives* 25(2):47–68.
- Baicker, Katherine, Michael E. Chernew and Jacob Robbins. 2013. "The Spillover Effect of Medicare Managed Care: Medicare Advantage and Hospital Utilization." *Journal of Health Economics* 32:1289–1300.
- Baicker, Katherine, Sendhil Mullainathan and Joshua Schwartzstein. 2015. "Behavioral Hazard in Health Insurance." *Quarterly Journal of Economics* 130(4):1623–1667.
- Berry, Steven. 1994. "Estimating Discrete Choice Models of Product Differentiation." *RAND Journal of Economics* 25:242–262.
- Berry, Steven and Panle Jia. 2010. "Tracing the Woes, an Empirical Analysis of the Airline Industry." *American Economic Journal: Microeconomics* 2(3):1–43.
- Brot-Goldberg, Zarek C, Amitabh Chandra, Benjamin R Handel and Jonathan T Kolstad. 2017. "Consumer Heterogeneity and Medical Care Price Responsiveness: Evidence and Implications for Optimal Insurance Design." *Quarterly Journal of Economics* 132(3):1261–1318.
- Brown, J., M. Duggan, I. Kuziemko and W. Woolston. 2014. "How Does Risk-Selection Respond to Risk-Adjustment: Evidence from the Medicare Advantage Program." *American Economic Review* 104(10).
- Cabral, M., M. Geruso and N. Mahoney. 2014. "Does Privatized Health Insurance Benefit Patients or Producers? Evidence from Medicare Advantage." NBER Working Paper No. 20470.
- Cabral, M. and N. Mahoney. 2014. "Externalities and Taxation of Supplemental Insurance: A Study of Medicare and Medigap."

- Carey, C.M. 2017. “Technological Change and Risk Adjustment: Benefit Design Incentives in Medicare Part D.” *American Economic Journal: Economic Policy* 9(1):38–73. Cornell University.
- CBO. 2012. Offsetting Effects of Prescription Drug Use on Medicare Spending for Other Health Care Services. Technical report Congressional Budget Office.
- Chandra, Amitabh, Jonathan Gruber and Robin McKnight. 2010. “Patient cost-sharing and hospitalization offsets in the elderly.” *American Economic Review* 100(1).
- Chernew, Michael E, Allison B Rosen and A Mark Fendrick. 2007. “Value-based insurance design.” *Health Affairs* 26(2):w195–w203.
- Crawford, G. 2012. “Accommodating Endogenous Product Choices: A Progress Report.” *International Journal of Industrial Organization* 30:315–320.
- 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, Christina M., Gautam Gowrisankaran and Robert Town. 2015. “Myopia and Complex Dynamic Incentives: Evidence from Medicare Part D.” NBER Working Paper No. 21104.
- Decarolis, F. 2015. “Medicare Part D: Are Insurers Gaming the Low Income Subsidy Design?” *American Economic Review* 105(4):1547–80.
- Decarolis, Francesco, Maria Polyakova and Steven P. Ryan. 2015. “The Welfare Effects of Supply-Side Regulations in Medicare Part D.” NBER Working Paper No. 21298.
- Draganska, M., M. Mazzeo and K. Seim. 2009. “Beyond Plain Vanilla: Modeling Joint Product Assortment and Pricing Decisions.” *Quantitative Marketing and Economics* 7(2).
- Duggan, Mark, Amanda Starc and Boris Vabson. 2016. “Who Benefits when Government Pays more? Pass-through in the Medicare Advantage Program.” *Journal of Public Economics* . NBER Working Paper No. 19989.
- Einav, Liran, Amy Finkelstein and Paul Schrimpf. 2015. “The Response of Drug Expenditures to Non-linear Contract Design: Evidence from Medicare Part D.” *Quarterly Journal of Economics* 130(2):841–899.
- Eizenberg, A. 2014. “Upstream Innovation and Product Variety in the U.S. Home PC Market.” *Review of Economic Studies* 81:1003–1045.
- Ellis, Randall P., Shenyi Jiang and W. Manning. 2015. “Optimal Health Insurance for Multiple Goods and Time Periods.” *Journal of Health Economics* 41:89–106.

- Ellison, Glenn. 2006. *Bounded rationality in industrial organization*. Vol. 42 Cambridge University Press p. 142.
- 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.
- Ericson, Keith and Amanda Starc. 2015. “Pricing Regulation and Imperfect Competition on the Massachusetts Health Insurance Exchange .,” *Review of Economics and Statistics* 97(3):667–682.
- Fan, Y. 2013. “Ownership Consolidation and Product Characteristics: A Study of the US Daily Newspaper Market.” *American Economic Review* 103(5):1598–1628.
- Finkelstein, A., L. Einav and M. Polyakova. 2016. “Private Provision of Social Insurance: Drug-specific Price Elasticities and Cost Sharing in Medicare Part D.” NBER Working Paper No. 22277.
- Finkelstein, A., M. Gentzkow and H. Williams. 2016. “Sources of Geographic Variation in Health Care: Evidence from Patient Migration.” NBER Working Paper 20789.
- Gaynor, M., J. Li and W. Vogt. 2007. “Substitution, Spending Offsets, and Prescription Drug Benefit Design.” *Forum for Health Economics and Policy* 10(2):1–31.
- Geruso, M. and T. J. Layton. 2018. “Selection in Health Insurance Markets and Its Policy Remedies.” *Journal of Economic Perspectives* 31(4):23–50.
- Geruso, Michael, Timothy J. Layton and Daniel Prinz. 2016. “Screening in Contract Design: Evidence from the ACA Health Insurance Exchanges.” NBER Working Paper No. 22832.
- Glazer, J. and T. McGuire. 2013. “A Welfare Measure of Offset Effects in Health Insurance.” *Journal of Public Economics* 96(5-6):520–523.
- Goldman, D. and Tomas J. Philipson. 2007. “Integrated Insurance Design in the Presence of Multiple Medical Technologies.” *American Economic Review* 97(2):427–432.
- Goldman, Dana P, Geoffrey F Joyce and Yuhui Zheng. 2007. “Prescription Drug Cost Sharing: Associations with Medication and Medical Utilization and Spending and Health.” *Journal of the American Medical Association* 298(1):61–69.
- Gowrisankaran, G., K. Norberg, S. Kymes, M. E. Chernew, D. Stwalley, L. Kemper and W. Peck. 2013. “A Hospital System’s Wellness Program Linked To Health Plan Enrollment Cut Hospitalizations But Not Overall Costs.” *Health Affairs* .

- Handel, Benjamin. 2013. "Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts." *American Economic Review* (103):2643–2682.
- Heiss, F., D. McFadden, J. Winter, A. Wupperman and B. Zhou. 2016. "Inattention and Switching Costs as Sources of Inertia in Medicare Part D." NBER Working Paper No. 22765.
- Ho, K., J. Hogan and F. Scott Morton. 2015. "The Impact of Consumer Inattention on Insurer Pricing in the Medicare Part D Program." NBER Working Paper No. 21028.
- Hoadley, J., L. Summer, E. Hargrave, J. Cubanski and T. Neuman. 2014. "Medicare Part D in Its Ninth Year: The 2014 Marketplace and Key Trends, 2006-2014."
- Hortascu, A., S.A. Madanizadeh and S. Puller. 2015. "Power to Choose? An Analysis of Consumer Inertia in the Residential Electricity Market." NBER Working Paper No. 20988.
- Hsu, J., J. Huang, V. Fung, M. Price, R. Brand, R. Hui, B. Fireman, W. Dow, J. Bertko and J. Newhouse. 2009. "Distributing 800 Billion: An Early Assessment Of Medicare Part D Risk Adjustment." *Health Affairs* .
- Ketcham, J., C. Lucarelli, E. Miravete and M. Roebuck. 2012. "Sinking, Swimming, or Learning to Swim in Medicare Part D." *American Economic Review* 102(6):2639–2673.
- Lavetti, K. and K. Simon. 2014. "Strategic Formulary Design in Stand-Alone and Integrated Medicare Advantage Prescription Drug Plans."
- Lustig, J. 2010. "Measuring welfare losses from adverse selection and imperfect competition in privatized medicare." Boston University Department of Economics.
- Maciejewski, Matthew L, Chris L Bryson, Mark Perkins, David K Blough, Francesca E Cunningham, John C Fortney, Sarah L Krein, Kevin T Stroupe, Nancy D Sharp and Chuan Fen Liu. 2010. "Increasing Copayments and Adherence to Diabetes, Hypertension, and Hyperlipidemic Medications." *The American Journal of Managed Care* 16(1):e20–34.
- Maciejewski, Matthew L, Joel F Farley, John Parker and Daryl Wansink. 2010. "Copayment reductions generate greater medication adherence in targeted patients." *Health affairs* 29(11):2002–2008.
- Manning, W., J. Newhouse, N. Duan, E. Keeler, B. Benjamin, A. Leibowitz and M. Marquis. 1987. "Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment." *American Economic Review* 77(3):251–277.
- McGuire, T., J. Newhouse and A. Sinaiko. 2011. "An Economic History of Medicare Part C." *Milbank Quarterly* .

- McGuire, Thomas G. 2011. Demand for health insurance. In *Handbook of health economics*. Vol. 2 Elsevier pp. 317–396.
- McWilliams, J. Michael, Alan M. Zaslavsky and Haiden A. Huskamp. 2011. “Implementation of Medicare Part D and Nondrug Medical Spending for Elderly Adults With Limited Prior Drug Coverage.” *The Journal of the American Medical Association* 306(4):402–409.
- McWilliams, J.M., L. Hatfield, M.E. Chernew, B. Landon and A. Schwartz. 2016. “Early Performance of Accountable Care Organizations in Medicare.” *New England Journal of Medicine* 374:2357–2366.
- Medpac. 2015a. “The Medicare Advantage program: Status report.”.
- Medpac. 2015b. Report to the Congress: Medicare and the Health Care Delivery System. Technical report Medpac.
- Medpac. 2016. Report to the Congress: Medicare and the Health Care Delivery System. Technical report Medpac.
- Miller, D. and J. Yeo. 2014. “A Dynamic Discrete Choice Model of Switching Costs in Medicare Part D.” Clemson University and Singapore Management University Unpublished Mimeo.
- Newhouse, J., M. Price, J. McWilliams, J. Hsu and T. McGuire. 2015. “How Much Favorable Selection is Left in Medicare Advantage?” *American Journal of Health Economics* .
- NICE. 2011. Hypertension in adults: diagnosis and management. Technical report National Institute for Health and Care Excellence.
- Nosal, K. 2011. “Estimating Switching Costs for Medicare Advantage Plans.” Mannheim.
- Polyakova, M. 2016. “Regulation of Insurance with Adverse Selection and Switching Costs: Evidence from Medicare Part D.” *American Economic Journal: Applied Economics* 8(3):165–195.
- Rothschild, Michael and Joseph Stiglitz. 1976. “Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information.” *Quarterly Journal of Economics* pp. 629–649.
- Starc, A. 2015. “Insurer Pricing and Consumer Welfare: Evidence From Medigap.” *RAND Journal of Economics* 45(1):198–220.
- Sweeting, A. 2010. “The Effects of Mergers on Product Positioning: Evidence from the Music Radio Industry.” *RAND Journal of Economics* 41(2):372–97.
- Tebaldi, P. 2017. “Estimating Equilibrium in Health Insurance Exchanges: Price Competition and Subsidy Design under the ACA.” 2017.

Town, R. and S. Liu. 2003. "The Welfare Impact of Medicare HMOs." *RAND Journal of Economics* 32(4):719–736.

Wollman, T. 2016. "Trucks without bailouts: Equilibrium product characteristics for commercial vehicles."

Wu, Y. 2016. "Supply Response to Consumer Inertia: Strategic Pricing in Medicare Part D." MIT Unpublished Memo.

Appendices

A Utilization and Insurer Spending

A.1 Reduced Form Relationship Between Plan Characteristics and Insurer Costs

We regress insurer costs on our key product characteristics: $x_{jmt} = \left[P_{jmt}^{ICR}, P_{jmt}^{Donut} \right]'$. The results are in Table A.1. In all specifications, we control for contract fixed effects, the beneficiary quintile of lagged total drug spending to capture underlying contract characteristics and consumer risk.⁵⁴ In our second and third specifications, we also control for demographic characteristics (age, race, and gender), which capture part of the observable risk. In our final, most conservative specifications, we also control for historical county-level FFS spending, which proxies for county level variation in medical services, including drugs, that might be driven by differences in patient preferences, medical care infrastructure and the physician culture (see Finkelstein, Gentzkow and Williams (2016)).

Table A.1: Plan Characteristics and Insurer Costs

Dependent Variable: Insurer Costs						
	(1)	(2)	(3)	(4)	(5)	(6)
P_j^{ICR}	-672.9***	-673.4***	-672.4***	-674.2***	-673.6***	-675.6***
	(95.23)	(139.9)	(94.65)	(140.9)	(94.93)	(140.4)
P_j^{Donut}	-25.39	-66.38**	-25.60	-65.09**	-25.09	-65.62**
	(18.02)	(30.78)	(18.02)	(30.55)	(17.94)	(30.47)
$1(MA) * P_j^{ICR}$		-1.014		-0.279		1.241
		(158.0)		(158.6)		(158.2)
$1(MA) * P_j^{Donut}$		54.86*		53.40		54.48*
		(32.68)		(32.49)		(32.47)
FFS Costs					0.388***	0.388***
					(0.0375)	(0.0374)
R-Squared	0.243	0.245	0.245	0.243	0.246	0.246
Observations	569,078	569,078	569,078	569,078	569,078	569,078
Year FE	X	X	X	X	X	X
Type FE	X	X	X	X	X	X
Demo. Controls		X	X		X	X
Plan FE	X	X	X	X	X	X

Notes: Table presents OLS regression models, where outcome variable is insurer costs. The unit of observation is an enrollee-year for the 2007-2009 period. The original data are obtained from a 10% sample of CMS prescription drug event files, aggregated to the enrollee-year level. We include year-level indicators and indicators for the quintile of 2006 spending (consumer types) in all specifications. In some specifications, we also control for 5-yr average per capita Medicare FFS spending and demographic controls for age categories, race, and gender. Standard errors are clustered at the contract level and are presented in parentheses. Statistical significance at the 10%, 5%, and 1% levels are denoted by *, **, and *** respectively.

The results show that higher costs for enrollees are associated with lower insurer drug costs: a \$1 increase in enrollee OOPC per days supply in the ICR is associated with \$675 less in insurer drug spending. The effect for the

⁵⁴Therefore, we are leveraging within contract variation across geographic space and over time.

donut hole is predictably smaller, as only a small fraction of enrollees enter the donut hole. Critically, we see in columns 2, 4, and 6 that the effect does not vary by plan type: MA-PD and stand-alone plans find it equally costly – in terms of prescription drug costs only – to increase plan generosity.⁵⁵ That is, on average, MA-PD plans do not have a cost advantage in the provision of their drug benefits and differential costs are not a likely source of the benefit design differential.

A.2 Behavioral Response and Elasticity Estimates

In order to estimate elasticities using the approach described in Dalton, Gowrisankaran and Town (2015), we estimate parameters from the following regression i:

$$Y_{it} = FE_i + \lambda_1 1\{2400 < Spend_{it} \leq Donut_t\} + \lambda_s 1\{Spend_{it} > Donut_t\} + v_{it}, \quad (17)$$

where Y_{it} is the weekly drug spending for individual i in week t , FE_i is an individual fixed effect, $Spend_{it}$ is the total drug spending in that year at the beginning of week t for individual i , $Donut$ is the dollar value of the initiation of the coverage gap phase of the benefit package in that year, and v_{it} is an i.i.d. error term. The individual fixed effects control for individual selection into the donut hole. Identification is driven by variation in the timing of entering the donut hole controlling for mean individual characteristics. Given our sample selection, λ_s coefficient measures the mean price impact of transitioning into the donut hole phase. We estimate this equation separately for PDP and MA-PD enrollees.

The parameter estimates from are presented in Table A.2. Consistent with our reduced form analysis, OOPC increase as consumers enter the donut hole: the average price increase by 15% for consumers in stand-alone PDPs and 13% for consumers in MA-PD plans. Consistent with the results presented in Figure 3, the parameter estimates show that upon entering the donut hole, PDP enrollees drop their spending an average of 10%. In contrast, MA-PD enrollees reduce their spending by 14% upon entering the donut hole; the larger response reflects both a great extensive margin demand elasticity and a shift to lower cost drugs.⁵⁶ Comparing these estimates to the estimates of the differences in OOPC between PDP and MA-PD plans implies a donut hole elasticity ranging from -0.53 (stand-alone plans) to -0.79 (MA-PD plans).

⁵⁵Table A.8 shows that these differences translate into lower OOPC for consumers.

⁵⁶There are two points to note here. First, we argue below that MA-PD enrollment leads to a causal increase in average consumption. Second, we find no evidence of an increased drop in consumption upon entering the donut hole among high offset drugs. Exploring heterogeneity in this effect is an interesting direction for future work.

Table A.2: Behavioral Response

	(1)	(2)	(3)
	ln(1 + Spending)	ln(1 + Days Supply)	ln(1 + OOPC)
At Kink	-0.0564*** (0.00377)	-0.0439*** (0.00331)	0.114*** (0.000999)
At Kink * MA	-0.00761 (0.00770)	0.000704 (0.00691)	-0.0123*** (0.00227)
Post Kink	-0.103*** (0.00217)	-0.0804*** (0.00190)	0.151*** (0.000684)
Post Kink * MA	-0.0364*** (0.00459)	-0.0227*** (0.00409)	-0.0208*** (0.00154)
N	5,220,060	5,220,060	2,452,070

Notes: Standard errors are in parentheses. All regressions include enrollee fixed effects and cluster standard errors at the enrollee level. Each observation is an enrollee-week for enrollees that have beginning-of-week spending in the interval [Kink - \$510, Kink) sometime between week 14 and week 30 of the calendar year and enrollee-weeks with beginning-of-week spending in the interval [Kink - \$510, Kink + \$490]. * p<0.05 ** p<0.01 *** p<0.001

A.3 Total Spending

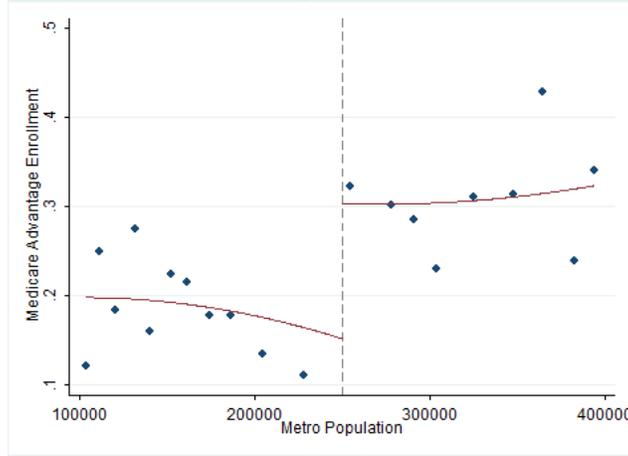
Our goal is to estimate the causal impact of MA-PD enrollment on total drug consumption, insurer, and enrollee drug costs. However, a naive estimate will be contaminated by selection, as MA-PD enrollees are likely healthier than non-MA enrollees. Therefore, on average, MA-PD enrollees will have lower drug expenditure than their counterparts in stand-alone PDPs for reasons other than plan design. This is likely to be true even once we control for a rich set of individual characteristics.

Following a series of papers (Afendulis, Chernew and Kessler (2013); Cabral, Geruso and Mahoney (2014); Duggan, Starc and Vabson (2016)), we rely on a statutory discontinuity in MA-PD plan reimbursement to identify the causal impact of MA-PD enrollment. For counties with relatively low fee-for-service (FFS) spending, payment is set equal to a floor. Beginning in 2003, differential floors were set for urban and rural counties – approximately two-thirds of counties are floor counties. Higher reimbursement in urban counties led to more plan entry and higher MA penetration rates (Duggan, Starc and Vabson (2016)). The evidence strongly indicates that this variation in MA penetration rates is driven by the differential MA subsidies and is not correlated with individual health risk or other demand side factors. Furthermore, because an urban county is somewhat arbitrarily defined as one with 250,000 or more in metro population, it is natural to focus the analysis on comparable counties near each side of the threshold. Consumers in urban floor counties close to the threshold are more likely to be enrolled in MA-PD plans than consumers in similar rural floor counties just to the right of the urban threshold.⁵⁷ In our reduced form analysis, we use the county urban/rural status as an instrument in a linear instrumental variable specification; our empirical strategy is a fuzzy regression discontinuity approach.

The variation we use in our IV specifications is highlighted in Figure A.1, which plots the probability of MA-PD

⁵⁷We will also use urban status to predict the inside share of MA-PD plans in the plan choice models.

Figure A.1: Effect of Population on MA Enrollment



Notes: Plots a binscatter with twenty population bins. We drop counties with FFS spending above the urban floor, and control for beneficiary age, sex, race, 2006 spending type, and county-level FFS spending. Lines represent a quadratic fit.

enrollment as a function of population. This figure depicts a binscatter plot with twenty population bins. We control for consumer demographics, including risk type, as well annual mean county-level FFS spending and plot the average probability of MA-PD enrollment. We fit quadratic curves on either side of the 250,000 population cutoff. We see a dramatic change in the probability of MA-PD enrollment just to the right of the discontinuity. We implement our identification strategy using an instrument variables framework. Specifically, we estimate:

$$y_{itjm} = X_{mt}^1 \beta_1 + X_{it}^2 \beta_2 + \beta_3 1(MA) + g_1(pop_{mt}) + \mu_{itj}, \quad (18)$$

$$1(MA) = X_{mt}^1 \gamma_1 + X_{it}^2 \gamma_2 + \gamma_3 1(urban_{mt}) + g_2(pop_{mt}) + v_{itj}, \quad (19)$$

where β_3 is the coefficient of interest, and X_{mt}^1 and X_{it}^2 are vectors of market and individual specific covariates, respectively. In all specifications, we control flexibly for metro area population. The dependent variables of interest, y_{itjm} , are total drug spending, consumer OOPC, and insured costs, where j represents a plan. We hypothesize that insured spending is causally higher in MA-PD plans, and consumer out-of-pocket costs lower. These relationships are due directly to plan design on the part of insurers; the overall impact of these changes on total expenditure is more ambiguous, as it depends on the size of the behavioral response, but likely to be positive as well.

To explore the impact of MA enrollment on consumption, we focus on the 2007-2009 time period. In all specifications, we control for the consumer quintile of 2006 drug spending, calculated at the national level. In our second and third specifications, we also control for demographic characteristics (age, race, and gender), which capture part of the observable risk. In our final, most conservative specifications, we also control for historical county-level FFS spending, which proxies for county level variation in medical services, including drugs, that might be driven by differences

in patient preferences, medical care infrastructure and the physician culture (see Finkelstein, Gentzkow and Williams (2016)). We focus the analysis on consumers living in counties with associated metro populations between 100,000 and 400,000.

Table A.3 reports the results of OLS regressions of total expenditure, OOPC, and insurer spending. These results are likely biased because of adverse selection into PDP plans – we report them in order to provide a benchmark to the IV estimates. In the bottom panel, we examine the impact on total expenditure. The first column, which controls only for year and the quintile of 2006 spending, shows that the average MA-PD enrollee has lower drug expenditures: total spending on drugs is \$252 less than their counterparts in stand-alone PDP plans. The average total expenditure for this sub-sample is \$1697, indicating that MA enrollees have 15% lower drug spending than PDP enrollees. This lower expenditure is associated with savings in the form of out-of-pocket costs to consumers (a reduction of \$178) and somewhat smaller reductions for insurers (\$74 per enrollee per year). The next two columns, which include demographic characteristics and county-level FFS spending, show that the effect is not attenuated by the inclusion of additional controls.

We use changes in MA reimbursement as an instrument for MA-PD coverage. In the first panel of Table A.3, we present the results of the first stage regressions that control for metro population using a cubic spline with knots in increments of 100,000 starting at 150,000. In all specifications, we find that Medicare eligibles in our dataset are 16-17% more likely to enroll in a MA-PD plan if they live in an urban county. Given an average MA market share of 25% within our sub-sample, this represents a large shift.⁵⁸ By exploring what happens to consumers who are exogenously shifted into MA-PD plans, we can isolate the impact of plan design on consumption.

The second panel of Table A.3 shows the estimated impact of MA-PD enrollment on insurer drug costs. Once we account for differential selection, MA-PD plans spend much more on drugs than stand-alone PDPs. The estimate of \$514 in column 4 is approximately half of average insurer spending across all plans (\$1031 per enrollee per year). This estimate is more attenuated in the final column (albeit not statistically different from the estimates in column 4), which includes historical, county-level FFS costs as an additional control. Here the estimates indicate that MA-PD plans spend \$388 more per year than stand-alone PDPs for an equivalent enrollee. As expected, historical FFS spending influences drug consumption: Finkelstein, Gentzkow and Williams (2016) find that approximately half of all variation in spending is due to place-specific supply factors. The following panels describe the impact of additional insurer spending on consumers. The third panel shows that a consumer enrolled in MA can expect to spend \$265 less per year on drugs, holding health risk constant. Consumer spending does not fall one-for-one with the increase in insurer spending, implying that the reduction in average out-of-pocket costs for consumers increases consumption, as confirmed in the final panel. In our preferred estimates, the causal impact of MA-PD enrollment is somewhat noisy, but implies a \$122 increase in drug consumption. On a base of \$1697 of drug spending per year, this represents a 7%

⁵⁸Furthermore, our instrument has a great deal of predictive power. The partial F-stat in the final specification is 509.02.

Table A.3: Impact of MA-PD Enrollment on Drug Spending

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS			IV		
First Stage, Dependent Variable: MA-PD Enrollment						
1 (Urban)				0.168*** (0.00785)	0.170*** (0.00785)	0.177*** (0.00787)
R-squared				0.026	0.036	0.037
Dependent Variable: Insurer Drug Expenditure						
1(MA-PD)	-74.21*** (3.969)	-76.25*** (3.973)	-73.32*** (3.972)	514.2*** (74.25)	506.7*** (73.35)	387.5*** (68.38)
FFS 5 Year Avg. Spend			0.430*** (0.0189)			0.506*** (0.0226)
R-Squared	0.217	0.219	0.221	0.114	0.119	0.159
Dependent Variable: Enrollee Out-of-Pocket Costs						
1(MA-PD)	-177.5*** (2.850)	-174.6*** (2.861)	-173.3*** (2.863)	-215.2*** (55.51)	-222.2*** (54.92)	-265.2*** (52.74)
FFS 5 Year Avg. Spend			0.198*** (0.0160)			0.183*** (0.0183)
R-Squared	0.193	0.195	0.195	0.193	0.194	0.192
Dependent Variable: Total Drug Expenditure						
1(MA-PD)	-251.7*** (5.851)	-250.9*** (5.870)	-246.6*** (5.873)	299.0*** (108.0)	284.6*** (106.7)	122.3 (100.7)
FFS 5 Year Avg. Spend			0.628*** (0.0298)			0.688*** (0.0343)
R-Squared	0.264	0.265	0.267	0.230	0.233	0.252
Year FE	X	X	X	X	X	X
Type FE	X	X	X	X	X	X
Demo. Controls		X	X		X	X
N	381921	381921	381921	381921	381921	381921
Sample	100-400K	100-400K	100-400K	100-400K	100-400K	100-400K

Notes: Table presents linear regression models, where outcome variables are insurer and enrollee costs and total expenditure levels. The unit of observation is at the enrollee-year level, for the 2007-2009 period. The original data are obtained from a 10% sample of CMS prescription drug event files, aggregated to the enrollee-year level. We restrict to those counties in the 100-400k metro population band. We include year-level indicators and indicators for the quintile of 2006 spending in all specifications. In some specifications, we also control for 5-yr average per capita Medicare FFS spending, from 2007. We also include controls for age, age squared, race, and gender as demographic controls. In addition, we include a spline of metro population. Standard errors are clustered at the product level and are presented in parentheses. Statistical significance at the 10%, 5%, and 1% levels are denoted by *, **, and *** respectively.

increase in spending. Total drug consumption increases *despite* a drop in consumer spending.

We hypothesize that the underlying mechanism driving an increase in drug consumption from MA enrollment is differences in MA-PD plan design intended to internalize the impact of offsets on non-drug medical spending. In support of this hypothesis, we show that the effect of MA enrollment on drug consumption is driven entirely by drugs believed to have large offsets *a priori*. We explore the total enrollee level of consumption of “Category 1” drugs, as classified by Chandra, Gruber and McKnight (2010). If these drugs are not taken when prescribed, a serious event, such as a hospitalization, is significantly more likely to occur within the next six months.⁵⁹

Table A.4 describes these results. About 40% of average expenditure (\$648.11) is concentrated in these Category 1 drugs. Consistent with previous specifications, the OLS results are biased downward due to advantageous selection into MA-PD plans. However, the IV specifications in columns 3-6 show a consistent pattern: MA-PD enrollees consume proportionally more of these “Category 1” drugs, due in large part to greater insurer expenditure. MA-PD enrollment leads to an additional \$156 in total spending on these drugs; on a base of \$648, this amounts to a 24% increase, versus 7% for total drug consumption. Put differently, all of the increased total expenditure in MA-PD plans is concentrated in these large offset drugs. Total expenditure in this category increases by \$156, while overall total expenditure increases by \$122, implying that MA-PD plans do not spend more on drugs that are unlikely to have large offsets.

B Additional Counterfactuals

Next, we consider budget neutral policies that attempt to internalize the externality generated by the PDP plans.⁶⁰ Our presumption is that CMS would like to increase drug consumption by PDP enrollees in order to both improve enrollee well-being and to reduce medical care costs. A natural policy to consider is a plan benefit generosity subsidy where CMS would cover some of the plan’s cost to increasing cost-sharing coverage. In order for this cost-sharing subsidy to

⁵⁹Category 1 drugs are “acute care drugs are those that, if not taken, will increase the probability of an adverse health event within a month or two.” These drugs comprise approximately 40% of total drug spending. Category 2 contains “chronic care medications are designed to treat more persistent conditions that, if not treated, will result in a potentially adverse health event within the year (examples include analgesics, antivirals, ACE inhibitors, medications, beta-blockers, hypertension drugs, statins, and glaucoma medications).” Category 3 are “medications that, while necessary to improve patients’ quality of life, will not result in an adverse health event if not taken, because they provide symptom relief as opposed to affecting the underlying disease process (examples are acne medications, antihistamines, motion sickness medications, cold remedies, relief of pain drugs).”

The classes included in Category 1 are Adrenal Corticosteroids, Aminoglycosides, Anaphylaxis Treatment Kits, Anesthesia, Anthelmintics, Antianginals, Antiarrhythmics, Antiasthmatics and broncodilators, Antibacterials, Miscellaneous, Antibiotics, Alkaloids, And Enzymes, Anticoagulants/thrombolytics, Anticonvulsants, Antidotes, Antimalarials, Antimetabolites, Antimycobacterials, Antineoplastics, Antiprotozoals, Antipsychotics/antimanics, Antitoxins/antivenins, Blood Components/substitutes, Blood Glucose Regulators, Cardiac Glycosides, Cardiovascular-renal, Cephalosporins, Chloramphenicol/derivatives, Coronary Vasodilators, Dna Damaging Drugs, Hypotension/shock, Lincosamides and macrolides, Ocular Anti-infective/anti-inflammatory, Penicillins, Polymyxins, Quinolones/derivatives, Repl/regs Of Electrolytes/water Balance, Respiratory Tract, Sulfonamides/related Compounds, Tetracyclines, Vascular Disorders, and Cerebral/peripheral. We exclude drugs that are believed to have differential selection effects, as described in Lavetti and Simon (2014). Drug lists for each category were compiled using lists from drugs.com. Respiratory tract drugs include drugs used to treat asthma and COPD. For drugs with multiple uses, the drug was only included under its primary usage (e.g. etanercept is sometimes used to treat Alzheimer’s Disease, but is much more commonly used for autoimmune diseases such as rheumatoid arthritis, psoriatic arthritis, plaque psoriasis and ankylosing spondylitis. Thus, it is not included on the list of Alzheimer’s drugs).

⁶⁰These calculations do not require knowledge of the “true” demand curve, from which we could derive welfare implications as in Glazer and McGuire (2013).

Table A.4: Impact of MA-PD Enrollment on Spending, Drugs with Large Offsets

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS			IV		
Dependent Variable: Insurer Drug Expenditures						
Mean	401.16					
SD	512.6					
1(MA-PD)	-18.63*** (3.118)	-18.30*** (3.122)	-17.52*** (3.124)	223.5*** (56.20)	229.6*** (55.66)	190.8*** (53.20)
FFS 5 Year Avg. Spend			0.126*** (0.0150)			0.156*** (0.0170)
Mean	0.046	0.047	0.047	0.005	0.005	0.018
Dependent Variable: Enrollee Out-of-Pocket Costs						
Mean	246.96					
SD	379.18					
1(MA-PD)	-58.56*** (1.848)	-56.57*** (1.849)	-56.42*** (1.848)	-27.68 (37.24)	-27.73 (37.24)	-34.43 (35.40)
FFS 5 Year Avg. Spend			0.0238** (0.0103)			0.0270** (0.0116)
R-Squared	0.064	0.065	0.065	0.063	0.064	0.065
Dependent Variable: Total Drug Expenditures						
Mean	648.11					
SD	802.67					
1(MA-PD)	-77.19*** (4.497)	-74.86*** (4.505)	-73.94*** (4.507)	195.8** (84.52)	201.9** (83.64)	156.4* (80.17)
FFS 5 Year Avg. Spend			0.150*** (0.0230)			0.183*** (0.0260)
R-Squared	0.064	0.065	0.066	0.043	0.044	0.051
Year FE	X	X	X	X	X	X
Type FE	X	X	X	X	X	X
Demo.		X	X		X	X
Controls						
N	322,066	322,066	322,066	322,066	322,066	322,066

Notes: Table presents parameter estimates and standard errors of the instrumental variable regression models, where outcome variables are insurer and enrollee costs and total consumption levels. The unit of observation is at the enrollee-year level, for the 2007-2009 period. We restrict to those counties in the 100-400k metro population band. We include year-level indicators and indicators for the quintile of 2006 spending in all specifications. We include controls for age, age squared, race, and gender as demographic controls. In addition, we include a spline of metro population. Standard errors are clustered at the product level and are presented in parentheses. Statistical significance at the 10%, 5%, and 1% levels are denoted by *, **, and *** respectively.

be budget neutral, CMS must also decrease the current premium subsidy, which will likely increase premiums faced by consumers. The impact of such a change depends on how consumers evaluate plans with greater generosity but higher premiums. While it is natural to consider the consumer surplus impact of these policies, such a calculation requires interpreting the utility parameters in the neoclassical context, which, given our earlier findings, is probably inappropriate. For this reason, we refrain from making consumer surplus statements here.

Consider a uniform cost-sharing subsidy for PDP plans, as shown in Table A.5. Mechanically, a subsidy alters both p_{jt} and $OOPC_{jt}$ if it is budget neutral and there is full pass-through; we can write the alternative premium and out-of-pocket costs as a function of the change in OOPC due to a change in the benefit generosity and the offset, which is given by $\frac{\partial q}{\partial x}(c - c')$. For a small change in \mathbf{x} (omitting subscripts for simplicity):

$$OOPC' = OOPC + \frac{\partial OOPC}{\partial \mathbf{x}},$$

$$p' = p + \frac{\partial OOPC}{\partial \mathbf{x}} - \frac{\partial q}{\partial \mathbf{x}}(c - c').$$

In this formulation, the offset savings are passed through completely to the consumer in the form of lower premiums, but the reduced premium subsidy is passed through to consumers in the form of higher premiums as well.⁶¹

A 1% cost-sharing subsidy would increase consumption by 7.2 days supply based on the behavioral elasticities in column 2. The implied offset, in column 3, is \$3.54.⁶² However, the cost-sharing subsidy applies to all of the infra-marginal units as well, and the total reduction in OOPC is \$10.15. Subtracting the offsets, this implies that premiums would have to increase by \$6.61 for the policy to be budget neutral. By contrast, the federal government could eliminate cost-sharing in the donut hole, as the ACA does. Using the calculations in Einav, Finkelstein and Schrimpf (2015), this would increase drug consumption by 8%, generating offsets amounting to \$52.54 per consumer. However, this policy is also expensive: while it reduces OOPC by \$356 per consumer, this reduction comes at a cost net of offsets of \$303. Therefore, if the policy is to be budget neutral, premiums will have to rise dramatically.⁶³

Furthermore, these policies reduce the market share of PDP plans. Consumers do not value the increased generosity at its full cost, as reflected in the measured decision utility; they prefer plans with lower premiums and higher cost-sharing. Therefore, we conclude that it will be difficult for the government to implement broad based changes to the Part D program aimed at reducing externalities that are budget neutral. This includes the recent Part D Enhanced Medication Therapy Management (MTM) Model, which incentivizes stand-alone plans to reduce Parts A and B spending among their enrollees. The actual financial incentives associated with this program are quite small. If plans reduce Parts A and B spending by 2% (about \$200 in 2008), they are eligible for a \$2 per member per month

⁶¹This gives us an upper bound of the potential welfare gain.

⁶²This is calculated as the additional spending multiplied by the 22% figure described above.

⁶³Exacerbating this is the fact that MA-PD plans become more generous in equilibrium, decreasing OOPC to consumers by \$97 per year. We note that the ACA policy is not budget neutral.

Table A.5: Counterfactual Policies

Uniform OOPC Reduction	Elasticity	Offset	Change in OOPC	Effective Cost (to Government)	% Change, PDP Penetration	Consumer Valuation, OOPC Reduction (UB)
1.00%	-0.54	3.54	10.15	6.61	-0.0010	4.30
2.50%	-0.38	6.23	25.28	19.06	-0.0030	10.70
5.00%	-0.33	10.82	50.51	39.69	-0.0064	21.38
10.00%	-0.30	19.67	100.94	81.28	-0.0132	42.74
25.00%	-0.29	47.53	252.30	204.77	-0.0342	106.82
50.00%	-0.29	95.06	504.60	409.54	-0.0744	213.64
75.00%	-0.31	152.43	757.27	604.84	-0.1224	320.61
Eliminate the Donut Hole		52.4	355.99	303.58	-0.0571	152.86

Notes: Results are calculated as described in Section 5.

Table A.6: Counterfactuals

	Baseline		“Sophisticates”:	
	MA	PDP	$\alpha_p = \alpha_p$	
	MA	PDP	MA	PDP
Premium	206.00	407.93	206.00	407.93
% Change	-	-	-	-
Insurer Rx Spend	1285.25	1211.62	1357.41	1337.64
% Change	-	-	0.0564	0.1038

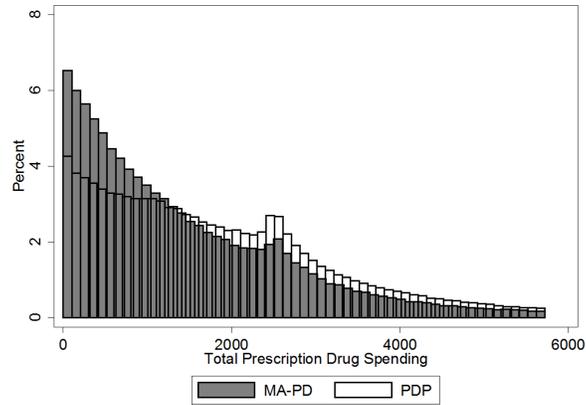
Notes: Results are calculated as described in Section 5. Means across markets are reported, as well as the % change from baseline. Drug spending represents the insured costs.

increase in their benchmark payment. Because firms only receive approximately one-tenth of the savings, they are unlikely to be incentivized to internalize the externality created by offsets. Furthermore, we note that a PDP plan that fully internalized the externality would only spend an additional \$153 per enrollee per year. Given our calculations, this would lead to savings in Parts A and B of about \$30. We find that a policy that provides a \$12.75 per member per month increase in the benchmark payment (and reduces Parts A and B spending by 0.3%) internalizes the externality, and would increase MA-PD enrollment by 3.4 percentage points. Furthermore, we note that a \$153 increase is likely to represent the firms’ entire profit margins, based on a 15% profit margin and the estimates in Ho, Hogan and Scott Morton (2015).

It is costly for the Medicare program to implement a budget neutral policy for two reasons: consumers are not sophisticated with respect to potential underutilization and most implementable policies fail to target marginal consumption, effectively leading to expensive OOPC reductions on infra-marginal units. These results are consistent with a model in which private insurers can better target and subsidize underutilized, high-value care. For example, while the cost-sharing subsidy we describe is uniform, applying to all drugs, private MA-PD insurers can implement more sophisticated contracts that better target increased utilization. We see evidence of this in our reduced form results; Figure A.3 shows that MA-PD plans have lower out-of-pocket costs for exactly those drugs likely to generate the largest offsets. While they increase the complexity of insurance contracts and may exacerbate plan choice frictions, targeted subsidies are more likely to be cost-effective. Therefore, it may be more reasonable to encourage MA-PD enrollment; based on our estimates in Table 4, we believe this can be done in a cost-effective way. For example, rather than closing the donut hole, the federal government could increase MA benchmarks by \$312 per year, plus the \$23 in implied offsets. This would increase MA-PD market share by 7.4%.

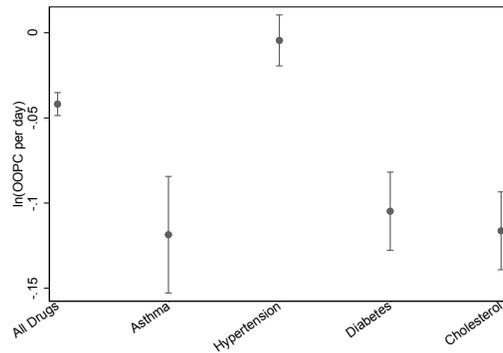
C Robustness Checks

Figure A.2: Histogram of Total Drug Spending by Plan Type, 2008



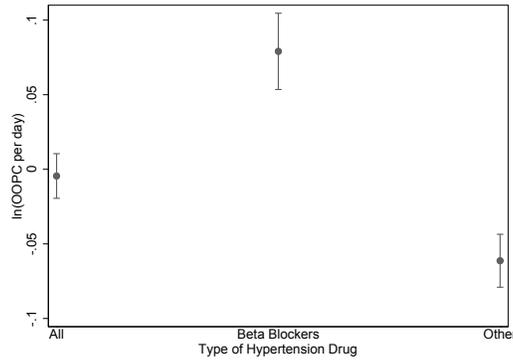
Notes: Plots a histogram of total spending by plan type. For visual simplicity, we drop enrollees spending more than the catastrophic limit (\$5726.25 in 2008). The initial coverage limit in 2008 was \$2510. N=981,813; 387,570 in MA-PD plans and 594,243 in stand-alone PDP plans.

Figure A.3: Out-of-Pocket Cost Effects by Drug Class



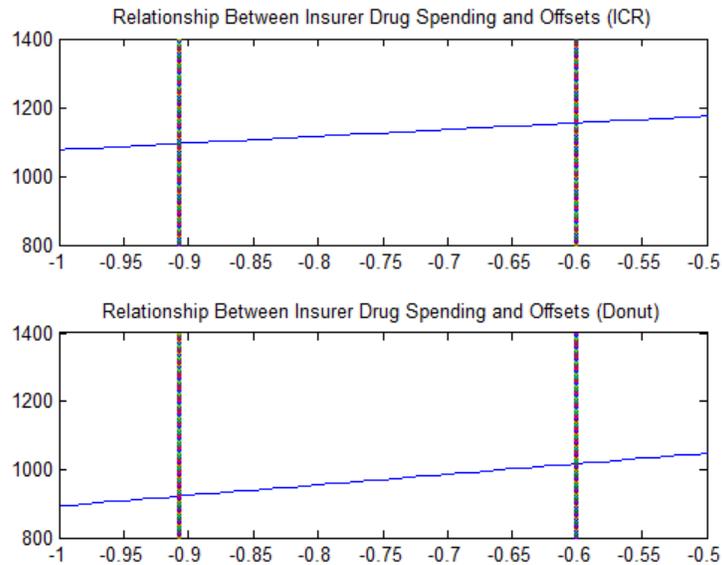
Notes: This figure plots the coefficient of the MA_PD interaction with each of four drug classes, as defined by the World Health Organization. Full regression results are available in Table 2.

Figure A.4: Price Effects by Drug Class



Notes: This figure plots the differences in prices by plan type. Other hypertension drugs include ACE inhibitors, angiotensin II receptor antagonists, renin inhibitors, antiadrenergic agents (centrally & peripherally acting), alpha-adrenergic blockers, aldosterone receptor antagonists, vasodilators and antihypertensive combination therapies. Standard errors are clustered at the plan-product level.

Figure A.5: Supply Side Identification



Notes: This figure plots the optimal level of insurer spending under alternative levels of θ from first-order conditions from both P^{ICR} and P^{Donut} , using average values of the derivatives of shares with respect to premiums and out-of-pocket costs.

Table A.7: Mechanisms

	(1)	(2)
Panel A: Main Results		
	Outcome: 1(90 Day)	
1(MA)	0.00272** (0.000929)	0.00212* (0.000939)
Constant	0.108*** (0.000654)	0.103*** (0.000654)
Observations	153,964,560	153,964,560
Adjusted R-Squared	0.096	0.097
Panel B: Main Results		
	Outcome: Total Cost/Day	
1(MA)	-0.00955 (0.00656)	-0.00468 (0.00657)
Constant	1.960*** (0.00539)	2.204*** (0.0366)
Observations	153,962,060	153,962,060
Adjusted R-Squared	0.590	0.590

Notes: Table presents linear regression models, where outcome variables are as described in each panel. The unit of observation is at the fill level, for the 2007-2009 period. The original data are obtained from a 10% sample of CMS prescription drug event files. We include year-level indicators and product fixed effects in all specifications. In some specifications, we also control the phase of the standard Part D benefit. Standard errors are clustered at the plan-product level. Statistical significance at the 10%, 5%, and 1% levels are denoted by *, **, and *** respectively.

Table A.8: Plan Characteristics and Out-of-Pocket Costs

Dependent Variable: Out-of-Pocket Costs						
	(1)	(2)	(3)	(4)	(5)	(6)
P_j^{ICR}	501.5*** (49.05)	638.9*** (56.45)	499.7*** (48.86)	640.6*** (56.44)	498.8*** (48.53)	637.3*** (55.80)
P_j^{Donut}	-2.442 (9.377)	5.911 (18.19)	-2.129 (9.311)	5.160 (18.22)	-1.758 (9.293)	6.468 (18.12)
$1(MA) * P_j^{ICR}$		-291.4*** (72.49)		-291.2*** (72.94)		-289.8*** (71.88)
$1(MA) * P_j^{Donut}$		-6.640 (16.54)		-6.057 (16.53)		-6.916 (16.52)
FFS Costs					0.281*** (0.0358)	0.281*** (0.0357)
R-Squared	0.228	0.230	0.230	0.228	0.230	0.231
Observations	569,078	569,078	569,078	569,078	569,078	569,078
Year FE	X	X	X	X	X	X
Type FE	X	X	X	X	X	X
Demo. Controls		X	X		X	X
Plan FE	X	X	X	X	X	X

Notes: Table presents OLS regression models, where outcome variable is consumer out-of-pocket costs. The unit of observation is an enrollee-year for the 2007-2009 period. The original data are obtained from a 10% sample of CMS prescription drug event files, aggregated to the enrollee-year level. We include year-level indicators and indicators for the quintile of 2006 spending (consumer types) in all specifications. In some specifications, we also control for 5-yr average per capita Medicare FFS spending and demographic controls for age categories, race, and gender. Standard errors are clustered at the contract level and are presented in parentheses. Statistical significance at the 10%, 5%, and 1% levels are denoted by *, **, and *** respectively.

Table A.9: IV Nested Logit Results

Quintile of 2006 Spending	(1)	(2)	(3)	(4)	(5)
Product Characteristics Approach					
Premium	-0.0326*** (0.00671)	-0.0518*** (0.00694)	-0.0633*** (0.00732)	-0.0626*** (0.00732)	-0.0526*** (0.00634)
Donut Hole Coverage	0.997*** (0.0953)	1.115*** (0.0896)	1.697*** (0.0985)	1.957*** (0.0989)	1.688*** (0.0897)
Has Deductible	-1.036*** (0.0328)	-1.050*** (0.0332)	-1.174*** (0.0356)	-1.213*** (0.0360)	-1.054*** (0.0325)
Plan Vintage	0.394*** (0.0209)	0.453*** (0.0222)	0.443*** (0.0237)	0.469*** (0.0250)	0.490*** (0.0241)
1 - σ	0.318*** (0.0110)	0.343*** (0.0111)	0.374*** (0.0114)	0.384*** (0.0116)	0.376*** (0.0103)
Observations	58,189	58,626	59,885	60,463	61,317
Adjusted R-Squared	0.378	0.356	0.251	0.173	0.242

Notes: Table presents instrumental variable regression models, where the outcome variable is the log of the plan share less the log of the outside share. The outside share is constructed as all Medicare eligibles not enrolled in a stand-alone Medicare Part D plan or MA-PD plan. We include plan fixed effects in all specifications. Instruments are the urban dummy, as well as premiums and out-of-pocket costs in other markets, where a market is defined as a county-year combination. Statistical significance at the 10%, 5%, and 1% levels are denoted by *, **, and *** respectively.

Table A.10: Unobserved Demand Heterogeneity

Quintile of 2006 Spending	(1)	(2)	(3)	(4)	(5)
Baseline Model					
Premium, Type 1	-0.1391 (0.0178)	-0.1275 (0.0053)	-0.0831 (0.0085)	-0.0897 (0.0071)	-0.0940 (0.0303)
OOPC, Type 1	-0.1123 (0.0167)	-0.0649 (0.0236)	-0.0435 (0.0031)	-0.02424 (0.0030)	-0.0591 (0.0061)
Premium, Type 2	-0.2325 (0.0167)	-0.2228 (0.0104)	-0.2548 (0.0107)	-0.2495 (0.0084)	-0.2845 (0.0550)
OOPC, Type 2	-0.1263 (0.0112)	-0.0826 (0.0088)	-0.0486 (0.0044)	-0.0303 (0.0035)	-0.0159 (0.0043)
γ (% of Type 2)	0.6553 (0.0498)	0.6617 (0.0018)	0.6680 (0.0330)	0.6740 (0.0240)	0.6108 (0.1342)
Allowing for Inertia					
Premium, Type 1	-0.1415 (0.0261)	-0.1445 (0.0243)	-0.1316 (0.0222)	-0.1320 (0.0206)	-0.1120 (0.0195)
OOPC, Type 1	-0.0965 (0.0134)	-0.0516 (0.0105)	-0.0298 (0.0076)	-0.0139 (0.0005)	0.0010 (0.0034)
Premium, Type 2	-0.1749 (0.0516)	-0.1717 (0.0519)	-0.1980 (0.0523)	-0.1917 (0.0532)	-0.1930 (0.0554)
OOPC, Type 2	-0.1199 (0.0261)	-0.0774 (0.0243)	-0.0435 (0.0222)	-0.0267 (0.0344)	-0.0007 (0.0039)
Plan Vintage	0.7688 (0.0578)	0.7456 (0.0623)	0.6300 (0.0654)	0.6390 (0.0704)	0.7779 (0.0843)
γ (% of Type 2)	0.6553 (0.2198)	0.6617 (0.2080)	0.6680 (0.2288)	0.6740 (1.7229)	0.6799 (0.2886)

Notes: Table presents results of a model following Berry and Jia (2010). For numerical stability, markets with less than 50,000 enrollees were dropped from the analysis.

Table A.11: Supply Results

	(1)	(2)	(3)
Panel A, Dep. Var.: Logged Premiums			
Lagged Market Share	0.855*** (0.00453)		0.902*** (0.00465)
Enrollment Growth Rate		-0.00281 (0.00180)	-0.0846*** (0.00187)
N	2,174,111	2,174,111	2,174,111
Panel B, Dep. Var.: Logged Insurer Costs			
Lagged Market Share	0.0103 (0.0142)		0.0131 (0.0143)
Enrollment Growth Rate		-0.0182** (0.00603)	-0.0186** (0.00615)
N	3,022,704	3,022,704	3,022,704

Notes: The unit of observation is the beneficiary year. All specifications include contract, state, and year fixed effects. Statistical significance at the 10%, 5%, and 1% levels are denoted by *, **, and *** respectively.

Table A.12: Supply Results

$\partial c / \partial OOPC$	(1)	(2)	(3)	(4)	(5)	(6)
	Non-LIS Distorted		ICR Price Only		Alt. Rebate	
Constant	-1.1871 (0.0081)	-1.3376 (0.0082)	-0.5762 (0.0046)	-0.8020 (0.0045)	-1.0481 (0.0065)	-1.3427 (0.0378)
MA		0.3814 (0.0411)		0.2559 (0.0173)		0.3341 (0.0397)
Plan-Market-Year Obs.	13,737	13,737	33,538	33,538	33,538	33,538
	(7)	(8)	(9)	(10)	(11)	(12)
	Basic Plans		Unobserved Heterogeneity		Accounting for Multi-Product Firms	
Constant	-0.9761 (0.0154)	-1.3427 (0.0163)	-0.9714 (0.0081)	-1.3427 (0.0079)	-0.9720 (0.0082)	-1.2798 (0.0080)
MA		0.2584 (0.0333)		0.4211 (0.0304)		0.3491 (0.0301)
Plan-Market-Year Obs.	28,975	28,975	33,538	33,538	33,538	33,538

Notes: Parameters are estimated using generalized method of moments as described in Section 4. Standard errors are calculated using a bootstrap that re-samples plans with replacement.

Table A.13: Supply Results, by Class

$\partial c / \partial OOPC$	Asthma	Hypertension	Diabetes	Cholesterol
Constant	-1.1700 (0.0132)	-1.3102 (0.0142)	-1.3370 (0.0133)	-1.3428 (0.0132)
MA	0.4437 (0.0781)	0.3605 (0.0545)	0.5761 (0.0591)	0.5764 (0.0513)
Plan-Market-Year Obs.	33,538	33,538	33,538	33,538

Notes: Parameters are estimated using generalized method of moments as described in Section 4. Standard errors are calculated using a bootstrap that re-samples plans with replacement and presented in parentheses.