NBER WORKING PAPER SERIES

SHARING THE BURDEN OF SUBSIDIZATION: EVIDENCE ON PASS-THROUGH FROM A SUBSIDY REVISION IN MEDICARE PART D

Colleen Carey

Working Paper 28529
http://www.nber.org/papers/w28529

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2021

This research received support from the Robert Wood Johnson Foundation Scholars in Health Policy Research Program. I am indebted to Marika Cabral, Michael Grubb, Paul Jacobs, Tim Layton, Sean Nicholson, Mark Shepard, Amanda Starc, Teresa Waters, and seminar participants at University of Pennsylvania, the NBER Summer Institute (Health Care), the NBER Insurance Group, the Northwestern Health Care Markets conference, the University of Michigan, and Cornell University for helpful comments. Anup Das, Megan Hyland, and Yang Zhang provided excellent research assistance. This research was approved by the Institutional Review Board of the University of Michigan. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

Sharing the Burden of Subsidization: Evidence on Pass-Through from a Subsidy Revision in Medicare Part D

Colleen Carey

NBER Working Paper No. 28529
March 2021
JEL No. H51,I11,L13

ABSTRACT

In many federally-subsidized insurance markets, insurers are subsidized on the basis of enrollee characteristics; in principle, subsidies that are "risk adjusted" in this way compensate insurers for ex ante differences in expected cost. Between 2010 and 2011, the subsidies in Medicare Part D were revised, sharply changing the subsidy for diagnoses and demographic characteristics. This paper uses the response of insurers to the subsidy update to estimate pass-through of government subsidies to two insurer choice variables: premiums and out-of-pocket costs. We find that diagnostic subsidies are passed-through at a rate of 40% to the out-of-pocket costs for relevant drugs. Premiums are not responsive to overall subsidies, but do reflect changes in the demographic component of subsidies.

Colleen Carey
Department of Policy Analysis and Management
Cornell University
MVR Hall, Room 298
Ithaca, NY 14853
and NBER
colleenmariecarey@gmail.com

1 Introduction

The degree to which taxes or subsidies are reflected in consumer prices – i.e., their rate of "pass-through" – is now understood to be a key statistic for assessing welfare (Weyl and Fabinger, 2013). Pass-through is of particular interest in public health insurance programs, where subsidies are distributed to private health insurers, as in Medicare Advantage, the Affordable Care Act Exchanges, or the setting of this study, Medicare Part D. There is often a popular perception that pass-through rates are low in such programs, constituting a "giveaway" to insurers or health care providers, especially in the presence of market power (Lenzner, 2013; Newhouse et al., 2007; Frank and McGuire, 2017). Complicating matters further, insurers could conceivably pass-through subsidies to any of a large number of strategic variables. Some strategic variables like premium affect all consumers, but others like out-of-pocket (OOP) cost for particular drugs are only of interest to the subset of individuals who take the drug. Even for an overall level of pass-through, whether pass-through is concentrated on the narrow or broad strategic variables has implications for the equity of the provision of this public service.

In this paper, we measure pass-through of government subsidies in Medicare Part D by exploiting a revision to the subsidy system that sharply changed subsidies for different diagnoses and demographic categories. We develop a theoretical framework to assess how subsidized insurers set both premiums and OOP costs. Within the model we assess how insurers would react to changes in subsidies that affect only a subset of individuals (similar to diagnosis-specific subsidies) as well as a change in subsidies for all enrollees. We empirically measure the pass-through of subsidy changes to Part D enrollees by comparing the change in plan premiums and typical OOP costs to the change in subsidies. Consistent with our theoretical model, we find that higher diagnostic subsidies result in lower OOP costs for the related diagnosis, at a pass-through rate of about 40 percent. Plan premiums do not respond to changes in diagnostic subsidies, but fall substantially in response to higher demographic subsidies.

A subsidy revision in Medicare Part D represents an ideal setting for studying the pass-through of government subsidies to consumers. Medicare Part D is a publicly-funded private prescription drug insurance benefit for twenty million elderly and disabled. The majority of Federal subsidies to insurers in Part D are "risk adjusted", meaning they aim to pay insurers the expected cost of treating an enrollee given her diagnosed conditions and demographics. For example, an average-premium plan in 2010 enrolling a 66 year old man whose medical claims reflect Multiple Sclerosis would receive a subsidy of $659. If a similar enrollee's medical claims instead reflect HIV/AIDS, the plan would receive $2217. In theory, plans are equally willing to enroll both men because the subsidy offsets the higher expected cost of the HIV/AIDS patient. The levels of the diagnosis-specific subsidies were calibrated using data from the early 2000s and then left in place through

2010, despite new drug entry and the onset of generic competition raising or lowering the costs of treating certain diagnoses (Carey, 2017). In 2011, the subsidy system was updated to again set subsidies equal to associated treatment costs. For the two men discussed previously, their insurer in 2011 receives $889 for the Multiple Sclerosis patient (a 35% increase) and $2081 for the HIV/AIDS patient (a 6% decrease). There were also recalibrations of the demographic-specific portion of subsidies. If this man lives in the community, his plan gets $95 more in demographic subsidies (a 28% increase). But if the man lives in a nursing home, the demographic portion of subsidies nearly triples from $515 to $1422. We demonstrate in Section 4 that the subsidies for many diagnoses and demographic categories were sizably raised or lowered as a result of the subsidy system revision.

We develop a theoretical model that predicts how Part D insurers should set premiums and OOP costs as a function of subsidies in Section 3. We start from the results of Weyl and Fabinger (2013), showing how pass-through of subsidies depends on market and demand characteristics. However, we alter the baseline model to consider the structure of Part D subsidies and the multivariate strategies of insurers choosing both the plan premium and a vector of OOP costs. We first consider a subsidy for a subgroup (e.g., a diagnosis), where the subgroup has differential preferences over one element of the strategy space (e.g., OOP costs for drugs that treat the diagnosis). We show that insurers will generally pass-through these "'narrow" subsidies to the related strategic variable, rather than to other strategic variables. We next consider a broad increase in all subsidies. We derive the conditions under which this "broad" subsidy will be passed through the "broad" strategic variable, premium. The model predicts that a diagnosis-specific subsidy in Part D is likely to be passed-through to related OOP costs, but has ambiguous predictions for the pass-through of the demographic portion of subsidies, which have features of both "broad" and "narrow" subsidies.

In the empirical portion of the paper, we set out an empirical model that leverages the large 2011 revision to the subsidy system to recover the rate of pass-through. We quantify the effects of the revision on demographic and diagnostic subsidies. We measure pass-through to premiums and OOP costs using a panel data model with fixed effects. To measure pass-through to premium, we consider how plan premiums responded to changes in plans' average per-enrollee subsidy (freezing a plan's enrollment in 2010 to avoid compositional changes). Plan fixed effects net out all plan-specific factors, identifying the effect of subsidies purely from the premium changes that co-occur with the 2011 revision. To measure pass-through to OOP costs, we examine how the typical OOP costs for drugs treating a diagnosis respond to the diagnosis's subsidy. Again, we use fixed effects to isolate the variation induced by the 2011 subsidy revisions, netting out all time-invariant plan $\times$ diagnosis factors or plan $\times$ year factors that are uniform across diagnoses. In additional specifications, we consider extra controls that accommodate the economic content of subsidy updates. Since a diagnosis's subsidy may be revised upward to accommodate an ongoing upward trend in

drug prices, we include a diagnosis price spline or a diagnosis-specific time trend.

We find that plan premiums did not respond to changes in the total average subsidy, with the point estimate suggesting $1 in increased subsidy is associated with a $0.01 decrease in plan premiums. However, when the diagnostic and demographic components of premiums are considered separately, we find that the pass-through rate for demographic subsidies is considerable: $1 in increased demographic subsidy is associated with a $0.74 reduction in plan premiums. An event study version of this analysis supports a causal interpretation; premiums of plans that went on to get subsidy increases as a result of the revision were trending similarly to plans that went on to get subsidy decreases. We estimate that plans respond to increases in diagnostic-subsidies by lowering the associated OOP costs, leading to a pass-through rate to OOP costs of 37-47%. Since demographic subsidies were approximately 40% of total subsidies in 2010, we compute an overall pass-through rate of about 53% ($= .4 * .74 + .6 * .4$).

We next turn to determining exactly how insurers adjust their benefit designs to pass through the diagnostic subsidies. We extend our panel-data model to predict the OOP cost, coverage, formulary placement of the universe of Part D drugs as a function of their diagnostic subsidy. We conduct this exercise separately for branded and generic drugs. We find that Part D insurers generally pass-through subsidies by beginning to cover branded drugs, and moving such drugs to more favorable formulary tiers. We find that there is little response for any of these outcomes among generic drugs. We discuss several potential explanations for this finding, including differences in the baseline benefit design for brands and generics, differences in the upstream market (monopolistic vs. competitive), and differences in demand. We also examine the response of utilization management tools such as prior authorization or requirements that beneficiaries try cheap alternatives before expensive drugs (step therapy). We find that, much like financial management tools like OOP costs, the use of "step therapy" falls when subsidies increase. However, the use of prior authorization increases. Because prior authorization creates a point of contact between the insurer and the provider, insurers may use it to ensure that providers are documenting the relevant diagnosis to ensure proper subsidy payment.

There are two features of Part D that are not reflected in our baseline pass-through model, but which could affect our interpretation of estimates. If a subsidy related to a particular characteristic increases, and insurers respond by reducing the related strategic variable, individuals with that characteristic may enroll in Part D for the first time. As explored in Cabral et al. (2018), if the new enrollees are healthier than the current enrollees, our pass-through estimates would tend to be biased upward. However, in Section 7.1 we rule out any economically meaningful enrollment response. Our pass-through estimate is also biased if drug demand is very elastic – i.e., if a reduction in OOP costs induces a large increase in drug utilization.[1]

[1]Einav et al. (2018) document a cross-sectional relationship between drug demand elasticity and cost-sharing, suggesting

elasticity on the order of 2%. In a back-of-the-envelope calculation, we show that accommodating this small drug-demand elasticity would only slightly alter our pass-through estimates.

This paper contributes to a set of recent reduced-form estimates of pass-through rates, both in health insurance (Cabral et al., 2018; Duggan et al., 2016) and beyond (Fabra and Reguant, 2014; Miller et al., 2017; Muehlegger and Sweeney, 2017; MacKay and Remer, 2018). Following the empirical models in this literature, pass-through is identified over time net of a product fixed effect, accommodating independent unobserved costs that could otherwise introduce bias in estimates (MacKay et al., 2014). It is novel in taking advantage of risk adjustment subsidy revision, a periodic feature of risk-adjusted health insurance markets, to obtain this key market descriptor. It improves on the cross-sectional analyses of Carey (2017) and Geruso et al. (2018), who showed that insurers provide more generous benefits for diagnoses made profitable by risk adjustment systems in Medicare Part D and the ACA Marketplaces, respectively. This paper isolates more credible over-time variation in subsidy levels, and adjusts the outcome variables to obtain a pass-through rate.

This paper's estimate of 53% for overall pass-through in Medicare Part D is very similar to that of Cabral et al. (2018), who use a revision to the capitation rate in Medicare Advantage to recover a pass-through rate of 54%. On the one hand, the two markets have the same consumers and are federally subsidized in a similar manner. However, Medicare Advantage is much less competitive than Medicare Part D, with an average HHI in MA markets of 0.6 as compared to 0.1 in Part D. Cabral et al. (2018) find that market power is the likely explanation for incomplete pass-through in Medicare Advantage (with estimated pass-through of 74% in the most competitive markets)[2], but that is unlikely to explain incomplete pass-through in Part D. One candidate explanation for incomplete pass-through is that Medicare Part D insurers face a much more concentrated upstream market than Medicare Advantage insurers. Of the $\sim$ 45 cents per dollar that Medicare Advantage and Part D insurers do not pass-through to consumers, monopoly drug firms may be better positioned than diffuse health care service providers to obtain a portion in the bargaining process.[3]

This paper is distinct in considering the breadth of choice variables through which Part D insurers can respond to subsidies, as well as distinguishing between the effects of narrowly-targeted vs. broad subsidies. In our theoretical model, diagnosis-specific subsidies are likely to be passed-through to OOP costs rather than premiums; insurers target the reductions on the variable most likely to pull in the individuals that

that insurers set higher cost-sharing for drugs with the most elastic demand.

[2] Following Cabral et al. (2018), we split Part D markets by Herfindahl-Hirschman index in Appendix Section A.5. However, while MA markets vary greatly in concentration with a sample standard deviation of 0.25, nearly all Part D markets are unconcentrated, with a 90-10 split of 0.16 to 0.09. Thus, we are underpowered to test for pass-through heterogeneity by concentration, and unsurprisingly find no differences.

[3] As we discuss in Section 6.2, pass-through is larger (more negative) for brands supplied monopolistically vs. generics supplied competitively. However, the differences in the baseline formulary placement of brands vs. generics, as well as potential differences in the demand functions, make this comparison uninformative for determining the effect of upstream market concentration.

qualify for the higher subsidy. We confirm this finding in our empirical model; plans adjust related OOP costs, but not premiums, in response to larger diagnosis-specific subsidies. However, we do find a large premium response to demographic-specific subsidies. Demand for particular drugs is unlikely to correlate with demographic characteristics, which in our model makes it more likely that insurers will pass-through demographic subsidies to premiums. Cabral et al. (2018) identify pass-through using a broad increase in subsidies. It is thus unsurprising that of their \$0.54 cents in pass-through of an incremental dollar in Medicare Advantage, \$0.45 comes as a reduction in premiums and only \$0.09 comes as an improvement in benefits.

This study also helps clarify the relationship between pass-through in insurance products and service-level selection in insurance. Insurers commonly have both broad and targeted components of their strategy space. If they have incentives to effectuate service-level selection, they will choose to pass-through to the targeted components, such as OOP costs for specific diagnoses. If instead they do not have these incentives, they will instead choose to pass-through subsidy increases broadly, either to premiums or to uniformly improving benefits. As we know from the work of Carey (2017), Geruso et al. (2018), and Brown et al. (2014), insurers are highly incentivized by diagnosis-based risk adjustment to effectuate service-level selection.

2 Medicare Part D

This section details the design of the Medicare Part D market, with special attention to insurer incentives and the diagnosis-specific subsidy system. We first describe how enrollees choose Part D plans and drugs. We then describe the insurers' plan benefit design problem and the regulations that constrain their actions. Finally, we review how Part D plans were paid in their first five years and the nature of the revision in 2011.

2.1 Enrollment and Drug Demand

Medicare Part D implements the managed competition model of public health insurance that underlies Medicare Advantage, Medicaid managed care, and the Affordable Care Act marketplaces. In the managed competition model, individuals choose among competing insurers offering a regulated benefit. Approximately half of Medicare beneficiaries are in the market for stand-alone Medicare Part D (i.e., no prescription drug coverage through a retiree benefit and not enrolled in a combined medical-drug Medicare Advantage plan). In 2010, they chose among an average of 45 insurance plans operating in their market. Markets are defined administratively by CMS as either a state or a group of states. Plans must accept everyone who enrolls at a uniform premium. Plans differentiate themselves both vertically (overall level of benefit generosity) and horizontally (level of coverage for competing drugs within a therapeutic class), subject to the regulations described in Section 2.2.

Because Medicare beneficiaries have very persistent drug utilization, choice of insurance plan commonly

incorporates enrollees' private information on predicted drug demand. There are several pieces of evidence for enrollees' private information. Firstly, prior to the onset of Medicare Part D in 2006, no free-standing prescription drug insurance existed for this population; Pauly and Zeng (2004) and Goldman et al. (2006) suggest the threat of adverse selection inhibited the development of such a market. Secondly, beneficiaries who remain uninsured despite eligibility for Part D appear to be positively selected (Yin et al., 2008; Levy and Weir, 2010); however, the presence of substantial government funding, covering 75% of Part D expenditure on average, means that most eligible beneficiaries enrolled. Thirdly, direct evidence on prescription drug utilization reflects substantial year-over-year persistence in drug needs (Hsu et al., 2009). An analysis by Heiss et al. (2013) finds that basing ones' choices entirely on last year's drug needs is the choice rule that minimizes *ex post* expenditures among a set of heuristics and rational expectations models. Finally, direct evidence gathered by Polyakova (2016) documents a substantial degree of asymmetric information, resulting in adverse selection into the most generous Part D plans.

The presence of private information in plan choice affects insurers' incentives because it means that individuals are aware of their diagnoses and drug needs at the time of enrollment, and are responsive to related benefits. In the next section, we explain insurers' strategic choices in Part D as well as applicable benefit design regulation.

2.2 Insurers and Drug Firms

Insurers recognize that Part D enrollees can forecast their drug needs. Since they must accept all applicants at a uniform preannounced premium, they cannot directly select enrollees. Instead, they must use their benefit designs –what drugs are covered and at what OOP costs– to attract profitable enrollees and deter those who will spend more than the subsidies the insurer receives for them.

Federal regulation constrains both choice of coverage and choice of OOP costs in hopes of providing access to an equitable benefit for all enrollees. For coverage, insurers must cover two drugs in each United States Pharmacopeia therapeutic class and all drugs in six "protected" classes (drugs for serious chronic illness). This regulation still allows considerable variation in coverage across plans, and plans are allowed to vary coverage to favor diagnoses with high diagnosis-specific subsidies. The plans we study in this analysis vary from covering 47 to 97 percent of drugs.

Out-of-pocket costs are also subject to regulation. Out-of-pocket (OOP) costs are defined in relation to the Part D "Basic Benefit", which is the coverage level funded by Federal subsidies. In the Basic Benefit, individuals' OOP costs depend on the zone of coverage, determined by their expenditure so far in the year: individuals pay a deductible, then 25% of drug expenditures in an "initial coverage zone", then 100%[4] of

[4]The Affordable Care Act began reducing this percentage in 2012 as discussed below.

drug expenditures in the doughnut hole, and finally 5% of drug expenditures after a catastrophic threshold. Plans can satisfy OOP cost regulation by either setting OOP costs to the Basic Benefit coinsurances or raising certain OOP costs and lowering others such that OOP costs still attain the Basic Benefit percentages on average. 90% of plans choose the latter strategy, which gives them considerable latitude to set lower OOP costs for certain diagnoses over others. Plans may also choose to offer "enhanced coverage", financed fully out of premiums, that reduces OOP costs below the Basic Benefit percentages in some zones of coverage.

Enrollees also pay a premium to their chosen plan. Premiums are computed from a *bid* that represents for each plan their expenditure on a "typical" enrollee. The premium is then set to $prem_i = (bid_i - \overline{bid}) + \gamma\overline{bid}$. In this equation, $\overline{bid}$ is the national average bid (weighted by last year's enrollment) and γ is a fixed percentage (36% in 2010). Plans that cover many drugs at low OOP costs spend more for a "typical" beneficiary and therefore have a higher bid; their premiums are higher by the full amount that their bid exceeds the national average bid.

Plans (or a contracted pharmacy benefit manager acting as their agent) negotiate with upstream drug firms over formulary placement and drug price. Plans obtain a discount or rebate when the plan sets a low OOP cost for the drug relative to competitors in the therapeutic class. The discount or rebate could be a percentage off the list price, or a quantity discount. Discounts and rebates are a closely-held trade secret and are not observable in the data. According to the consultant Milliman, the Part D setting is institutionally distinct, such that an insurer and drug firm may set distinct prices and rebate terms for Part D as compared to the employer-sponsored market (Dieguez et al., 2018).

Because plans set coverage and OOP costs for approximately 5000 drugs, they have a relatively fine-grained tool for attracting or deterring potential enrollees who prefer certain drugs (Geruso et al., 2018). In the next section, we explore the diagnosis-specific subsidies meant to make insurers indifferent between all enrollees.

2.3 Diagnosis-Specific Subsidies

Diagnosis-specific subsidies, as well as government subsidies in general, play a critical role in Part D market design. In the absence of any subsidization, many individuals who know their (persistent) drug needs are inexpensive would not wish to pool with those with high expected expenditures. The high degree of government subsidies to the Part D market induces the healthy to voluntarily enroll, facilitating a balanced risk pool and providing financial protection for unexpected drug needs.

To see why subsidies are based on diagnoses and demographics, suppose Medicare had simply paid each Part D plan the average expenditure for each individual: approximately $1200. Within the benefit design regulations above, insurers would have designed benefits to disproportionately attract healthy beneficiaries

and deter the sick. Instead, Medicare conditions its subsidies on diagnoses and demographics: subsidies to plans are higher for enrollees with high-cost diagnoses or in high-cost demographic categories, and lower for those who are relatively healthy. Subsidies that vary with individuals' expected health status are known as "risk adjustment". A recent literature has pointed out the weaknesses of basing subsidies exclusively on diagnostic and demographic factors. These factors may not directly predict demand for insurance (Layton, 2017); alternatively, such subsidy systems may incompletely adjust for predictors of economic choices such as service elasticity (Einav et al., 2016) or inertia (Bijlsma et al., 2014). Still, diagnosis-based subsidy systems can be easily computed by regulators and can significantly reduce the scope for selection (Newhouse et al., 2013). Recent work such as Layton et al. (2018) shows how the risk adjustment calibration process can be further improved to neutralize selection incentives.

A subsidy system such as Part D's contains three distinct elements: diagnostic definitions, weights representing the relative cost of each diagnosis, and a conversion from weights to subsidies. The first diagnosis-specific subsidy system was calibrated prior to Part D's beginning in 2006 and is detailed in Robst et al. (2007). The diagnostic definitions, built up from ICD-9 codes, were borrowed from the subsidy system used in Medicare Advantage; in addition to diagnoses, individuals were grouped by demographics: age, sex, and originally entitled to Medicare due to disability. The subsidy system designers obtained a sample of prescription drug and medical claims from Federal retirees (incurred in 2000) and disabled Medicaid beneficiaries (incurred in 2002). They applied the Part D Basic Benefit to each individual's claims to simulate the expenditure of a Part D plan for these individuals.

To set relative cost weights for diagnoses and demographics, they ran the following regression:

$$\mathcal{E}_i/\overline{\mathcal{E}} = \sum_x \omega_x \delta_{ix} + \sum_g \omega_g \delta_{ig} + \varepsilon_i \tag{1}$$

In this expression, $\mathcal{E}_i/\overline{\mathcal{E}}$ is the simulated Part D expenditure for this Federal retiree or disabled Medicaid beneficiary, normalized by the sample mean expenditure. δ_{ix} and δ_{ig} are 0/1 flags for the 84 diagnoses[5] or demographic categories, and the coefficients ω_x and ω_g are the relative weights for each. A fixed factor increases the weight for low-income or long-term institutionalized individuals, since such individuals generally have more severe forms of diagnoses. An individual with a weight of one is expected to spend the sample average $\overline{\mathcal{E}}$.

The subsidy a plan receives for an individual is the product of the plan's bid and the sum of the individual's demographic and diagnostic weights. Scaling weights by a plan's bid allows subsidies to increase with the overall generosity of a plan's benefit design.

[5]Robst et al. (2007) refer to 87 diagnoses; we disregard two related to Cystic Fibrosis because of extreme rarity, and we treat as a single diagnosis two that were constrained in Equation 1 to have the same coefficient.

To see how the original subsidy system works, suppose an insurance plan enrolls a 66-year-old man (never disabled, not low-income, not institutionalized). His medical claims from the previous year reflect an Infectious Disease. The total weight for this man is the ω_x for Infectious Disease, 0.073, and the ω_g for his demographic category, 0.355. A plan that bids the national average for 2010 ($1060) would receive $454 for this man. A more generous plan bidding $1500 would receive $642.

As explored in Carey (2017), technological change in the form of the entry of new molecules and the onset of generic competition (among other forces) caused actual treatment costs in Part D to drift from the subsidy weights set in the initial calibration. Therefore, Medicare revised the subsidy system for 2011.

2.4 The Subsidy System Revision

The subsidy system revision, detailed in Kautter et al. (2012), altered the diagnostic definitions and recalibrated the weight associated with each diagnosis (the conversion of weights to subsidies remained the same). Firstly, diagnostic definitions were altered by reorganizing the ICD-9 codes. For example, the diagnoses *Quadriplegia* and *Motor Neuron Disease & Spinal Muscular Atrophy* in the old subsidy system are collapsed into one diagnosis – *Spinal Cord Disorders* – in the new system. *Chronic Renal Failure*, on the other hand, is expanded from one diagnosis to four subtypes. Various forms of cancer are completely reorganized.

In addition, each diagnosis and demographic category now comes in five subtypes for disabled $\times$ low-income status and long-term institutionalized. This is because those factors can dramatically change the expenditure associated with a given diagnosis. A subsidy weight for each subtype was intended to better align a diagnosis's subsidy and a plan's expenditures for that diagnosis.

Finally, Equation 1 was reestimated using the claims and diagnoses of individuals enrolled in free-standing Part D in 2008.[6] The introduction described the change in subsidies for two diagnoses – *HIV/AIDS* and *Multiple Sclerosis* – which were defined by the same ICD-9 codes in both the new and old systems. The subsidy update for those two diagnoses suggests that many diagnoses received much larger or smaller subsidies in 2011 relative to 2010. Later, we develop evidence that this is indeed the case.

We have seen that, firstly, beneficiaries' plan choices are characterized by private information on their drug needs; secondly, insurers can attract individuals by generous benefit design for drugs that treat their diagnoses; and, finally, the subsidy system and its revision provide variation over time in the subsidy a plan receives for each diagnosis and demographic category. In the next section, we explore this interaction in a simple theoretical model.

[6]The 2010 Affordable Care Act gradually closed the doughnut hole in the Basic Benefit through a combination of increased insurer liabilities and drug firm discounts. To compensate insurers for the increased liability, the risk adjustment was again recalibrated for 2012. This revision, detailed in the Appendix, resulted in very minor changes to subsidies.

3 Conceptual Model

While original results date to Marshall (1890), the pass-through of taxes or subsidies to consumer prices (premium and OOP costs in our framework) has recently been revived in the theoretical literature (Weyl and Fabinger, 2013; Fabinger and Weyl, 2015). In this section, we apply the recent results to show how insurers in Part D would respond to changes in the subsidies. We adapt the canonical model to determine how insurers might differentially pass-through to premiums and OOP costs. In addition, we discuss how adverse selection into Part D or drug demand elasticity might affect our interpretation of estimated pass-through rates.

3.1 Subsidy Pass-Through

As a baseline model, we greatly simplify the Part D market. Suppose Medicare beneficiaries can be represented as a set of homogeneous individuals with the same level of illness, each consuming one unit of prescription drugs. In this case, we can ignore individual-level heterogeneity and represent an enrollee's expenditure in plan p by by m_p, where m_p is the sum of premium and OOP costs for one unit of drugs in plan p. Let $D(m_p)$ represent enrollment in plan p when it sets the sum of premium and OOP costs to m_p, with $D' < 0$ and $D'' > 0$. For each individual enrolled in insurance, the insurer receives subsidy r for their enrollment, purchases the drugs at κ. Insurer profits are thus $D(m_p)[r - \kappa + m_p]$.

Assume insurers are symmetric: $m_p = m$. Weyl and Fabinger (2013) show that an insurer's optimal choice for m for a large range of demand systems can be captured by the following first order condition:

$$\text{FOC: } r - \kappa + m = -\theta \frac{D(m)}{D'(m)} = \theta\mu$$

In this expression $\mu = -\frac{D(m)}{D'(m)}$ and $\theta \in [0, 1]$ measures the level of insurer competition, with 0 representing perfect competition and 1 representing a monopoly insurer.[7] Implicitly differentiating the first order condition we obtain a simple relationship for the response of m to a change in r

$$\frac{\partial m}{\partial r} = -\frac{1}{1 - \theta\mu'} = -\rho$$

The pass-through rate ρ[8] depends crucially on $\mu' = -\frac{D'^2 - DD''}{D'^2}$. The sign of μ' determines whether, in imperfect competition, the rate of pass-through is above or below 1. Mathematically, μ' takes the sign of the second derivative of the log of demand. To build intuition, consider a monopoly insurer ($\theta = 1$). If the log

[7] Weyl and Fabinger (2013) allow θ to vary with m, but for simplicity we will not allow this dependence. They show that several imperfectly competitive markets, such as homogeneous products oligopoly, are characterized by a θ that does not vary with m.

[8] Following the literature, we define ρ as a positive parameter, such that the effect of a subsidy increase on m is $-\rho$.

of D is linear, $\mu' = 0$, and pass-through to OOP costs would be 1, or perfect pass-through.[9] An increase in subsidy would cause the insurer to reduce m one-for-one, leading to the same markup $(r - \kappa + m)$ but a larger share. The demand elasticity is lower at this new lower m. The economic implication of log-concavity is that the same unit decrease in OOP costs would lead to a region of even less elasticity than in the log-linear case. In response to the inelasticity, the insurer chooses a higher m than is chosen in the log-linear case, leading to pass-through between zero and one.

One of the immediate consequences of Weyl and Fabinger's equation is that pass-through approaches one as markets become more competitive, meaning as θ approaches 0. Furthermore, if pass-through is below one, it will approach one from below. In our setting, this prediction means that pass-through should be larger (more negative) in Part D markets that are more competitive. We test this prediction in Appendix Section A.5.

3.2 Subsidy Pass-Through to Premiums vs. Out-of-Pocket Costs

In the previous section, we simplified an individual's spending in Part D to m, the sum of both premiums and OOP costs. In practice, plans set a multidimensional vector comprised of the plan premium and the OOP costs c_x for each of X diagnoses. All enrollees pay the premium, while enrollees only consider the OOP costs for the diagnoses they have. We consider two kinds of subsidies: *broad*, meaning they affect all enrollees, or *targeted*, meaning they affect only a subgroup.

In order to consider how pass-through differs across these dimensions, we will consider a setting where there are two diagnoses; individuals have only one diagnosis of the two and it is known to them at the time of plan selection. Because diagnoses are known, demand for enrollment among those with diagnosis 1 depends on the sum $\text{prem} + c_1$ while demand for those with diagnosis 2 depends on the sum $\text{prem} + c_2$. Using the sum of these variables is necessary because models with multiple non-additive strategic variables generally cannot be solved analytically.[10] However, because the key parameters always appear added together, profit maximization only identifies the *sum* of premium and OOP costs that is optimal; the optimal division between the two parameters is not given by the model. Still, we can make progress by assuming a plan is setting optimal premium and OOP costs at baseline and then reacts to an increase in a subsidy.

Assume that the demand functions for diagnoses 1 and 2 are identical, as are the initial subsidy r and the price of drugs to treat the diagnoses κ. Plans are constrained to a single premium prem_p for all enrollees but

[9] An example of log-linear demand is $D(m) = exp(am)$ for $a < 0$. Then $ln(D(m)) = am$, with the first derivative of $ln(D(m))$ equal to a and a second derivative of zero. The elasticity of demand is equal to $-am$, which is lower at lower m.

[10] Rothman (2015) studies pass-through when firms are setting a two-part tariff, which is analogous to Part D plans choosing an annual premium for enrollment as well as OOP costs that are incurred upon filling prescriptions. In contrast to the model we consider here, the two elements of a two-part tariff cannot be combined additively as arguments for demand or profits. Rothman shows that there are no clear predictions for pass-through in this setting; it depends on the covariance of demand for the two goods. An increase in the cost of the first good could lead the firm to actually *increase* the second's price (rather than decrease it).

can set different OOP costs c_{1p} and c_{2p}. However, given identical demand, subsidy, and drug prices, plans initially set the same OOP costs c_p. Thus, we denote the sum of premiums and OOP costs as m_p for both diagnoses, and furthermore let ϕ represent the per-person profits at the initial subsidy level: $\phi = r - \kappa + m_p$.

$$\Pi^{\text{pre}} = 2D(\text{prem}_p + c_p)[r - \kappa + \text{prem}_p + c_p] = 2D(m_p)[r - \kappa + m_p] = 2D(m_p)\phi$$

3.2.1 Pass-Through of a Targeted Subsidy

Suppose then that there is an exogenous increase of ε in the subsidy for diagnosis 1. Insurers could consider lowering either premium, OOP costs for that diagnosis, or OOP costs for the other diagnosis. Using the model developed in the previous section, we know that plans will lower the *sum* of premium and OOP costs by $\rho\varepsilon$. We will show that pass-through to c_{1p} creates higher profits than pass-through to c_{2p} or prem_p.

We first show that if the subsidy for diagnosis 1 increases, the plan prefers reducing c_{1p} to reducing c_{2p}. Intuitively, the revenue lost by reducing c_{1p} is offset by the increased subsidy for those with diagnosis 1; there is no such offset for those with diagnosis 2.

$$\Pi^{\text{post,pass-through to } c_{1p}} > \Pi^{\text{post,pass-through to } c_{2p}}$$

$$\overbrace{D(m_p - \rho\varepsilon)[\phi + \varepsilon - \rho\varepsilon]}^{\Pi \text{ for diag 1}} + \overbrace{D(m_p)\phi}^{\Pi \text{ for diag 2}} > \overbrace{D(m_p)[\phi + \varepsilon]}^{\Pi \text{ for diag 1}} + \overbrace{D(m_p - \rho\varepsilon)[\phi - \rho\varepsilon]}^{\Pi \text{ for diag 2}}$$

$$D(m_p - \rho\varepsilon)[\phi + \varepsilon - \rho\varepsilon - \phi + \rho\varepsilon] + D(m_p)[\phi - \phi - \varepsilon] > 0$$

$$D(m_p - \rho\varepsilon) - D(m_p) > 0$$

The inequality holds as long as $D' < 0$.

Next, we derive the conditions under which the plan prefers reducing c_{1p} to reducing prem_p after an increase in the subsidy for diagnosis 1.

$$\Pi^{\text{post,pass-through to } c_{1p}} > \Pi^{\text{post, pass-through to prem}}$$

$$\overbrace{D(m_p - \rho\varepsilon)[\phi + \varepsilon - \rho\varepsilon]}^{\Pi \text{ for diag 1}} + \overbrace{D(m_p)\phi}^{\Pi \text{ for diag 2}} > \overbrace{D(m_p - .5\rho\varepsilon)[\phi + \varepsilon - .5\rho\varepsilon]}^{\Pi \text{ for diag 1}} + \overbrace{D(m_p - .5\rho\varepsilon)[\phi - .5\rho\varepsilon]}^{\Pi \text{ for diag 2}}$$

On the RHS of the inequality, the premium decrease is half the size of the decrease for c_{1p}, which leads to the insurer reducing weighted prices by the same amount in the two situations.[11] Rearranging the inequality

$$\overbrace{[D(m_p - \rho\varepsilon) + D(m_p) - 2D(m_p - .5\rho\varepsilon)]}^{+ \text{ by convexity of } D}\overbrace{\phi}^{+} + \overbrace{[D(m_p - \rho\varepsilon) - D(m_p - .5\rho\varepsilon)]}^{+}[\varepsilon - \rho\varepsilon] > 0$$

This inequality will always hold when pass-through is less than 1 (as we find empirically), and can still hold

[11] See Appendix Section A.1 for a discussion of this assumption, as well as an exploration of alternatives.

even with pass-through rates slightly above 1. This finding suggests that targeted subsidies will be applied to the strategic variable most important in the demand function of the affected population. For diagnostic-specific subsidies, this suggests that pass-through will be to the OOP costs rather than the premium.

3.2.2 Pass-Through of a Broad Subsidy

Another possibility is that a broad subsidy increase raises r for both diagnoses. Since both enrollees and plans care only about the sum of OOP cost and premiums, insurers are indifferent between reducing OOP cost or premiums. However, it is uncertain whether the plan would prefer to pass-through the subsidy increase equally between the two diagnoses or focus the reductions on a single diagnosis.

$$\Pi^{\text{post,pass-through to both}} > \Pi^{\text{post, pass-through to } c_{1p}}$$

$$\overbrace{2D(m_p-\rho\varepsilon)[\phi+\varepsilon-\rho\varepsilon]}^{\Pi \text{ if pass-through to both}} > \overbrace{D(m_p-2\rho\varepsilon)[\phi+\varepsilon-2\rho\varepsilon]}^{\Pi \text{ for diag 1 if pass-through to } c_{1p}} + \overbrace{D(m_p)[\phi+\varepsilon]}^{\Pi \text{ for diag 2 if pass-through to } c_{1p}}$$

$$2D(m_p-\rho\varepsilon)[\phi+\varepsilon]-2D(m_p-\rho\varepsilon)[\rho\varepsilon]-D(m_p-2\rho\varepsilon)[\phi+\varepsilon]+2D(m_p-2\rho\varepsilon)[\rho\varepsilon]-D(m_p)[\phi+\varepsilon]>0$$

$$\underbrace{[2D(m_p-\rho\varepsilon)-D(m_p-2\rho\varepsilon)-D(m_p)]}_{-\text{ by convexity of } D}\underbrace{[\phi+\varepsilon]}_{+}+\underbrace{2[D(m_p-2\rho\varepsilon)-D(m_p-\rho\varepsilon)]}_{+}\underbrace{[\rho\varepsilon]}_{+}>0$$

The inequality only holds if demand is not very convex. To understand the economic intuition, if a plan passes the subsidy through to both consumer types, demand for those with diagnosis 1 increases from $D(m_p)$ to $D(m_p-\rho\varepsilon)$. If the plan targets pass-through only on the type 1 consumers, there is an incremental demand increase from $D(m_p-\rho\varepsilon)$ to $D(m_p-2\rho\varepsilon)$, which is larger than the initial increase due to convexity. If demand is very convex, the plan gets a large demand response from those with diagnosis 1 and decides to concentrate all subsidy increases on them. Note that demand that is very convex is not log-concave.

3.3 Predictions for Pass-Through in Part D

In the previous sections, we generated theoretical predictions for an insurer setting premiums and diagnosis-specific OOP costs who faces an increase in subsidy. When the subsidy is tied to a particular diagnosis, the insurer will generally choose to reduce the related OOP costs. The insurer's optimal strategy when subsidies increase broadly for all enrollees is less clear, and dependent on parameters. In this section, we apply these predictions to the Part D setting.

Our empirical analogue to the "targeted" subsidy increase considered in the model is a diagnostic increase. However, there is no exact analogue over our sample period to a "broad" subsidy increase. The demographic component of Part D subsidies is closer to a "broad" subsidy, in that it affects enrollees irrespective of diagnoses. In particular, the characteristics that define demographic groups do not strongly predict increased demand for particular drugs, and thus an insurer cannot differentially appeal to particular demographic

groups by selective improvement to OOP costs.

An important assumption in the model is that demand functions for the two diagnoses are identical. Suppose that the subsidy for diagnosis 1 increases, but demand for diagnosis 1 is perfectly inelastic. The insurer would direct pass-through to diagnosis 2. If the Part D subsidy revisions are concentrated on diagnoses where demand is particularly inelastic, that could lead Part D insurers to direct pass-through to alternative diagnoses where demand is elastic. Carey (2017) shows that subsidy inaccuracies are related to new drug entry and the onset of generic competition, which would not obviously be correlated with demand parameters.

Some studies of Part D demand have suggested that consumers are less responsive to OOP costs than they are to premium (Abaluck and Gruber, 2011). The low salience of OOP costs is a countervailing force that could lead plans to concentrate pass-through on premiums. In Appendix Section A.2 we adapt the models of Agarwal et al. (2014) and Heidhues et al. (2016) to include a parameter that allows OOP costs to be less salient in demand than premium. As salience of OOP costs falls, insurers are more likely to concentrate pass-through on premium rather than OOP costs.

The differing institutions surrounding premium-setting and the determination of OOP costs could also influence pass-through to each. Plans submit a detailed "bid" to CMS; while the bid must be supported by data on the plan's expected cost, the plan has the ability to set a profit margin. Thus, while premiums must be rationalized by cost, a plan can pick its optimal premium within a range. Out-of-pocket costs are somewhat more constrained. Plans using coinsurances are setting OOP costs as a percent of list prices; plans typically pick round numbers (25, 33, 50) as coinsurance rates. And drug list prices are very similar across plans, limiting the possibility of desired OOP cost adjustments to be generated by price adjustment. For plans using tiered copays, the number of tiers is usually four or five, meaning that while a plan can adjust copays by moving drugs between tiers as well as changing the copay associated with each tier, the plan has to choose from a set of discrete copays for each drug. Conlon and Rao (2020) show that the typical pass-through models only apply to a setting with continuous choice variables; when choice variables are discrete, pass-through rates can be above or below the predictions under continuity.

Finally, we note that 13% percent of Part D enrollees have no diagnoses. If these individuals are not responsive to OOP costs for any particular diagnosis, premium reductions are the only way to attract such individuals.

3.4 Interpreting the Pass-Through Rate

Finally, once we obtain a pass-through rate, there are two further considerations in interpreting it: adverse selection into Part D and drug demand elasticities.

If new individuals join Part D in response to the revision and have a different level of illness than current enrollees, then the estimated pass-through rate could be inaccurate. Cabral et al. (2018) explore this possibility in the context of Medicare Advantage. Suppose the subsidy revision raises the subsidy for a given diagnosis from S_{x10} to S_{x11}, and that plans then lower OOP costs for drugs treating that diagnosis. Assume that some uninsured individuals choose to join Medicare Part D due to the improved benefits for their diagnosis, and assume they have milder forms of the diagnosis, as would be suggested by Polyakova (2016). In this case, we would measure a plan's subsidy increase as $S_{x10} - S_{x11}$, but would neglect the increase in profitability for diagnosis x among new enrollees. We would estimate a pass-through rate that is larger in magnitude (more negative) than the truth, because we would be comparing a change in OOP cost to an underestimate of the change in subsidy. In Section 7.1 we consider this potential source of bias.

Another distortion to pass-through rates arises if drug demand is strongly responsive. Above we modeled enrollment into Part D, because enrollment in a plan is the key condition for subsidy receipt. However, in reality there is a second stage in which drug demand as a function of OOP costs is realized. Consider again a subsidy that increases from S_{x10} to S_{x11}, and suppose we find that OOP costs fall from c_{x10} to c_{x11}. If individuals with x increase the quantity of drugs that treat x, this would partially offset the measured increase in subsidy. In this case, our estimated pass-through rate would be overstated. In Section 7.2 we examine the response of drug demand to the subsidy revision.

4 Measuring Subsidy Updates

We now move to testing the predictions generated by our theoretical model. In this section, we describe a substantial subsidy system revision in Medicare Part D. The subsidy update is the independent variable in our estimation of pass-through rates.

4.1 Data

This research combines Medicare claims data with the publicly-available Part D benefit designs. Our Medicare claims dataset provides medical and prescription drug claims for a 5% panel of Part D enrollees between 2008 and 2012. The medical claims enable us to assign diagnoses to individuals in the exact same way as Medicare: if an individual has a medical claim with a specified ICD-9 code in year $t-1$, the subsidy given to their Part D plan in year t will reflect that diagnosis. Diagnoses can only be observed for individuals enrolled in fee-for-service Medicare (not Medicare Advantage) because claims from Medicare Advantage enrollees are not released to researchers for the relevant years.

The benefit designs of all Part D plans are contained in the Prescription Drug Plan Formulary files for the years 2009 through 2012. The Formulary files contain coverage and, if covered, OOP costs for all drugs

and all plans. For all covered drugs a "list" price for the plan is also reported, but the price is before an unobserved rebate. In 2011, rebates amounted to about 15% of the list price of branded drugs in Part D (Office of the Inspector General, 2018).

4.2 Measurement of Change in Individual-Level Subsidies

In order to examine the pass-through of subsidy changes, we determine the change in subsidy for each characteristic (demographic category or diagnosis) in the risk adjustment system. As discussed in Section 2.3, this step is nontrivial because the revision also altered diagnosis definitions and the treatment of demographic categories. To measure the updates, we calculate the total subsidies for Part D enrollees in 2011 under both the new and old subsidy systems. Both new and old subsidies are based on the same underlying data; only the risk adjustment system varies. We introduce the notation S_i(system, data year) to denote individual i's subsidy based on a given subsidy system and a given year's data[12]. $\Delta S_i = S_i(new, 2011) - S_i(old, 2011)$ is the difference in an individual's subsidy induced solely by the subsidy revision. We measure ΔS_i for 764,621 individuals enrolled in free-standing Part D in 2011 and in fee-for-service Medicare in 2010, and report its distribution in Panel (a) of Figure 1.[13] The average change in subsidy holding characteristics fixed is -$31.[14]

The variants S_i^{diag}(system, data year) and S_i^{demo}(system, data year) denote the diagnostic and demographic components of the subsidy[15]. Panels (b) and (c) in Figure 1 reports the distribution of ΔS_i^{diag} and ΔS_i^{demo}. We first note that the overall decrease in total subsidy is the sum of an increase in demographic subsidies of about $91 and a $122 average decrease in diagnostic subsidies.[16] We also point out the bimodal distribution of the change in demographic subsidies; many demographic categories had only small changes in their risk adjustment rate, but demographic risk adjustment for those who are long-term institutionalized increased by several hundred dollars. For diagnostic risk adjustment, there is a point mass at no change for individuals who have no diagnoses under either system.

[12] To simplify, we refer to a data year. But since the diagnostic information used in risk adjustment is retrospective while the demographic information is concurrent, "data year 2011" refers to 2011 demographics and diagnoses from 2010 medical claims.

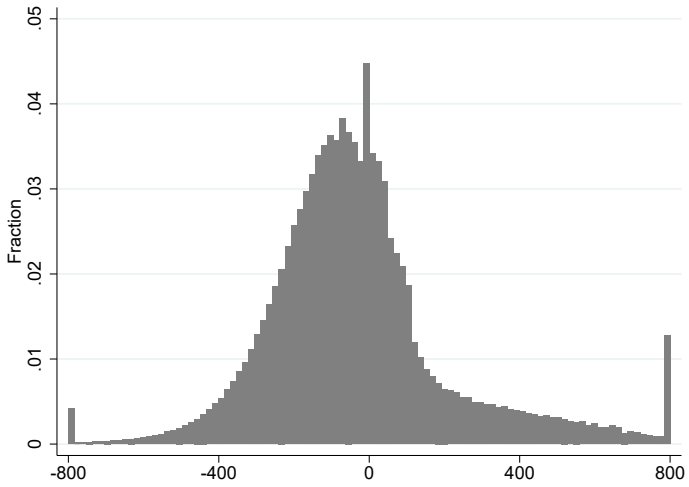
[13] Appendix Figure A.3 reports the distribution of $S_i(old, 2011)$ and $S_i(new, 2011)$ reports the distribution of total subsidies, as well as its demographic and diagnostic components.

[14] However, if we instead examine the realized subsidy for the Part D population in 2010 and 2011, the change in subsidy is much smaller at -$18, representing about 2% of the average 2010 subsidy. There are a number of possible explanations for this fall in average subsidy. While total drug costs are slightly increasing over the time period, the subsidies are designed only to compensate plans for their outlays. Much of the increase in total drug costs occurred for high-priced drugs for which plan outlays are limited by government reinsurance.

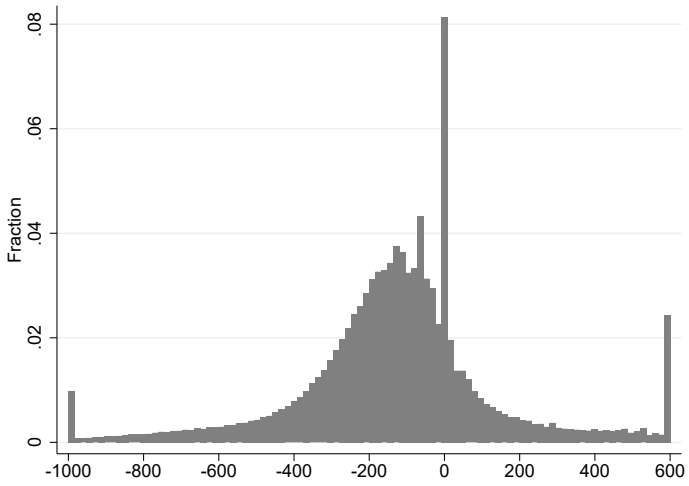
[15] The demographic and diagnostic components are not entirely additively separable under either system. In the old system, low-income subsidy receipt and long-term institutionalization are factor multipliers for the total subsidy. In the new system, the subsidy weight for each diagnosis and age-gender category is defined for five subpopulations: low-income subsidy × disabled and long-term institutionalized. We define the diagnostic subsidy as the component attributable to diagnoses alone in the old system, and diagnosis × subpopulation in the new system. The demographic component is defined as the total subsidy less the diagnostic subsidy.

[16] This decrease in diagnostic subsidies can arise if the diagnoses included in the risk adjustment system simply explain less of the variance in spending in the new model than they were in the old model, such that more spending loads on the demographic factors.

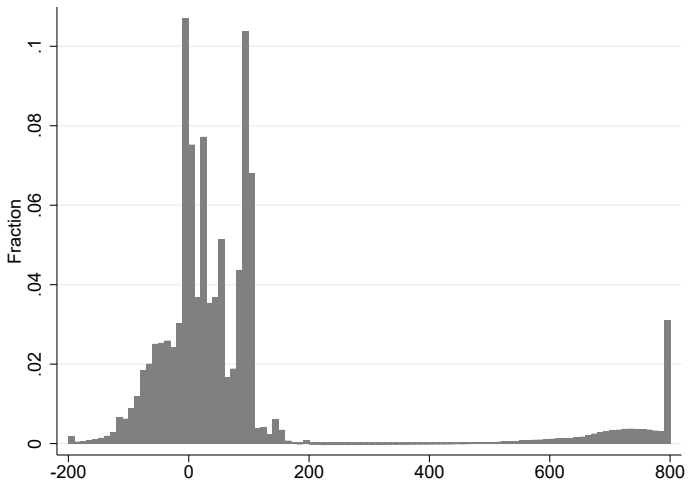
Figure 1: Distribution of Change in Individuals' Subsidies Induced by 2011 Subsidy Revision



(a) Change in Total Subsidy



(b) Change in Diagnostic Subsidy



(c) Change in Demographic Subsidy

This figure depicts the difference between individuals' subsidies induced by the revision of the subsidy system. The histogram in Panel (a) reports the difference in total subsidy, ΔS_i, while Panel (b) reports the difference for the diagnostic component and Panel (c) the demographic component. Each histogram displays 100 bins and is bottom- and top-coded to the minimum and maximum x-axis values.

4.3 Measurement of Change in Plan and Diagnosis Subsidies

In Section 5 we develop an estimation model to measure subsidy pass-through to enrollee premiums and OOP cost. To measure the pass-through rate to premiums, we compare plan premiums to average subsidies in each year 2009-2012. However, in order to isolate changes in a plan's average subsidy that are attributable to the subsidy system revision, we measure average subsidy in each year using plan j's 2010 enrollment and the enrollees' 2010 characteristics: $\overline{S_{jt}} = \frac{1}{N_{j2010}} \sum_{i \in j \text{ in } 2010} S_i(\text{system in year } t, 2010)$. E.g., $\overline{S_{j2011}}$ measures the average subsidy a plan would expect if their 2011 enrollment was comprised of the exact same individuals as their 2010 enrollment and those individuals had the same diagnoses and demographics. This measure removes changes to a plan's enrollment that may endogenously arise from its strategic behavior.

To measure the pass-through rate to OOP costs, we measure the subsidy a plan receives for each diagnosis in the years 2009-2012. For the years 2009-2010, the subsidy for each diagnosis x (as defined by the old system) is given by the weights ω_x from Equation 1 multiplied into dollars using the national average bid. However, because diagnoses are reorganized in 2011, we use the data to estimate the average change in diagnostic subsidy associated with each diagnosis under the old definitions. To do so, we regress ΔS_i^{diag} on a set of dummies for the 84 diagnoses δ_{ix}.

$$\Delta S_i^{diag} = S_i^{\text{diag}}(\text{new, } 2011) - S_i^{\text{diag}}(\text{old, } 2011) = \sum_x \delta_{ix} U_x + \varepsilon_i \tag{2}$$

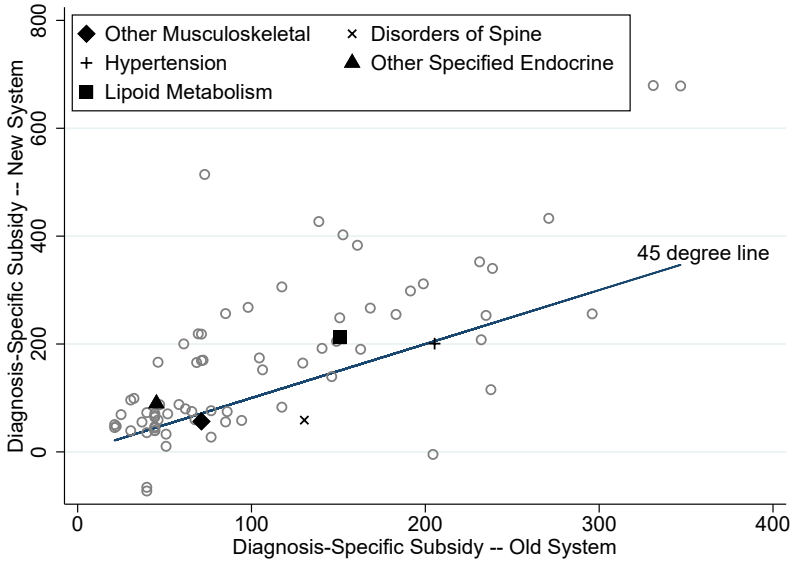
The coefficients U_x capture the change in diagnostic subsidy associated with diagnosis x, which we refer to as the "subsidy update" for diagnosis x. Our estimate of the subsidy associated with diagnosis x in the years 2011 and 2012 simply adds U_x to the subsidy value for 2009 and 2010. We estimate Equation 2 using 764,621 individuals enrolled in free-standing Part D in 2011 and in fee-for-service Medicare in 2010.

Appendix Table A.2 reports each diagnosis's old subsidy as well as the subsidy update U_x and its robust standard error as estimated by Equation 2. The diagnoses are sorted by the magnitude of the old subsidy. Note that standard errors are quite small relative to coefficients; we nearly always reject the hypothesis that a diagnosis's subsidy is the same under both systems.

Two figures illustrate the subsidy updates between 2010 and 2011. Figure 2 graphs subsidies before (x-axis) and after (y-axis) the subsidy update, with the 45 degree line, which would imply no update, provided for reference. The five most common diagnoses are labeled. Figure 3 shows the magnitude of subsidy updates U_x across diagnoses, sorted by the magnitude of old subsidies. Updates are economically large; in addition, it is clear that subsidy updates are not strongly related to the magnitude of old subsidies.[17]

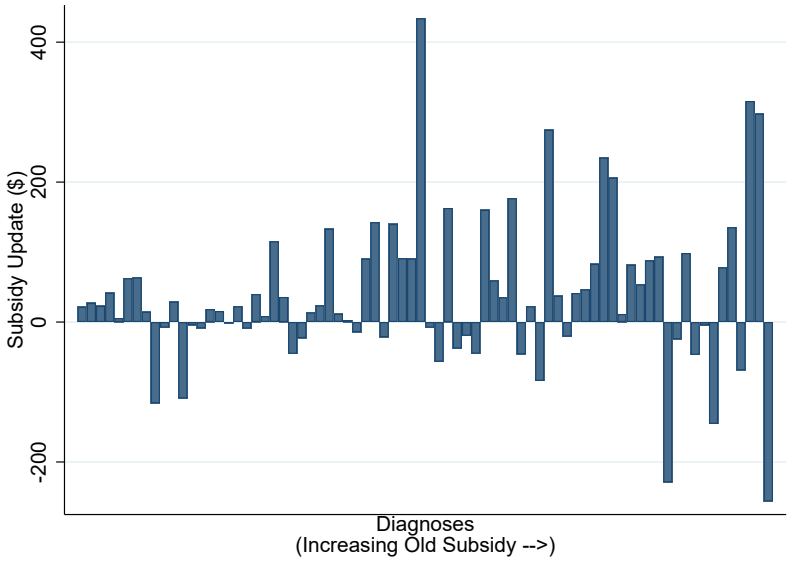
[17] While Figure 2 suggests a slight positive association between the magnitude of the old subsidy and the size of the subsidy update, the association is negative when we include HIV/AIDS, and is statistically indistinguishable from zero.

Figure 2: Diagnostic Subsidies Pre- and Post- Recalibration



Each marker in this figure represents one of 75 diagnoses (excluding HIV/AIDS for scale). The diagnosis's 2010 subsidy is measured along the x-axis, and the diagnosis's 2011 subsidy along the y-axis. The five most common diagnoses are identified in the legend.

Figure 3: Magnitude of Subsidy Updates



This figure displays the subsidy updates reported in Table A.2. The 76 diagnoses used in later analyses are arrayed along the y axis by increasing 2010 subsidy level.

5 Measuring Pass-Through of Medicare Part D Subsidies

In this section, we exploit the subsidy changes measured in the previous section to estimate the pass-through of subsidies to premiums and OOP costs.

5.1 Pass-Through of Subsidies to Premiums

To estimate pass-through of subsidies to premiums, our dependent variable is annual premiums for plan j in year t. The independent variable is the average subsidy a plan receives using t's subsidy system and plan j's 2010 enrollment: $\overline{S_{jt}} = \frac{1}{N_{j2010}} \sum_{i \in j \text{ in } 2010} S_i(\text{system in year } t, 2010)$. We also examine how premiums respond to the diagnostic and demographic components of subsidies.

$$\text{premium}_{jt} = \eta \overline{S_{jt}} + \delta_j + \delta_t + \varepsilon_{jt} \tag{3a}$$

Our model also includes plan (δ_j) and year (δ_t) fixed effects. The plan fixed effects account for all time-invariant aspects of plan premiums, while the year fixed effects account for annual differences in the Part D program. Our equation is analogous to a two-way fixed effects difference-in-difference model: we compare the change in premiums for plans receiving positive subsidy updates relative to plans that receive negative subsidy updates. In this setting, the "parallel trends" assumption is that plans whose average subsidy was increased or decreased by the subsidy revision would have otherwise evolved similarly to plans whose average subsidy was similar under both subsidy systems.

Given our fixed effects, our key identifying variation arises from the difference in subsidy induced by the 2011 revision: $\Delta\overline{S_j} = S_{j2011} - S_{j2010}$. In the left-hand column of Figure 4, we show a scatter plot of the average change in subsidy between 2010 and 2011 (x-axis) and the change in annual basic premium between 2010 and 2011 (y-axis). Each marker represents a plan, and the size of the marker is the plan's enrollment. Panels (a) and (c) reveal that there is no relationship between the change in plan's premiums and the change in average total or diagnostic subsidy. However, panel (e) shows a strong negative relationship between changes in demographic subsidies and premiums.

To support a causal interpretation of the relationship between subsidy and premium, the right-hand column of the Figure 4 reports the results of the event study version of Equation 3a.

$$\text{premium}_{jt} = \sum_{\tau=2009,2011,2012} \eta_\tau \Delta\overline{S_j} + \delta_j + \delta_t + \varepsilon_{jt} \tag{3b}$$

In this event study, the subsidy update is used to predict changes in premiums between all three year pairs, e.g. η_{09} measures the change in premiums between 2009 and 2010 associated with a one dollar

(future) subsidy update. If η_{09} is small or zero, plans that will go on to have positive subsidy updates changed premiums similarly in the pre-period to those that will go on to have negative subsidy updates. The coefficients η_{11} and η_{12} measure the impact of the subsidy update, with η_{12} estimated separately in case plans do not fully respond in the first year. η_{10} is normalized to zero. The standard errors are clustered at the market level and the regression is weighted by the number of enrollees in plan j in year t.

The η coefficients are reported in Figure 4. Panels (b) and (d) show that premiums are not responsive to the average subsidy or its diagnostic component. However, a \$1 increase in a plan's average demographic subsidy is associated with a large reduction in the plan's premiums in 2011, partially reversed in 2012. There is no evident trend between 2009 and 2010 that explains this post-period pattern.

5.2 Pass-Through of Subsidies to Out-of-Pocket Costs

To measure the pass-through of subsidies to OOP costs, we first measure typical OOP costs associated with each diagnosis in the subsidy system. Next, we propose our empirical model and report its results.

5.2.1 Measurement of Out-of-Pocket Costs

To begin, we characterize the typical demand for drugs that treat diagnosis x among those with the diagnosis using the prescription drug claims in a fixed pre-revision year (2010). We then calculate the OOP cost of taking the typically demanded drugs in each plan 2009-2012.[18]

To characterize the typical demand for drugs that treat diagnosis x, we first determine the set of such drugs $\mathcal{D}_x$. We identify this set using a one-versus-all classifier that takes advantage of our large sample of individuals' diagnoses and their prescription drug claims; we report full details in Appendix Section A.4.1. We next total the months' supply for each drug $d \in \mathcal{D}_x$ in each zone of coverage z in year 2010 of the claims, averaging across all those who have the diagnosis I_{x2010}.[19]

$$\text{months}_{dz2010} = \frac{1}{I_{x2010}} \sum_i \text{months}_{idz2010}$$

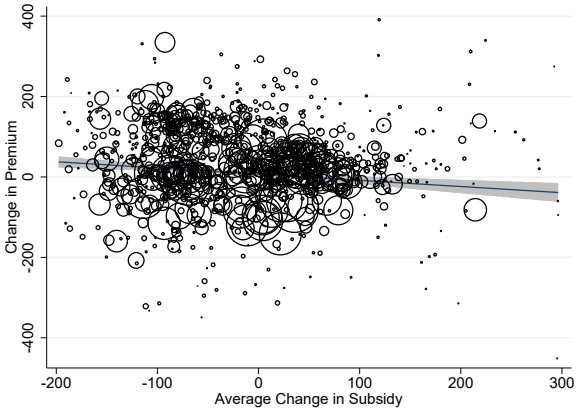
The use of zone-specific quantities accounts for the fact that those with expensive diagnoses tend to incur most of their demand in later zones (e.g., the donut hole or catastrophic zones). We use 2010 claims because demand after the revision may be induced by changes in OOP costs, and thus calculating the outcome under 2010 demand more cleanly estimates pass-through of subsidy to OOP costs.

Next, we compute the OOP costs in each plan for the estimated diagnosis-specific typical demand. The OOP costs W for diagnosis x in plan j in year t are equal to the dot product of typical (2010) demand for

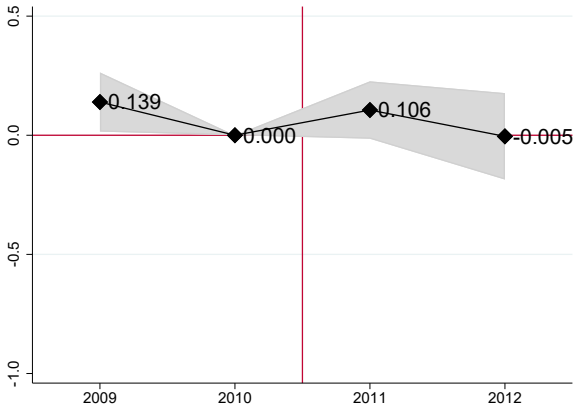
[18] Part D enrollees who get the low-income subsidy also receive a cost-sharing subsidy that further reduces their OOP costs. We define OOP cost to be the amount before cost-sharing subsidies, since this is the plan's choice variable.

[19] A difficulty arises here about how to handle those who take drugs that treat the diagnosis but do not have an ICD-9 code for the diagnosis in their medical claims; our approach is described in Appendix Section A.4.2.

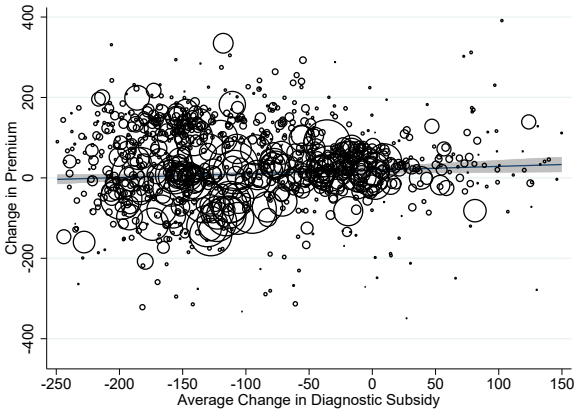
Figure 4: Impact of Change in a Plan's Average Subsidy On Premiums



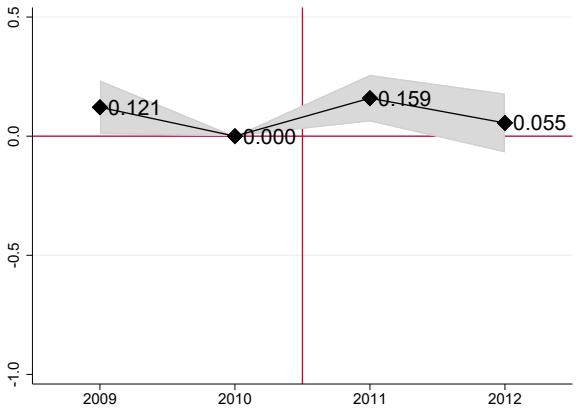
(a) Changes in Total Subsidy



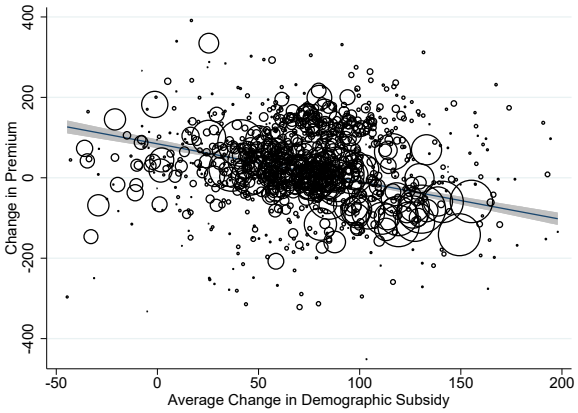
(b) Event Study: Total Subsidy



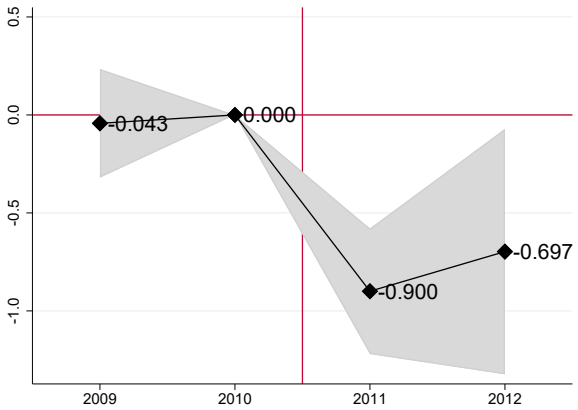
(c) Changes in Diagnostic Subsidy



(d) Event Study: Diagnostic Subsidy



(e) Changes in Demographic Subsidy



(f) Event Study: Demographic Subsidy

Panel (a) reports the scatter plot of a plan's change in premiums between 2010 and 2011 (y-axis) and the change in a plan's average total subsidy ($\Delta\overline{S_j} = S_{j2011} - S_{j2010}$, holding the plan's enrollment and the enrollee characteristics fixed at their 2010 levels) on the x-axis. Panel (c) and (e) use the same y-axis but use the diagnostic (c) or demographic (e) components of subsidy. The marker size represents the number of enrollees in 2010, and the weighted least squares line is also reported. Panels (b), (d), and (f) report the event study coefficients estimated by Equation 3b representing the impact of the change in total (b), diagnostic (d), and demographic (f) subsidies on annual plan premiums.

each drug that treats x and j's OOP costs in year t, summed across the zones of coverage.

$$W_{xjt} = \sum_{z} \sum_{d \in \mathcal{D}x} \text{months}_{dz2010} * \text{OOP cost}_{jdzt}$$

If a plan does not cover the drug in any zone of coverage, we impute OOP cost equal to the average price of the drug in all plans in that year. We will examine coverage as an outcome variable in Section 6.2.

Figure 5 depicts a histogram of year-over-year changes in annual OOP costs W_{xjt} for diagnosis-plan combinations. Two patterns in these histograms suggest the importance of the revision. The first is that the peak at no change is notably lower at revision (panel b) than in the other two years. The second pattern is that the share of plan × diagnosis combinations with reductions in OOP costs is much higher between 2010 and 2011, when many diagnoses received a subsidy increase.

5.2.2 Empirical Model for Pass-Through of Diagnostic Subsidies to Out-of-Pocket Costs

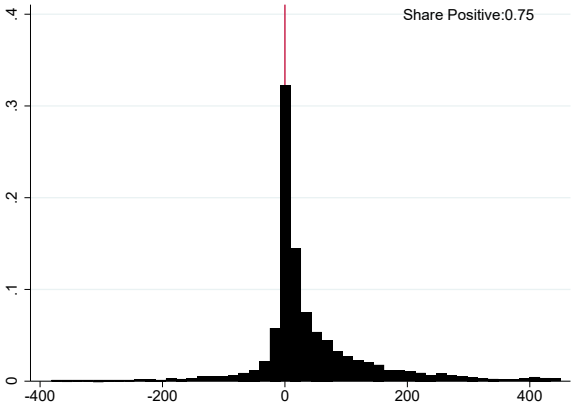
The effect of subsidies on OOP costs is estimated using Equation 4a.

$$\overbrace{W_{xjt}}^{\substack{\text{typical OOP for diag } x \\ \text{in plan } j \\ \text{in year } t}} = \beta \overbrace{S_{xt}}^{\substack{\text{subsidy} \\ \text{for diag } x \\ \text{in year } t}} + \overbrace{\delta_{xj}}^{\substack{\text{diag} \times \text{plan} \\ \text{fixed} \\ \text{effect}}} + \overbrace{\delta_{jt}}^{\substack{\text{plan} \times \text{year} \\ \text{fixed} \\ \text{effect}}} + \lambda \overbrace{X_{xjt}}^{\substack{\emptyset \text{ or} \\ \text{price} \\ \text{spline} \\ \text{or diag} \\ \text{trend}}} + \varepsilon_{xjt} \tag{4a}$$

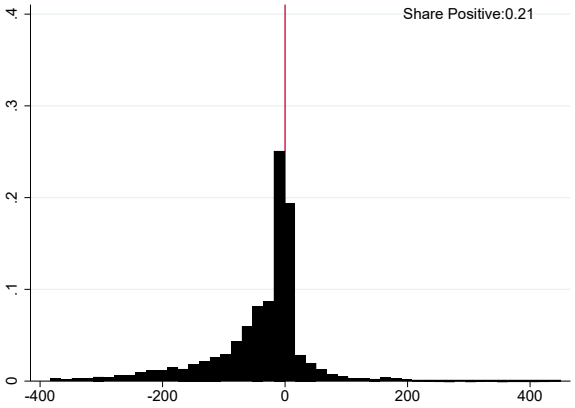
The pass-through rate is captured by β, which gives the change in typical annual OOP costs for a \$1 increase in subsidy.[20] We include two sets of fixed effects: plan × diagnosis and plan × year. The fixed effect δ_{xj} represents all time-invariant demand or supply factors that affect the OOP costs for this plan × diagnosis observation. This controls for unobserved drug efficacy and side effects, marginal cost of production, or (time-invariant) market power of the drug's maker. In addition, it controls for the plan's time-invariant preferences for individuals with this diagnosis, such as a strong negotiating position with the relevant drug firms. The plan × year fixed effect corrects for any plan-level changes that treat all diagnoses equally, such as a change in plan strategy that affects all OOP costs, and can also account for unobserved cost or demand shocks that are not specific to diagnoses (e.g., brand effects, administrative costs). Given these fixed effects, our identification comes within plan × diagnosis observations as subsidies vary over time due to the revision. The key identifying assumption is that the pattern of OOP costs for diagnoses that received no or very small subsidy changes is a good counterfactual for OOP costs for diagnoses with large subsidy changes.

[20] As discussed in Miller et al. (2017) and Muehlegger and Sweeney (2017), under oligopolistic competition pass-through of a subsidy or cost shock is the sum of the direct effect on the firm's price and the indirect effect of competitors' price changes inducing further price adjustment. However, only the net effect of these shocks can be obtained when the subsidy or cost shock is industry-wide, as it is in our setting and that of Miller et al. (2017).

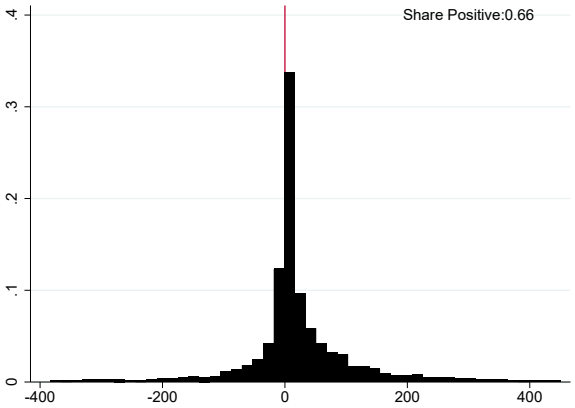
Figure 5: First Differences in Typical Out-of-Pocket Cost in Consecutive Year Pairs



(a) $W_{xj2010} - W_{xj2009}$ ($)



(b) $W_{xj2011} - W_{xj2010}$ ($)



(c) $W_{xj2012} - W_{xj2011}$ ($)

Each panel reports a histogram of first differences in typical annual OOP costs for a given diagnosis in a given plan between consecutive year pairs. The subsidy revision takes place between 2010 and 2011.

We weight each observation by the plan enrollment and the number of individuals with the diagnosis.[21]

We cluster our standard errors in two ways: at the plan × year and at the diagnosis × market. The first clustering recognizes that plan benefit designs must comply with regulations requiring an actuarial value of 25%. While our plan × year fixed effect will absorb all positively correlated changes, these requirements may induce negative cross-sectional correlation in OOP costs within a plan × year; if OOP costs exceed 25% of price for some diagnoses, the OOP costs for others must be lowered to compensate. The second clustering allows arbitrary correlation in how plans in the same market design benefits for a diagnosis. Errors may be serially correlated across time in a diagnosis × market due to market-specific differences in diagnostic subtype or treatment preferences, or cross-sectionally correlated across plans due to competition.

We report three specifications. The simplest uses only the subsidy and the fixed effects ($X_{xjt} = \emptyset$). Our preferred specification adds controls that recognize the economic content of subsidy updates. Suppose that drug prices for a particular diagnosis have been rising since 2000. Price rises through 2008 will be incorporated into a positive subsidy update, but benefit designs in 2009-2012 will reflect the continued increase in price between 2009 and 2012. This could induce a positive correlation between subsidies and the error term of Equation 4a. In the below, we explicitly condition on the annual (pre-rebate) price for the typical demand for diagnosis x in plan j in year t. In particular, we control for a linear spline in price at ventiles of the price distribution.

In a more general sense, any persistent trends between 2000 and 2008 that affect benefit design can generate a spurious correlation between the subsidy update and OOP costs that is *not* via the pass-through of diagnosis-specific subsidies.[22] We therefore consider models with a time trend in diagnosis x. In the presence of the time trend, we are identifying β from any deviation from trend that occurs between 2010 and 2011.

We also estimate an event study implementation of Equation 4a. Similar to Equation 3b, the event study interacts the change in subsidy with year indicators to predict OOP costs in each year. Each β_τ represents the change in OOP costs for that diagnosis × plan combination in year τ for a \$1 increase in the subsidy

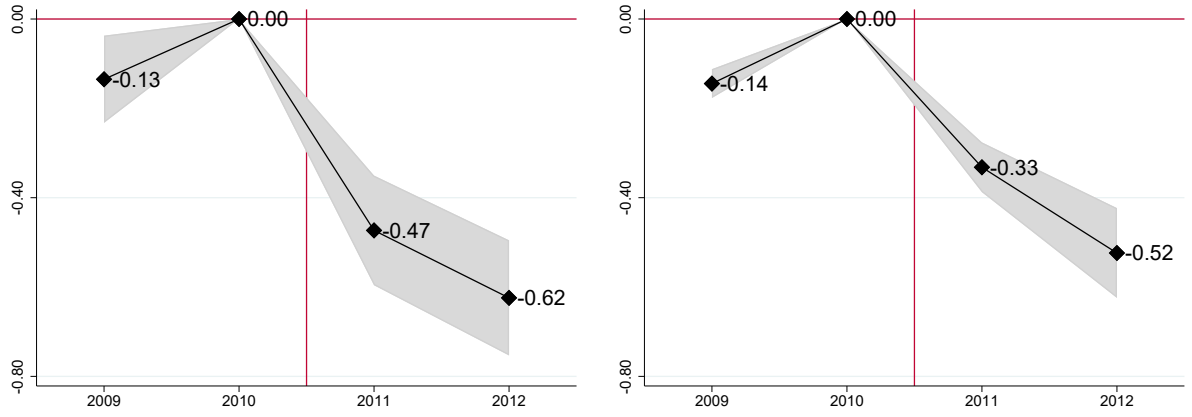
[21] In the framework of Solon et al. (2015), these weights recover the average partial effect of subsidies in the presence of unmodeled heterogeneity across plans and diagnoses in the response of agents to the change in incentives. Such heterogeneity would result if plans reoptimize benefits for popular diagnoses but ignore the long tail of uncommon diagnoses, or if larger plans are more likely to reoptimize.

[22] To see why, suppose each year since 2000 insurers have simply raised the OOP costs for drugs that treat diagnosis x a fixed amount I_x: $OOP_{xt} = OOP_{x00} + I_x t$. Medicare's subsidy recalibration process finds that the diagnosis-specific costs in 2008 are a linear function of OOP costs plus some error: $\omega_{x08} = \rho OOP_{x08} + \nu_{x08}$, where ν_{x08} captures all the other features of costs in 2008. Subsidy in 2011 is set to ω_{x08}. Insurers continue to raise OOP costs, so for example the change in OOP costs between 2010 and 2011 is simply I_x. In our analysis, we will calculate that the subsidy update is $U_x = \omega_{x08} - \omega_{x00} = \rho(8I_x) + \nu_{x08} - \nu_{x00}$. If we use this subsidy update to identify Equation 4a we will find that the change in subsidy and change in OOP costs are correlated through I_x. But there is no "pass-through" in this setting – it is simply that a time trend in OOP costs influences both the updated subsidy and the change in benefit designs that interests us.

between 2010 and 2011.[23]

$$\overbrace{W_{xjt}}^{\substack{\text{typical}\\ \text{OOP/coverage}\\ \text{for diag } x\\ \text{in plan } j\\ \text{in year } t}} = \sum_{\tau=09,11,12} \beta_\tau \overbrace{U_x}^{\substack{\text{subsidy}\\ \text{update}\\ \text{for diag } x}} \mathbb{1}(t=\tau) + \overbrace{X_{xjt}}^{\substack{\emptyset \text{ or}\\ \text{price}\\ \text{spline}}} + \overbrace{\delta_{xj}}^{\substack{\text{diag}\times\text{plan}\\ \text{fixed}\\ \text{effect}}} + \overbrace{\delta_{jt}}^{\substack{\text{plan}\times\text{year}\\ \text{fixed}\\ \text{effect}}} + \varepsilon_{xjt} \tag{4b}$$

Figure 6: Event Study: Impact of Diagnosis Subsidy Updates Over Time



(a) Typical Out-of-Pocket Costs ($), no controls

(b) Typical Out-of-Pocket Costs ($), price spline

These figures depict the coefficients from estimation of Equation 4b. Panel (a) has no controls, while Panel (b) controls for a price spline.

We present the event study coefficients in Figure 6. We find that diagnoses that go on to have positive subsidy updates were slightly raising their OOP costs between 2009 and 2010. This movement is what we would expect if, prior to the subsidy revision, these diagnoses were becoming more expensive and plans were raising OOP costs to compensate. After the revision, there is a large apparent reduction in OOP costs. While ideally the coefficient on β_{09} would be statistically zero, it's clear from these event studies that response to the subsidy update is large relative to pre-existing trends.

5.3 Comparing Pass-Through to Premiums and Out-of-Pocket Costs

The event studies in Figures 4 and 6 support the interpretation of η_t and β_t in Equations 3a and 4a as pass-through rates. Table 1 reports these estimated pass-through rates. Panel A summarizes our findings for pass-through to premiums: no pass-through of total subsidy or diagnostic subsidy, but a pass-through

[23]Estimating the specification with the diagnosis-specific trend from Equation 4a would require normalizing another β_τ to zero, because the three year-over-year changes cannot identify three β_τ coefficients as well as a diagnosis-specific trend. Thus, we exclude the diagnosis-trend specification from event study estimation.

Table 1: Pass-Through to Premiums and Out-of-Pocket Costs

Panel A: Premium			
total subsidy	-0.0072		
	(0.0628)		
diagnostic subsidy		0.0606	
		(0.0506)	
demographic subsidy			-0.7390*
			(0.2310)
fixed effects	plan, year		
N	4,824 plan × year obs		

Panel B: Typical Annual OOP Cost			
Specification:	none	price spline	diagnosis trend
diagnostic subsidy	-0.466**	-0.369**	-0.467**
	(0.0508)	(0.0319)	(0.114)
fixed effects	plan × diagnosis, plan × year		
N	397,100 plan × diagnosis × year obs		

This table reports the pass-through rates estimated for premiums using Equation 3a and for OOP costs using Equation 4a. In panel A, the dependent variable is annual premiums in the years 2009-2012, and the independent variable is the plan's average subsidy (total, diagnostic component, or demographic component) using the risk adjustment system in year t and the plan's enrollment in 2010. Analyses are weighted by plan enrollment and standard errors are clustered on the market. In panel B, the dependent variable is the typical annual OOP cost for a diagnosis in a plan in the years 2009-2012, measured as described in Section A.4.2. The independent variable is the subsidy for that diagnosis, measured as described in Section 4.3. In Panel B, analyses are weighted by plan enrollment and the number of Part D enrollees who have the diagnosis, and standard errors are two-way clustered on plan×year and diagnosis×market. +, * and ** represent significance at the 10, 5 and 1 percent levels.

rate of 74% of demographic subsidies to premium (SE 23pp). In Panel B, all three specifications estimating pass-through of diagnostic subsidy to OOP costs find a rate of about 40%, with overlapping confidence intervals.

In Section 3.2 we developed a theoretical model suggesting that diagnosis-specific subsidies should be passed through exclusively to the diagnosis's OOP costs. Table 1 reports considerable pass-through of diagnosis-specific subsidies to OOP costs (40%) but no statistically-significant pass-through to premiums. Thus, these empirical findings are consistent with our theoretical model. When diagnosis-specific subsidies increase, insurers focus pass-through on the strategic variable that most affects those with the diagnosis.

The analysis suggests that the demographic component of subsidies is passed-through to premiums. Our theoretical model suggested that pass-through to premium was more likely for a broad subsidy that affected all enrollees. The demographic component of subsidies is not exactly a broad subsidy as represented in the model: it is specific to a given demographic category. However, the characteristics that define the demographic groups – age, sex, disability, long-term institutionalization – are *not* clearly linked to any particular drugs. Thus, the insurer cannot differentially attract individuals in the given demographic category through selective reductions in OOP costs. In the theoretical model, we suggested that the insurer may be indifferent between passing-through a broad subsidy to premium or to a broad reduction of OOP costs. However, any lower salience of OOP costs would push the insurer to favor reducing premium.

One final possibility to consider is how subsidies are passed-through to overall OOP costs. This is not tested explicitly in Equation 4a because of the plan $\times$ year fixed effect. However, we can collapse across diagnoses x to create the average OOP costs in a plan. Using the average OOP costs in a plan as the dependent variable in Equation 3a, we find no response of average OOP costs to subsidies.[24]

6 Mechanisms of Insurers' Benefit Design Response

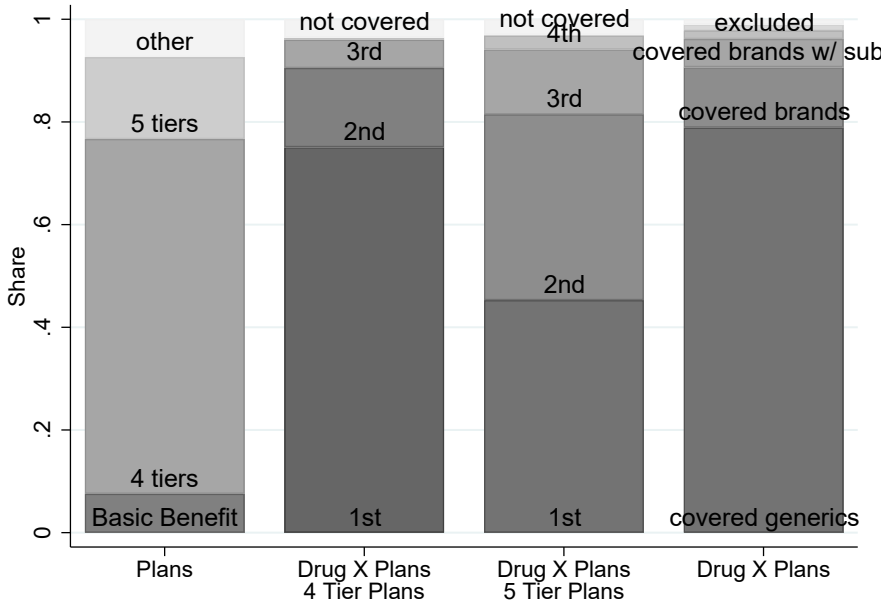
Having established in Section 5.2 that insurers indeed reduce the OOP costs for drugs that treat diagnoses with subsidy increases, we now turn to the exact margins of insurers' responses. We find that insurers largely achieve OOP cost reductions by improving benefits for branded drugs – covering slightly more branded drugs and moving branded drugs to lower formulary tiers. Insurers may also choose to reduce non-price "utilization management" barriers, or may choose to increase them to offset moral hazard induced by lower OOP costs; we find mixed results on the response of utilization management to subsidy changes.

[24] Results available upon request.

6.1 Empirical Model for Benefit Design Response

In Section 5.2, we aggregated OOP costs for drug × plan observations using the typical annual demand in order to correctly measure the pass-through rate for an annual diagnostic subsidy. In this section, we return to outcomes at the level of the drug × plan × year. This change in the unit of observation allows us to subset on key drug characteristics such as brand/generic status, as well as to study outcomes such as formulary tier that only have ordinal meaning within a plan. The disadvantage of this approach is that the effect of a diagnosis's subsidy on a single drug's OOP cost cannot be interpreted as a pass-through rate. However, the direction and magnitude of these effects indicate exactly how insurers are responding.

Figure 7: Distribution of Benefit Design Outcomes: Number of Tiers, Proportion of Drugs on Each Tier, and Coverage Status



The left-most bar represents the share of plans (2009-2012), weighted by enrollment, in each of four categories of plan design: Basic Benefit, four tiers, five tiers, or another plan design. The two central bars represent the share of plan × drug × year observations on each tier of coverage among four tier and five tier plans. The right-most bar represents the share of plan × drug × year observations in each of six categories: covered generics, covered brands, covered brands with an exact generic substitute, and the uncovered complements of each. The second, third, and fourth bars weight each plan × drug × year observation with the plan enrollment and the number of individuals who take the drug in Medicare Advantage (the same weighting applied in Equation 3).

We first describe the benefit designs in use in Part D plans over the sample period. The first column in Figure 7 shows the distribution of plans across four types, weighted by plan enrollment. Recall from Section 2.2 that insurers can trivially satisfy the Part D actuarial value regulations by offering the "Basic

Benefit" of 25% coinsurance for all drugs. Ten percent of plans are taking this option and thus not exercising any diagnosis-specific latitude.[25] For plans choosing to exercise more fine-grained control of benefits, drugs are commonly placed on "tiers". Drugs on the same tier in a plan have a uniform dollar copay (or, less commonly, coinsurance rate), with the copay rising with the tier. Favorable tier placement is traded off for higher discounts in price negotiations with drug manufacturers. In 2010, the modal benefit design has four tiers: a generics tier, a preferred brands tier, a non-preferred brands tier, and a specialty tier.[26] The first three tiers have flat copays while the specialty tier, where expensive and less-common drugs are placed, usually has a coinsurance.

The second column of Figure 7 shows the distribution of drug × plan × year observations across the four tiers among four-tier plans. We weight each observation by the product of plan enrollment and the number of MAPD enrollees who take the drug, which is the same weighting we use in estimation below and analogous to the weighting for Equation 4a.[27] The fourth tier, commonly the specialty tier, is depicted but too small to be visible. A substantial fraction of other plans have five tiers, which usually implies a preferred and unpreferred generics tier; the weighted distribution of drugs across tiers in these plans is depicted in the third column of the figure (again the top tier is not visible).

We will examine benefit design outcomes separately for brands and generics because both the formulary treatment and upstream markets differ between these two drug types. The last column of Figure 7 reports the weighted distribution of drug × plan × years by coverage status for three drug types. "Brands" are drugs whose ingredients are only sold in a drug with a brand name in the calendar year. Among drugs whose ingredients are sold in generic drugs, we distinguish between brands with a substitute (i.e., branded Lipitor after the entry of generic atorvastatin) and generics (atorvastatin). More than half of the observations are covered generic drugs, generally occupying the first tier in four tier plans and the first two tiers in five tier plans. Covered brands as well as covered brands with substitutes occupy the upper tiers. Weighted rates of coverage are high for generics (98%) and moderate for brands (91% of brands without substitutes, 84% of brands with substitutes).

This figure motivates our four key benefit design outcomes Y_{djt}. The first is the monthly OOP cost for a drug d in the initial coverage zone in plan j in year t. The diagnosis-level outcome W_{xjt} in the pass-through estimation reflects costs in all four zones of coverage, but OOP costs in the initial coverage zone account for the majority of all OOP costs. If drug d is not covered by plan j, the OOP cost is imputed to the average price for the drug among covering plans. Thus, our next outcomes decompose changes in OOP cost into

[25] When analyzed separately, "Basic Benefit" plans tend to adjust benefit design parameters less than other plans, as expected, but standard errors are large due to small sample size. Results available upon request.

[26] For 2010 only, we have an auxiliary dataset that reports plans' descriptions of the type of drugs on each tier. This dataset is not available for other years. Not all four-tier plans in 2010 are using these tier types.

[27] For these drug-level analyses, we use utilization in Medicare Advantage to reduce sensitivity to the subsidy revision.

its two components: whether the drug is covered, and the OOP cost conditional on coverage. Finally, we measure the tier of coverage of the drug in this plan.

Our estimation equation evaluating the impact of diagnosis-specific subsidies on benefit design outcomes is:

$$\overbrace{Y_{djt}}^{\substack{\text{benefit design for drug } d\in\mathcal{D}_x \\ \text{in plan } j \\ \text{in year } t}} = \beta \overbrace{S_{xt}}^{\substack{\text{subsidy for diag } x \\ \text{in year } t}} + \overbrace{\delta_{dj}}^{\substack{\text{drug}\times\text{plan} \\ \text{fixed} \\ \text{effect}}} + \overbrace{\delta_{jt}}^{\substack{\text{plan}\times\text{year} \\ \text{fixed} \\ \text{effect}}} + \lambda \overbrace{X_{djt}}^{\substack{\emptyset \text{ or} \\ \text{price} \\ \text{spline} \\ \text{or drug} \\ \text{trend}}} + \varepsilon_{djt} \tag{3}$$

This equation is a direct analog to Equation 4a at the plan × drug level, so we highlight only the differences in estimation details. We use the drug-diagnosis linkage described in Appendix Section A.4.1 to determine the set of drugs that treat each diagnosis x. We use a plan × drug fixed in order to isolate the change in benefit design associated with the over-time variation in subsidies. The analyses are weighted by plan j's enrollment times the number of individuals who take drug d in Medicare Advantage. When estimating this equation, we drop Basic Benefit plans, which are not actively designing benefits, and we focus on generics and brands without substitutes, which represent more than 95% of drug demand in this sample.

6.2 Response of Benefit Designs to Subsidy Revision

Before reporting the results of Equation 3, we support a causal interpretation by estimating an event study analogous to Equation 4b for each outcome and subsample and reporting the results in Appendix Figures A.4 and A.5. We again note that we cannot estimate an event study that includes the drug trend. The event studies with no controls suggest some potential divergent pretrends, although all still show a sharp drop in 2011. The event studies with a price spline, our preferred specification, suggest parallel pre-trends with a clear difference in 2011.

Panel (a) of Figure 8 reports our estimate of the effect of subsidies on OOP cost for branded and generic drugs. Each coefficient is estimated in a separate regression. We can see that there is almost no response by generics with copays. Instead, the reductions in OOP costs are driven by branded drugs.[28]

Out-of-pocket cost is determined by both coverage and the OOP cost conditional on coverage. Panel (b) reports the coefficients, in percentage points, when the outcome is coverage. For both types of drugs, the magnitudes of the changes in coverage are very small: in the model with no controls, $1 in increased subsidy reduces coverage for generics by -0.00346 percentage points, off of mean of 98%, or increases coverage for brands by 0.00412 percentage points off a mean of 91%. Thus, even a hundred-dollar increase in subsidy would change coverage by less than half of a percentage point for either drug type. Panel (c) shows how the subsidy

[28] In contrast to the findings in Table 1, the coefficients differ for the branded drugs across specification. Our view is that the price or trend controls are the preferred model.

increase affects OOP costs conditional on coverage. These coefficients are not statistically distinguishable from those in panel (a), suggesting that the primary way that plans are responding to the subsidy increase is via altering OOP costs rather than adding or removing drugs from their formularies. Finally, panel (d) uses formulary tier as the outcome (which exists only for covered drugs). While the coefficients reported in Panel (d) of Figure 7 are raw (in units of "tier"), we note that the standard deviation of tier for brands is very close to 1, so those coefficients approximate the effect in standard deviations. We can see that there is little response of tier among generics, because most generics are already on the lowest tier. However, branded drugs are significantly more likely to be moved to lower tiers in response to subsidy increases. The effect size suggests that a $100 increase in subsidy would reduce a branded drug's tier by 5% of a standard deviation.

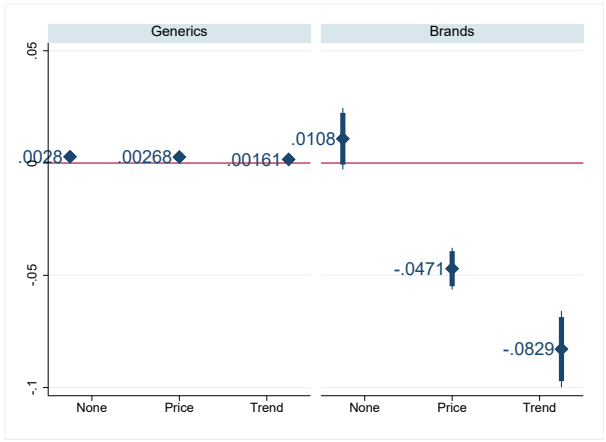
Overall, these results suggest that the rate of pass-through is greater for brands than generics. There are several reasons why this could occur. Firstly, Part D plans commonly cover all generics on the first tier, and so there simply is not very much scope for plans to move generics to other tiers, distorting pass-through for those drugs (Conlon and Rao, 2020). Secondly, generics are generally purchased from competitive upstream suppliers while brands are commonly supplied by upstream monopolies. While it is possible to solve for pass-through in the presence of an upstream monopoly, whether this rate is greater or less than the rate when the upstream market is competitive depends on demand curvature parameters. Finally, brands are much more expensive for Part D enrollees. If this greater expense means individuals are more sensitive to OOP costs for these drugs when enrolling in plans, this would alter the curvature of enrollment demand for these drugs, which would alter pass-through rates.

6.3 Response of Utilization Management to Subsidy Revision

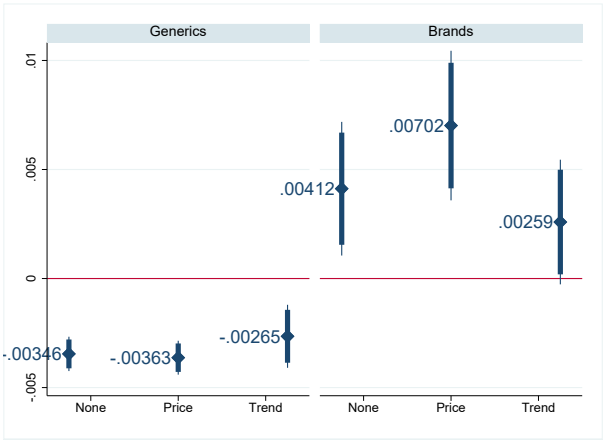
Insurers also have non-financial strategic variables known as utilization management restrictions. These restrictions affect the quantity demanded of a drug by requiring a patient to first try a cheaper competitor (step therapy), requiring a plan's agreement prior to prescription filling (prior authorization), or limiting the quantity of the drug that can be purchased under the insurance (quantity limits). It's possible that insurers would reduce these restrictions in response to a subsidy increase, just as they reduce OOP costs, in order to attract the enrollment of individuals who want to take these drugs. Alternatively, an insurer could pair OOP cost reductions with utilization management increases, to reduce the moral hazard impact of the OOP cost reductions. This strategy could arise if utilization management is less observable to potential enrollees than OOP costs.

Appendix Figure A.6 reports the results of estimating Equation 3 using indicators for the three utilization management tools. While utilization management tools are concentrated among branded drugs, we report the effects for the same branded and generic subsamples as in Section 6.2. Panel (a) reports the results for

Figure 8: Impact of Diagnosis-Specific Subsidies on Benefit Design Outcomes by Drug Subsample



(a) Out-of-Pocket Cost ($)



(b) Coverage (p.p.)

(c) Out-of-Pocket Cost | Coverage ($)

(d) Tier

This figure depicts the β coefficients from Equation 3. Each coefficient represents the effect of a $1 increase in the subsidy for the diagnosis the drug treats on the drug's benefit design outcomes in a plan $\times$ year. Each panel represents a different outcome. Within each panel, the coefficients are estimated separately for generics and brands. Three specifications are reported: no extra controls ("none"), a price spline ("price"), and a drug time trend ("trend").

step therapy, which is overwhelmingly used only among brands. Increases in subsidy reduce the use of step therapy for branded drugs, on the order of a 15% reduction in the prevalence of step therapy for a $100 increase in subsidies. Thus, step therapy and OOP costs react similarly to subsidy increases.

In Panel (b), we report the effects of subsidy on the use of prior authorization. We generally find increases in the use of prior authorization as subsidy increases (although they are relatively small in magnitude). Thus, it is possible that insurers institute prior authorization requirements in part to counter moral hazard induced by lower OOP costs. In addition, prior authorization is unique among utilization management measures in creating an opportunity for an interaction between an insurer and a prescriber. The insurer could use this opportunity to ensure the prescriber has noted the diagnostic code that generates the subsidy before authorizing this prescription. The rewards to noting the diagnostic code grow in the subsidy value.

In Panel (c), we report the effects of subsidy on the use of quantity limits. Quantity limits are by far the most commonly used utilization management tool, even common among generic drugs. However, the most common quantity limits still allow 30 days supply every 30 days, suggesting that many quantity limits allow the standard course of treatment and are not important in practice. These results are mixed and sensitive to specification, although coefficients are small relative to the mean.

Overall, our conclusion is that larger subsidies result in a small increase the use of prior authorization, perhaps due to the importance of noting diagnoses, and reduce the use of step therapy.

7 Interpreting the Pass-Through Rate

In Section 3, we considered two potential issues that would affect our interpretation of our pass-through estimates: a change in the pool of enrolled individuals in response to the subsidy revision, and a large drug-demand response to reductions in OOP costs after diagnostic subsidy increases. In this section, we argue that neither issue is economically important in our setting.

7.1 Response of Enrollment to Subsidy Revision

The first issue is whether individuals with diagnoses or demographics that receive subsidy increases differentially begin to enroll in Part D in 2011. Enrollment into Part D is not mandatory, although there is a "late enrollment" premium surcharge upon enrollment for individuals who choose to remain uninsured. Enrollment of previously-uninsured individuals can affect our measurement of the change in subsidies. Suppose Part D plans lower OOP costs by $0.40 after a diagnostic subsidy increase of $1, and uninsured individuals with that diagnosis choose to join Part D as a result. If those individuals have milder forms of the diagnosis and consume less prescription drugs, the enrollment of these individuals could make the *effective* increase in the subsidy larger than $1 (because the enrollees induced by the subsidy increase are cheaper than the

incumbent enrollees). A significant enrollment response would suggest that the true pass-through rates are closer to zero than what is reported in Table 1.

The variation in the subsidies over time lets us directly estimate the enrollment response. Our analysis uses as its key explanatory variable the subsidy change for each risk adjuster. For the diagnosis risk adjusters, we use the diagnosis-specific subsidy changes U_x estimated in Equation 2 and reported in Appendix Table A.2. For the demographic risk adjusters, we use the 2011-2010 change in demographic subsidies for 72 demographic bins created by the interaction of five enrollee types (disabled $\times$ dual eligible and long-term institutionalized) and the age-sex bins used in demographic risk adjustment. Our key dependent variable is the rate of Part D enrollment among Medicare beneficiaries who have this risk adjuster (i.e., have this diagnosis or are in this demographic bin) and are in the market for Part D, meaning the union of those enrolled in Part D and those not enrolled but without other prescription drug insurance.[29] In Appendix Figure A.7, the left-hand figures show a scatter plot of the changes in diagnostic and demographic subsidies and the change in the enrollment rate between 2010 and 2011. We find that enrollment increases by about 1.8 percentage points1, but this increase is unrelated to the increase in either diagnostic of demographic subsidies. The right-hand figures report on the estimation of an event study, Equation 4, which interacts the change in the subsidy for either a diagnosis or demographic category (indexed by r) with year dummies. The event study shows there is no response of enrollment to subsidy increases.

$$\text{enrollment rate}_{rt} = \sum_{\tau=2009,2011,2012} \eta_\tau U_r + \delta_r + \delta_t + \varepsilon_{rt} \tag{4}$$

Because I can rule out significant enrollment effects of the subsidy revision, I can rule out economically-meaningful selection effects. At the upper bound of the confidence interval, a one-dollar increase in a diagnosiss subsidy increases enrollment among individuals with that diagnosis by 0.0001 percentage points, such that even a hundred dollar increase would increase it by 0.01 percentage points. A one-dollar increase in a demographic groups subsidy increases enrollment among that demographic group by at most 0.002 percentage points, such that a hundred-dollar increase would increase it by only 0.2 percentage points. The upper-bound enrollment effects are too small to create an economically-meaningful selection effect.

7.2 Response of Drug Demand to Subsidy Revision

Another issue affecting interpretation of our pass-through rates is the possibility of a large drug-demand response. In Section 3, we examined enrollment demand, holding drug demand conditional on enrollment fixed at one unit per enrollee. However, pass-through to OOP costs can both increase enrollment among

[29] Thus, we exclude individuals enrolled in Medicare Advantage Part D plans, those receiving the retiree drug subsidy, and those known by Medicare to have "other creditable prescription drug coverage".

those with the diagnosis (modeled) and increase drug demand conditional on enrollment (unmodeled). If this second effect is present and large, it would distort our measurement of pass-through. E.g., if a subsidy increased by $1 and OOP costs fall by $0.40, but a drug demand response increased plans' outlays, the pass-through estimate should reflect the subsidy increase net of the increased plan outlays.

In Appendix Section A.6 we describe our estimation of the elasticity of drug demand, using the subsidy system update as an input cost shock in an instrumental variables (IV) demand estimation. By exploiting the revision of the subsidy system in an individual-level panel of drug demand, we recover an estimate of overall elasticity while flexibly controlling for time-invariant individual-level preference heterogeneity. Our elasticity estimates are 2-3%, somewhat lower than previous estimates from the demand change when Part D enrollees encounter a 100% coinsurance rate in the coverage gap (Einav et al., 2015; Jung et al., 2014). However, our lower elasticity captures how total annual demand responds to relatively small changes in OOP costs, rather than the demand response within the year when a beneficiary encounters the large, salient, and temporary increase in OOP costs at the onset of the coverage gap.

From the estimates in Appendix Table A.1, we can calculate the expected increase in drug demand. Using the IV estimates in the first column, we find that $1 in (annual) subsidy increase reduces (monthly) OOP costs by $0.01553. A $1 decrease in monthly OOP costs increases the months' supplied of the drug by 0.00141. Thus, a $1 increase in the subsidy increases the months supplied of the drug by $0.01553 * 0.00141 = 0.000022$ months. On average, the demand-weighted mean plan outlay for the drug demand used in Appendix Table A.1 is $63 per month supply. Thus, the estimated drug demand response would increase plan outlays by 0.000022 months $*$ $63 per month $*$ 12 months =$0.017 per year. Using instead the IV estimates in the third column of Appendix Table A.1, the increase in plan outlays is $0.013.

This demand response would not significantly change our pass-through estimate. If we take $0.40 as our focal pass-through estimate from Table 1, accommodating this drug demand response would mean that we acknowledge that for each $1 increase in diagnostic subsidy, $0.017 is spent by the firm on increased drug demand. Of the remaining $0.983, $0.40 is passed-through to lower OOP costs, for an updated pass-through rate of ($0.40/$0.983 =) $0.407.

References

Abaluck, Jason T. and Jonathan Gruber, "Choice Inconsistencies Among the Elderly: Evidence from Plan Choice in the Medicare Part D Program," *American Economic Review*, 2011, *101* (4), 1180–1210.

Agarwal, Sumit, Souphala Chomsisengphet, Neale Mahoney, and Johannes Stroebel, "Regulating Consumer Financial Products: Evidence from Credit Cards," *The Quarterly Journal of Economics*, 11 2014, *130* (1), 111–164.

Bijlsma, Michiel, Jan Boone, and Gijsbert Zwart, "Competition Leverage: How the Demand Side Affects Optimal Risk Adjustment," *RAND Journal of Economics*, 2014, *45* (4), 792–815.

Brown, Jason, Mark Duggan, Ilyana Kuziemko, and William Woolston, "How Does Risk Selection Respond to Risk Adjustment? New Evidence from the Medicare Advantage Program," *American Economic Review*, 2014, *104* (10), 3335–64.

Cabral, Marika, Michael Geruso, and Neale Mahoney, "Does Privatized Medicare Benefit Patients or Producers? Evidence from the Benefits Improvement and Protection Act," *American Economic Review*, Forthcoming 2018.

Carey, Colleen, "Technological Change and Risk Adjustment: Benefit Design Incentives in Medicare Part D," *American Economic Journal: Economic Policy*, February 2017, *9* (1), 38–73.

Conlon, Christopher T. and Nirupama L. Rao, "Discrete Prices and the Incidence and Efficiency of Excise Taxes," *American Economic Journal: Economic Policy*, 2020, *forthcoming.*

Dieguez, Gabriela, Maggie Alston, and Samantha Tomicki, "A Primer on Prescription Drug Rebates: Insights into Why Rebates Are a Target for Reducing Prices," May 2018. Available at `http://www.milliman.com/uploadedFiles/insight/2018/Prescription-drug-rebates.pdf`.

Duggan, Mark, Amanda Starc, and Boris Vabson, "Who Benefits when the Government Pays More? Pass-Through in the Medicare Advantage Program," *Journal of Public Economics*, September 2016, *141*, 50–67.

Einav, Liran, Amy Finkelstein, and Maria Polyakova, "Private Provision of Social Insurance: Drug-Specific Price Elasticities and Cost Sharing in Medicare Part D," *American Economic Journal: Economic Policy*, August 2018, *10* (3), 122–53.

_ , _ , and Paul Schrimpf, "The Response of Drug Expenditures to Non-Linear Contract Design: Evidence from Medicare Part D," *The Quarterly Journal of Economics*, 2015, *130* (2), 841–899.

_ , _ , **Raymond Kluender, and Paul Schrimpf**, "Beyond Statistics: The Economic Content of Risk Scores," *American Economic Journal: Applied Economics*, April 2016, *8* (2), 195–224.

Fabinger, Michal and E. Glen Weyl, "A Tractable Approach to Pass-Through Patterns," March 2015.

Fabra, Natalia and Mar Reguant, "Pass-Through of Emissions Costs in Electricity Markets," *American Economic Review*, September 2014, *104* (9), 2872–99.

Frank, Richard G. and Thomas G. McGuire, "Regulated Medicare Advantage And Marketplace Individual Health Insurance Markets Rely On Insurer Competition," *Health Affairs*, 2017, *36* (9), 1578–1584.

Geruso, Michael, Timothy J. Layton, and Daniel Prinz, "Screening in Contract Design: Evidence from the ACA Health Insurance Exchanges," *American Economic Journal: Economic Policy*, Forthcoming 2018.

Goldman, Dana E., Geoffrey Joyce, Pinar Karaca-Mandic, and Neeraj Sood, "Adverse Selection in Retiree Prescription Drug Plans," *Forum for Health Economics and Policy*, 2006, *9* (1).

Heidhues, Paul, Botond Koszegi, and Takeshi Murooka, "Exploitative Innovation," *American Economic Journal: Microeconomics*, February 2016, *8* (1), 1–23.

Heiss, Florian, Adam Leive, Daniel McFadden, and Joachim Winter, "Plan Selection in Medicare Part D: Evidence from Administrative Data," *Journal of Health Economics*, December 2013, *32*, 1325–1344.

Hsu, John, Jie Huang, Vicki Fung, Mary Price, Richard Brand, Rita Hui, Bruce Fireman, William Dow, John Bertko, and Joseph P. Newhouse, "Distributing $800 Billion: An Early Assessment of Medicare Part D Risk Adjustment," *Health Affairs*, 2009, *28* (1), 215–225.

Jung, Kyoungrae, Roger Feldman, and A. Marshall McBean, "Demand for Prescription Drugs Under Nonlinear Pricing in Medicare Part D," *International Journal of Health Care Finance and Economics*, 2014, *14*, 19–40.

Kautter, John, Melvin Ingber, Gregory C. Pope, and Sara Freeman, "Improvements in Medicare Part D Risk Adjustment: Beneficiary Access and Payment Accuracy," *Medical Care*, 2012, *50* (12), 1102–1108.

Layton, Timothy, "Imperfect Risk Adjustment, Risk Preferences, and Sorting in Competitive Health Insurance Markets," *Journal of Health Economics*, December 2017, pp. 259–280.

Layton, Timothy J., Thomas G. McGuire, and Richard C. van Kleef, "Deriving risk adjustment payment weights to maximize efficiency of health insurance markets," *Journal of Health Economics*, September 2018, *61*, 93–110.

Lenzner, Robert, "ObamaCare Enriches Only The Health Insurance Giants and Their Shareholders," *Forbes.com*, 2013.

Levy, Helen and David Weir, "Take-up of Medicare Part D: Results from the Health and Retirement Study," *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 2010, *65* (4), 492–501.

MacKay, Alexander and Marc Remer, "An Empirical Model of Consumer Affiliation and Dynamic Price Competition," January 2018. Available at `alexandermackay.org`.

_ , Nathan Miller, Marc Remer, and Gloria Sheu, "Bias in reduced-form estimates of pass-through," *Economics Letters*, 05 2014, *123*, 200202.

Marshall, Alfred, *Principles of Economics*, New York: Macmillan, 1890.

Miller, Nathan H., Matthew Osborne, and Gloria Sheu, "Pass-through in a concentrated industry: empirical evidence and regulatory implications," *The RAND Journal of Economics*, 2017, *48* (1), 69–93.

Muehlegger, Erich and Richard L. Sweeney, "Pass-Through of Input Cost Shocks Under Imperfect Competition: Evidence from the U.S. Fracking Boom," November 2017. NBER Working Paper 24025.

Newhouse, Joseph P., Erica Seiguer, and Richard G. Frank, "Was Part D a Giveaway to the Pharmaceutical Industry?," *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*, 2007, *44* (1), 15–25.

_ , J. Michael McWilliams, Mary Price, Jie Huang, Bruce Fireman, and John Hsu, "Do Medicare Advantage Plans Select Enrollees in Higher Margin Clinical Categories?," *Journal of Health Economics*, 2013, *32*, 1278–1288.

Office of the Inspector General, "Increases in Reimbursement for Brand-Name Drugs in Part D," June 2018. Available at `https://oig.hhs.gov/oei/reports/oei-03-15-00080.pdf`.

Pauly, Mark V. and Yuhui Zeng, "Adverse Selection and the Challenges to Stand-Alone Prescription Drug Insurance," in David M. Cutler and Alan M. Garber, eds., *Frontiers in Health Policy Research*, Vol. 7, NBER Books, 2004, pp. 55–74.

Polyakova, Maria, "Regulation of Insurance With Adverse Selection and Switching Costs," *American Economic Journal: Applied Economics*, July 2016, *8* (3), 165–195.

Robst, John, Jesse Levy, and Melvin Ingber, "Diagnosis-Based Risk Adjustment for Medicare Prescription Drug Plan Payments," *Health Care Financing Review*, 2007, *28* (4), 15–30.

Rothman, Dov, "A NOTE ON THE ECONOMICS OF PASS-THROUGH WITH TWO-PART TARIFF PRICING," *Journal of Competition Law & Economics*, 05 2015, *11* (2), 401–408.

Solon, Gary, Steven J. Haider, and Jeffrey Wooldridge, "What Are We Weighting For?," *Journal of Human Resources*, 2015, *50* (2), 301–316.

Weyl, E. Glen and Michal Fabinger, "Pass-Through as an Economic Tool: Principles of Incidence under Imperfect Competition," *Journal of Political Economy*, 2013, *121* (3).

Yin, Wesley, Anirban Basu, James X. Zhang, Atonu Rabbani, David O. Meltzer, and G. Caleb Alexander, "The Effect of the Medicare Part D Prescription Benefit on Drug Utilization and Expenditures," *Annals of Internal Medicine*, 2008, *148* (3).

A Appendix: For Online Publication

A.1 Pass-through to Premium vs. OOP costs

In Section 3.2, we compared a plan's profits if it passed-through an ε increase in the subsidy for diagnosis 1 to OOP costs for diagnosis 1 to its profits if it passed it through to premium. We assumed that the insurer would reduce the subsidy for diagnosis 1 by $\rho\varepsilon$ and would reduce the premium by $0.5\rho\varepsilon$.

We justify this assumption by considering the pass-through rate we would recover if we applied our estimation model to this setting. Consider our procedure for estimating the pass-through to premium. Our measurement of average subsidy increase (ΔS_j) for this plan would find an increase of 0.5ε. If the plan passes-through the subsidy to c_{1p}, there is no change in premium, and we would (correctly) estimate zero pass-through to premium. If the plan passes-through the subsidy to premium, we would (correctly) estimate pass-through to premium equal to ρ.

Scenario	pass-through to c_{1p}	pass-through to premium
"Data" for estimating pass-through to premium		
plan's average subsidy increase	0.5ε	0.5ε
change in premium	0	$0.5\rho\varepsilon$
estimated pass-through to premium	0	ρ
"Data" for estimating pass-through to OOP costs		
subsidy increase for diagnosis 1	ε	ε
change in c_{1p}	$\rho\varepsilon$	0
estimated pass-through to OOP costs	ρ	0

If we instead hypothesize that the insurer would reduce premium by $\rho\varepsilon$, it holds without additional assumptions that the insurer is better off passing through to c_{1p}.

$$\Pi^{\text{post,pass-through to } c_{1p}} > \Pi^{\text{post, pass-through to prem}}$$

$$D(m_p - \rho\varepsilon)[\phi + \varepsilon - \rho\varepsilon] + D(m_p)\phi > D(m_p - \rho\varepsilon)[\phi + \varepsilon - \rho\varepsilon] + D(m_p - \rho\varepsilon)[\phi - \rho\varepsilon]$$

$$D(m_p)\phi - D(m_p - \rho\varepsilon)[\phi - \rho\varepsilon] > 0$$

We know this inequality holds because, at baseline, m_p was optimal for the insurer, and thus higher profit than reducing the total price by $\rho\varepsilon$.

A.2 Reduced Salience of OOP costs

Adapting the models of Agarwal et al. (2014) and Heidhues et al. (2016), we can allow one of the variables to be less salient in demand by a factor of $\psi \in [0, 1]$. As in Section 3.2.1 we consider a baseline scenario where there are two known, mutually exclusive diagnoses with the same demand function, subsidy r, and drug prices κ, in which case insurers set the same OOP costs for both diagnoses c_p. To represent that insurers potentially choose differently when OOP costs are less salient, let ϕ' represent the baseline per-person profits.

$$\Pi^{\text{pre}} = 2D(\text{prem}_p + \psi c_p)\phi'$$

We again consider an increase in the subsidy for diagnosis 1 of ε.

$$\Pi^{\text{post,pass-through to } c_{1p}} > \Pi^{\text{post, pass-through to prem}}$$

$$\overbrace{D(\text{prem}_p + \psi c_p - \rho\varepsilon)[\phi' + \varepsilon - \rho\varepsilon]}^{\Pi \text{ for diag 1 if pass-through to OOP}} + \overbrace{D(\text{prem}_p + \psi c_p)\phi'}^{\Pi \text{ for diag 2 if pass-through to OOP}}$$

$$> \overbrace{D(\text{prem}_p - .5\rho\varepsilon + \psi c_p)[\phi' + \varepsilon - .5\rho\varepsilon]}^{\Pi \text{ for diag 1 if pass-through to premium}} + \overbrace{D(\text{prem}_p - .5\rho\varepsilon + \psi c_p)[\phi' - .5\rho\varepsilon]}^{\Pi \text{ for diag 2 if pass-through to premium}}$$

Rearranging the inequality

$$\overbrace{[D(\text{prem}_p + \psi c_p - \psi\rho\varepsilon) + D(\text{prem}_p + \psi c_p) - 2D(\text{prem}_p - .5\rho\varepsilon + \psi c_p)]}^{+ \text{ by convexity if } \psi=1,\ - \text{ if } \psi=0} \overbrace{\phi'}^{+}$$

$$+ \overbrace{[D(\text{prem}_p + \psi c_p - \psi\rho\varepsilon) - D(\text{prem}_p - .5\rho\varepsilon + \psi c_p)]}^{+ \text{ if } \psi>0.5}[\varepsilon - \rho\varepsilon] > 0$$

If OOP costs are less salient than premium, it is possibly optimal for plans to pass-through to premium instead of OOP costs.

A.3 2012 Risk Adjustment Recalibration

The Affordable Care Act gradually increased Part D coverage by raising the actuarial value of coverage in the doughnut hole in 2010 from 0% in 2010 to 75% in 2020. The increase is partially generated by a requirement from drug firms to reduce the cost of branded drugs purchased in the donut hole by 50%; the remainder is generated by increased insurer liability (i.e., reduced patient cost-sharing). In 2011 and 2012, insurers were required to pay 7% and 14% respectively for generic drugs purchased in the donut hole. In 2011 risk adjustment revision, the increase to insurer liability in the donut hole was incorporated into the dependent variable prior to estimating Equation 1. However, since the increase in insurer liability would affect expensive and cheap diagnoses differently, CMS decided to recalibrate risk adjustment weights again

in 2012. They used the same set of diagnoses and the same calibration data (2008), but further altered the dependent variable to accommodate the incremental increase in liability for generic drugs.

The 2012 recalibration was quite modest compared to the 2011 revision; risk adjustment weights were often the same up to the thousandth place between the two years. To see the similarity of the 2011 and 2012 systems, we repeat the analysis of Section 4 to determine the change in subsidies between the 2011 and 2012 risk adjustment systems for the same fixed population (764,621 individuals enrolled in Part D in 2011 and fee-for-service Medicare in 2010). Appendix Figure A.1 mimics the histograms in Figure 1, reporting the difference in total subsidy, demographic subsidy, and diagnostic subsidy between 2011 and 2012 using the same x-axis scale and bin sizes. More than 95% of the 2011-2012 changes are within $50 of zero, whereas for the 2010-2011 revision only 10% of total and diagnostic changes are within $50 of zero, and 24% of demographic changes.

It is easy to incorporate the 2012 recalibration into the right-hand side variables when $t = 2012$ for our key equations ($\overline{S_{jt}}$ in Equation 3a and S_{xt} in Equation 4a). However, the 2012 recalibration is so minor that results[30] are virtually unchanged. Because the event study implementation of our key results (Equations 3b and 4b) does not easily accommodate the 2012 recalibration, and because it is so small relative to the 2011 revision, we do not incorporate the 2012 recalibration into our main specification.

A.4 Measuring Typical Annual Out-of-Pocket Costs: Detail

A.4.1 Associating Drugs and Diagnoses

While drugs are relatively closely linked to diagnoses, there is no reference work we can consult that tells us which drugs treat which diagnoses. Instead, we take advantage of our large claims datasets to estimate the empirical association of prescription drug claims and diagnoses derived from matched contemporaneous medical claims. In particular, we run a linear probability model to predict whether an individual takes a given ingredient combination (we abstract from differences in strength and route of administration) using flags for the 84 diagnoses.

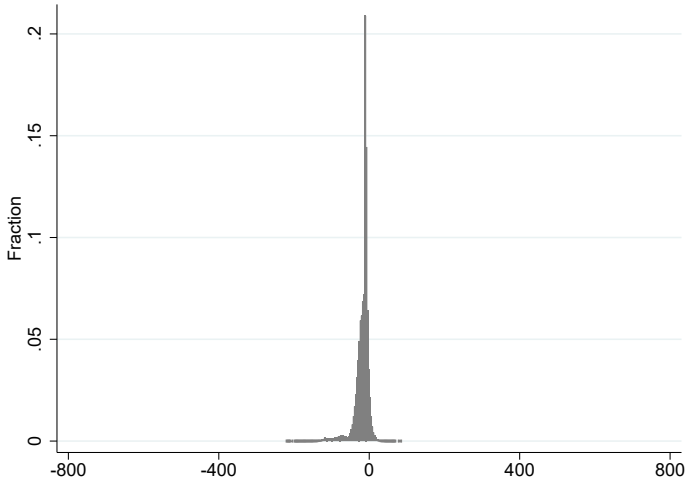
$$\overbrace{T_{ict}}^{\substack{\text{1 if ind } i \\ \text{takes ing combo } c \\ \text{in year } t}} = \sum_x \gamma_{cx} \overbrace{\delta_{ixt}}^{\substack{\text{1 if ind } i \\ \text{has diag } x \\ \text{in year } t}} + \varepsilon_{ict} \tag{5}$$

We estimate these models on the prescription drug and medical claims of Part D enrollees in 2007-2012[31]: more than five million beneficiary × year observations in all. We restrict to 791 ingredient combinations

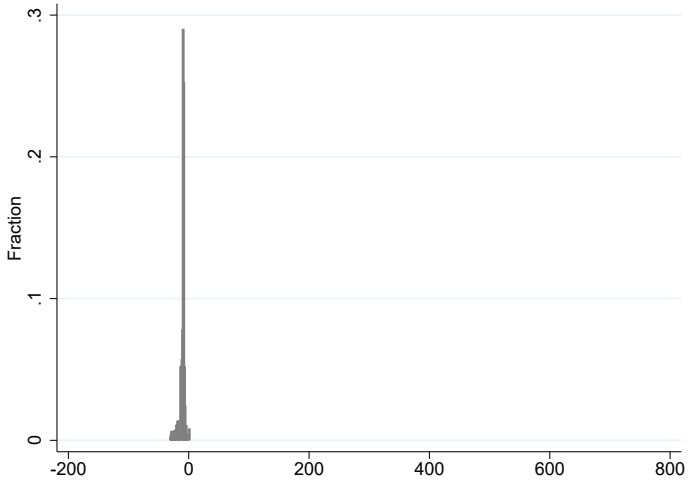
[30] Available upon request.

[31] It is possible that the joint distribution of diagnostic codes and drug utilization adjusts endogenously to subsidy system incentives (e.g., if a diagnosis's subsidy rises, Part D plans increase efforts to ensure providers code it.) When we use only 2007-2010 to associate drugs with diagnoses, we find nearly the same correspondence and very similar empirical results. A downside of using only pre-period years is that drugs introduced in 2011 and 2012 cannot be included in analysis.

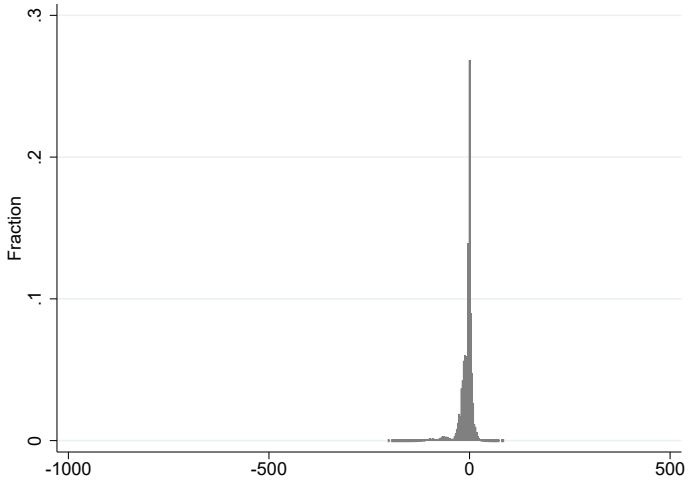
Figure A.1: Distribution of Change in Individuals' Subsidies Induced by 2011 Subsidy Revision



(a) 2011-2012 Change in Total Subsidy



(b) 2011-2012 Change in Demographic Subsidy



(c) 2011-2012 Change in Diagnostic Subsidy

This figure depicts the difference between individuals' subsidies induced by the 2012 recalibration of the 2011 subsidy system. The histogram in Panel (a) reports the difference in total subsidy, ΔS_i, while Panel (b) reports the difference for the demographic component and Panel (c) the diagnostic component. Each histogram displays 100 bins and is bottom- and top-coded to the minimum and maximum x-axis values.

taken by at least 200 beneficiaries in one of the years. Each coefficient γ_{cx} gives the marginal increase in the probability of taking the ingredient combination c associated with having the diagnosis x. We define an ingredient combination as "treating" the diagnosis that most strongly predicts taking it; i.e., ingredient combination c treats the diagnosis x with the largest γ_{cx}. We assign all drugs containing that ingredient combination to set $\mathcal{D}_x$, the set of drugs that treat diagnosis x. On average, the largest coefficient (i.e., the one for the treating diagnosis) exceeds the second largest coefficient by a factor of three. Eight of 84 diagnoses are not found to be "treated" by any ingredient combination we study; these diagnoses tend to be catch-alls (*Other Neurological Conditions, Coagulation Defects and Other Specified Blood Diseases*) or diagnoses, such as *Pelvic Fracture*, where drugs are used for general symptoms such as pain or infection but not for the underlying diagnosis.

A.4.2 Calculating Typical Demand: Detail

Let d index drugs, z index zones of coverage, i index individuals, j index plans, and x index diagnoses. $\mathcal{T}_{xt}$ are the set of individuals taking a drug that treats diagnosis x in year t, and $\mathcal{F}_{xt}$ are the set of individuals with a flag for diagnosis x in the same year. Recall that diagnostic flags are generated by medical encounters, and that individuals may fill prescriptions without regard to the presence of related diagnostic flags. The two sets can diverge for a number of reasons: individuals may be choosing no treatment for a given diagnosis; individuals may have a chronic diagnosis well-controlled by drug therapy for which they did not seek a medical encounter; medical providers, whose payment is independent of the diagnostic flags, may not record them accurately; and the algorithm that assigns drugs to diagnoses in Section A.4.1 may mistakenly assign drugs to a diagnosis with which they are not well connected.

We choose to characterize demand as the total utilization of drugs that treat a given diagnosis, normalized by the number of people who have the diagnosis. The first item is denoted as months_raw_{dzt}:

$$\text{months\_raw}_{dzt} = \sum_{i \in \mathcal{T}_{xt}} \text{months}_{idzt} | d \in \mathcal{D}_x$$

$$\text{months}_{dzt} = \frac{\sum_{i \in \mathcal{T}_{xt}} \text{months}_{idzt} | d \in \mathcal{D}_x}{\sum_{i \in \mathcal{F}_{xt}} 1}$$

In words, months_{dzt} is the months supply of related drugs per person with the diagnosis. The advantage of this definition is that $\sum_z \sum_d \text{months\_raw}_{dz}$ = total demand. This holds because every drug d is in $\mathcal{D}_x$ for some diagnosis x. If we instead summed utilization from only those with the diagnosis, a significant fraction of Part D utilization would be excluded from the measure.

A.5 Pass-Through Rates by Market Concentration

In Section 3, we described theoretical results suggesting that pass-through is lower (closer to zero) in less competitive markets. To test this hypothesis, we compute the Herfindahl-Hirschman index for each PDP region in 2010. We calculate HHI using the market shares of each "contract" operating in the PDP region; parent insurers often sponsor multiple plans within a single contract, so using the contract is a better measure of market power than plans. For the plan × diagnosis analyses measuring pass-through to OOP costs, we also consider an alternative metric that evaluates each diagnosis in each PDP region as an independent market, and compute the HHI using each contract's share of the total number of individuals with each diagnosis in the PDP region. On either measure, the markets are relatively unconcentrated, with a median HHI of 0.1. We split the sample into high-concentration and low-concentration subsamples based on whether the plan or plan × diagnosis is in a market with above or below average concentration. However, not only is concentration low in Part D on average, it is almost uniformly low. For the market-based HHI measure, the 90-10 split ranges from only 0.16 to 0.09. For the diagnosis-market measure, the 90-10 split is similar, ranging from 0.2 to 0.09.

Figure A.2 reports the estimates of the effect of subsidy on premiums and out-of-pocket costs for subsamples defined by market concentration. We find that the pass-through of subsidies is very similar in both subsamples and cannot be statistically distinguished. However, because even the high-concentration markets are relatively unconcentrated, we conclude that we are underpowered to test differences in pass-through by market concentration.

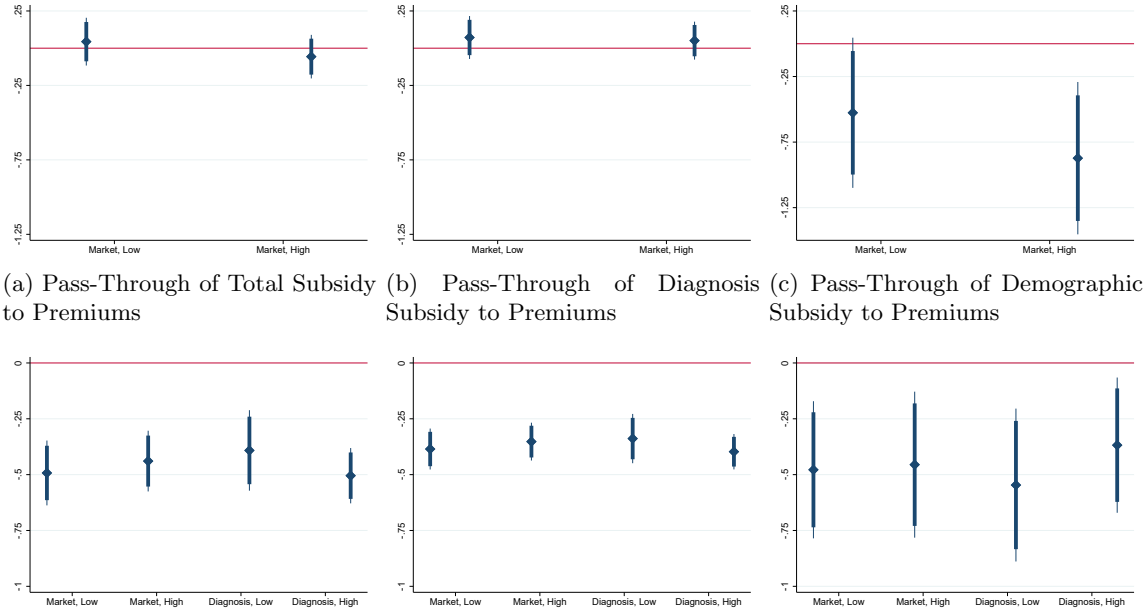
A.6 Recovering Demand Elasticities

To determine the response of drug demand to the subsidy revision, we implement an instrumental variables analysis. The first stage predicts OOP costs for drug d in year t using models similar to Equation 3: using the subsidy for the diagnosis in year t and either a price spline or a linear drug trend. The second stage predicts an individual's demand – months supplied – for drug d in year t. Our panel data is balanced across t for id combinations – i.e., if an individual i takes a particular drug d only in 2010, months_{idt} is imputed to zero for years 2009, 2011, and 2012. A balanced panel includes both the intensive and extensive margins of months supply, but results are similar when we use only the intensive margin.

$$\overbrace{\text{months}_{idt}}^{\substack{\text{demand}\\ \text{for } i\\ \text{for drug } d\\ \text{in year } t}} = b\overbrace{\widehat{\text{OOP cost}_{idt}}}^{\substack{\text{predicted}\\ \text{OOP cost for } d\\ \text{in } t}} + \overbrace{X_{djt}}^{\substack{\text{price}\\ \text{spline}\\ \text{or drug}\\ \text{trend}}} + \text{Fixed Effects} + \varepsilon_{idt} \tag{6}$$

$$\text{OOP cost}_{idt} = \beta S_{xt} + X_{djt} + \text{Fixed Effects} + \varepsilon_{idt}$$

Figure A.2: Heterogeneity by Market Concentration



(a) Pass-Through of Total Subsidy to Premiums

(b) Pass-Through of Diagnosis Subsidy to Premiums

(c) Pass-Through of Demographic Subsidy to Premiums

(d) Pass-Through to OOP Costs, no controls

(e) Pass-Through to OOP Costs, price spline

(f) Pass-Through to OOP Costs, diagnosis trend

The top panel represents the coefficients from estimation of Equation 3a for subsidies (panel (a)), the diagnostic component of subsidies (b), and the demographic component of subsidies (c), split by subsamples defined by the HHI of the market (PDP region), where HHI is defined by market shares at the contract level. The bottom panel represents the coefficients from estimation of Equation 4a, with no extra controls (panel (a)), a price spline (panel (b)), or a diagnosis time trend (panel (c)). Within each panel, the first two coefficients report for markets with low (below-median) or high HHI. The second two coefficients report for markets with low (below-median) or high HHI, where HHI is defined at the diagnosis-market level using the number of individuals with the given diagnosis in the contract.

We consider two sets of fixed effects. The first uses an individual, a drug × plan, and a plan × year fixed effect. The first stage in this specification is more similar to Equation 3. In the second set, we use an individual × drug fixed effect and an individual × year fixed effect. This second set is better suited to capturing individual heterogeneity – both time-invariant taste for a particular drug and overall demand for drugs in a particular year. We two-way cluster ε_{idt} on individuals and drugs.

We estimate this equation on more than 35 million individual × drug × year observations between 2009 and 2012. We drop individuals who receive the low-income subsidy because their out-of-pocket costs are subsidized by the government. Our results are reported in Table **??**. The top panel estimates the relationship between out-of-pocket cost and months supply using ordinary least squares, and the bottom panel reports the full instrumental variables model. As is common, the co-determination of out-of-pocket cost and demand biases our OLS coefficients towards zero. In the IV, the first stage recovers estimates similar to what is implied by a weighted average of the estimates in Figure 8. Our second stage, which is only significant (and based on an significant first stage) when we control for a price spline, implies elasticity estimates of about -2%. Previous research has computed elasticities using the increased out-of-pocket costs at the coverage gap, and has found larger estimates: -30% to -50% in Einav et al. (2015) and -14% to -36% in Jung et al. (2014). Similar to those papers, we are estimating elasticity among those who do not receive the low-income subsidy, meaning those in the top half of the Medicare income and wealth distribution. The benefit design changes we study may induce less of a utilization response because they are less salient for beneficiaries than the large discrete changes in out-of-pocket costs at the coverage gap; in addition, beneficiaries entering the coverage gap may be able to delay purchases for a few weeks or months until the following contract year, a strategy unavailable to beneficiaries in the setting we study.

A.7 Supplementary Figures and Tables

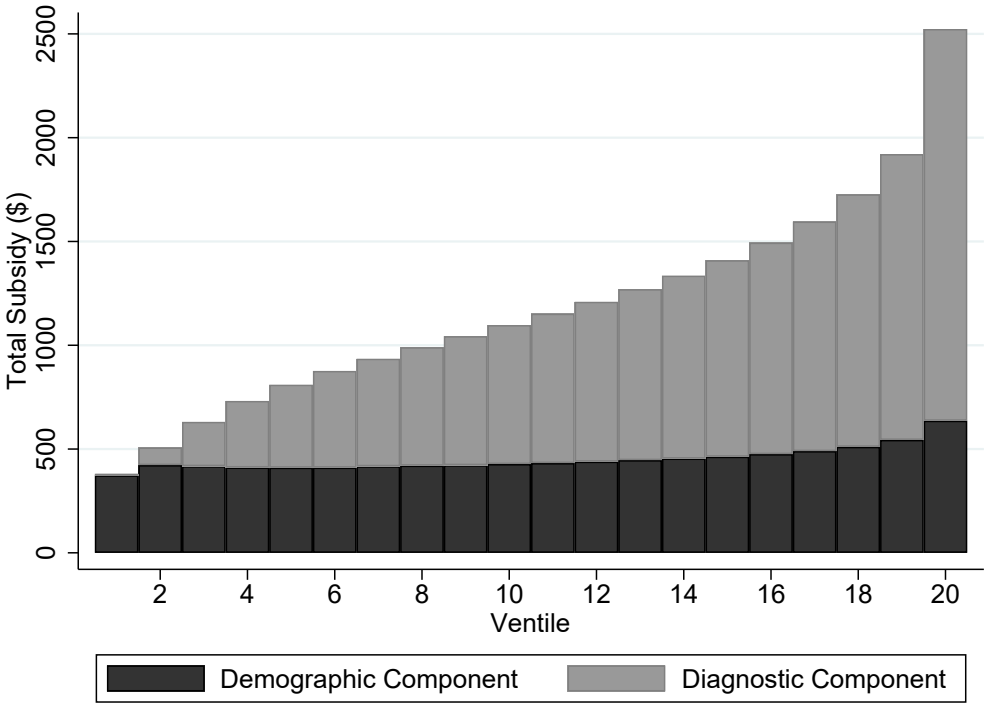
Table A.1: Demand Elasticity: Instrumental Variables for Out-of-Pocket Cost

	OLS			
	Months	Months	Months	Months
OOP Cost ($)	-0.00034+	-0.00060*	-0.00020*	-0.00039**
	(0.00019)	(0.00027)	(0.00010)	(0.00014)
spline for price	X		X	
drug trend		X		X
FEs	Individual Plan X Year Plan X Drug		Individual X Drug Individual X Year	
Implied ε (%)	-0.64	-1.14	-0.38	-0.74

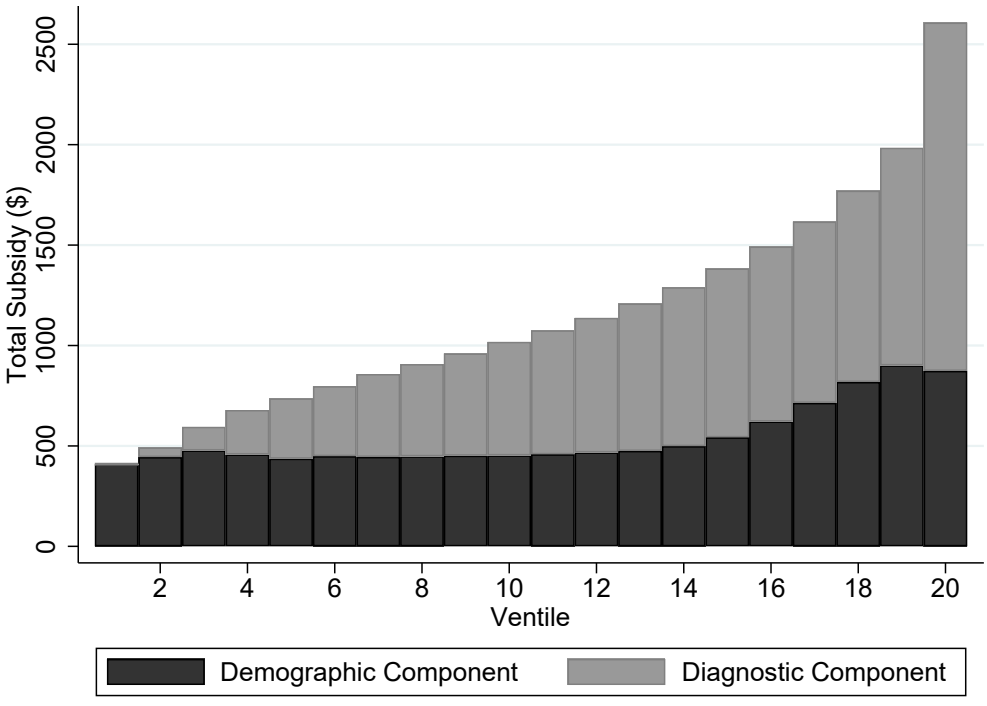
	IV			
First Stage				
	OOP Cost ($)	OOP Cost ($)	OOP Cost ($)	OOP Cost ($)
Subsidy ($)	-0.01553+	-0.00660	-0.01449+	-0.00463
	(0.00882)	(0.00572)	(0.00806)	(0.00444)
Second Stage				
	Months	Months	Months	Months
OOP Cost ($)	-0.00141*	0.08069	-0.00121*	0.07460
	(0.00064)	(0.09674)	(0.00048)	(0.10408)
spline for price	X		X	
drug trend		X		X
FEs	Individual Plan X Year Plan X Drug		Individual X Drug Individual X Year	
Implied ε (%)	-2.67	153.01	-2.29	141.46
N	35,469,096			

This table reports the results of estimating Equation 6 on the months supplied to each individual between 2009 and 2012. The top panel estimates ordinary least squares with the stated controls (price spline or drug trend) and fixed effects. The bottom panel instruments for out-of-pocket cost using subsidy and reports both the first and second stage of estimation. Standard errors are two-way clustered at the individual and drug. +, * and ** represent significance at the 10, 5 and 1 percent levels.

Figure A.3: Distribution of Total Subsidies Under the 2010 and 2011 Risk Adjustment Systems



(a) 2010 Risk Adjustment Subsidies ($S_i(\text{old}, 2011)$)



(b) 2011 Risk Adjustment Subsidies ($S_i(\text{new}, 2011)$)

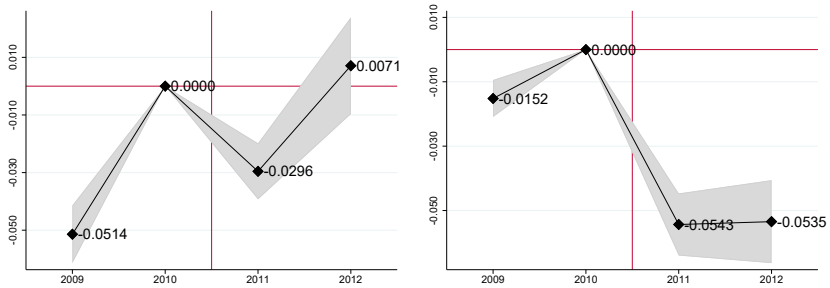
These figures report the average subsidy within ventiles of the distribution for 764,621 individuals enrolled in Part D in 2011 and fee-for-service Medicare in 2010. Panel (a) uses the individuals' 2011 data and the 2010 risk adjustment system while panel (b) uses the same 2011 data and the 2011 risk adjustment system. The demographic component of the subsidy is in black, while the diagnostic component is in gray.

Table A.2: Old Subsidies and Subsidy Update for Each Diagnosis

Diagnosis	Old Subsidy ($)	Subsidy Update ($)	SE
HIV/AIDS	1889	-256	14
Age<65 & Schizophrenia	347	298	3
Multiple Sclerosis	331	316	9
Parkinson's Ds	296	-69	4
Leukemia	271	136	41
Diabetes w/ Comps	239	78	1
Opportunistic Infections	238	-146	8
ADD	235	-5	6
Congestive Heart Failure	232	-47	1
Schizophrenia	231	99	5
Hypertension	205	-25	1
Dementia w/ Depression	204	-229	3
Kidney Transplant	199	93	7
Dsr of Immunity	191	88	8
Rheumatoid Arthritis	183	54	2
Inflamm. Bowel Ds	168	82	4
Esophageal Ds	163	12	1
Metastatic Acute Cancers	161	206	6
Age<65 & Other Major Psych. Dsrs	153	235	2
Lipoid Metabolism	151	47	1
Asthma and COPD	151	83	1
Open-angle Glaucoma	149	41	1
Other Major Psych. Dsr	146	-21	1
Motor Neuron Ds/Atrophy	141	38	15
Psoriatic Arthropathy	139	275	12
Dsr of Spine	130	-84	1
Myocardial Infarction/Unstable Angina	129	23	1
Other Psych.	117	-46	3
Seizure Dsr & Convulsions	117	177	2
Osteoporosis	106	35	1
Severe Hematological Dsr	105	59	5
Migraines	98	160	3
Incontinence	94	-45	2
Heart Arrhythmias	86	-20	1
Polycythemia Vera	85	-38	8
Hepatitis	85	163	6
Other Upper Respiratory Ds	77	-8	1
Muscular Dystrophy	77	-57	16
Major Organ Transplant	73	434	12
Other Endocrine	72	91	2
Psoriasis	71	140	3
Polyneuropathy exc. Diabetic	71	91	2
Other Musculoskeletal	71	-22	1
Inflamm. Spondylopathies	69	143	3
Chronic Renal Failure	68	91	1
Infectious Ds	68	-14	3
Mononeuropathy/Abnormal Movement	66	3	1
Female Stress Incontinence	62	12	3
Connective Tissue Dsr	61	133	3
Cerebral Hemorrhage/Stroke	58	24	1
Vascular Retinopathy exc. Diabetic	52	14	2
Huntington's Ds	51	-23	12
Vertebral Fracture w/o Spinal Injury	51	-45	3
Nephritis	47	36	7
Salivary Gland Ds	46	8	5
Lung Cancer	46	115	1
Other Spec. Endocrine	45	40	1
Chronic Skin Ulcer exc. Decubitus	44	-9	2
Pancreatic Ds	44	16	3
Fecal Incontinence	44	19	6
Cellulitis & Skin Ds	44	-2	1
Quadriplegia	44	23	4
Urinary Obstruction	44	-4	2
Bullous Dermatoses	44	-9	1
"Empyema, Abscess, & Lung Ds"	40	-109	16
Bronchitis & Congenital Lung Dsr	40	29	1
Polymyalgia Rheumatica	40	-8	3
Pneumonias	40	-116	5
Macular Degeneration & Retinal Dsr	37	15	1
Vascular Disease	32	64	1
Ulcer & Gastro Hemorrhage	31	6	2
Vaginal & Cervical Ds	31	63	2
Pulmonary Embolism & Thrombosis	25	42	2
Larynx/Vocal Ds	22	23	8
Impaired Renal Function	21	27	1
Bone Infections	21	22	4

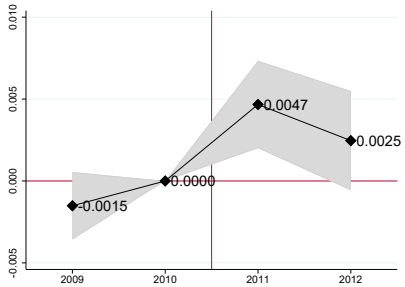
This table reports the diagnosis-specific coefficients from Equation 2 . The second column reports the subsidy for the diagnosis in a plan bidding the national average bid under the 2010 system. The next columns report the subsidy update, the coefficient from 2, and its robust standard error. Only the 76 diagnoses used in later analyses are reported.

Figure A.4: Event Study: Impact of Subsidy Updates Among Brands, Plan × Drug Analyses

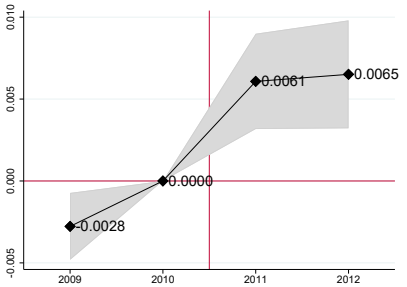


(a) OOP cost ($), no controls

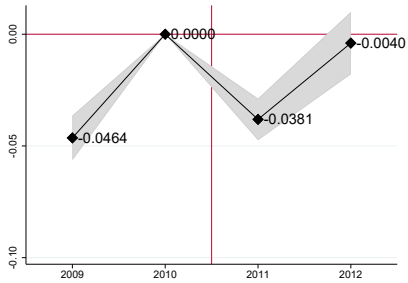
(b) OOP cost ($), price spline



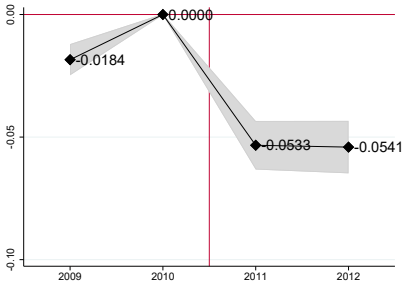
(c) Coverage (p.p.), no controls



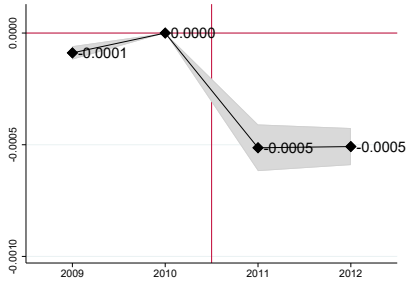
(d) Coverage (p.p.), price spline



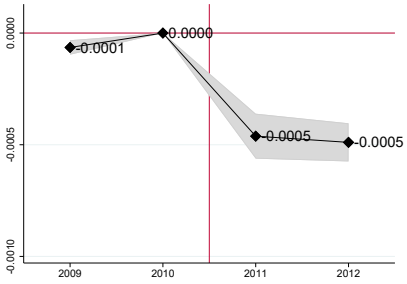
(e) OOP cost | Covered ($), no controls



(f) OOP cost | Covered ($), price spline



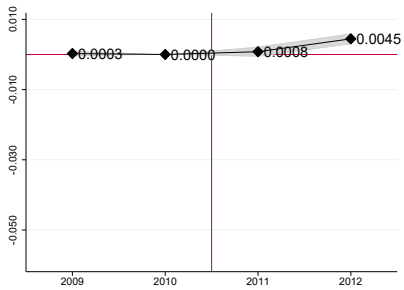
(g) tier, no controls



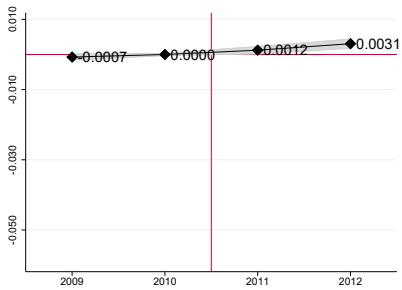
(h) tier, price spline

These figures depict the coefficients from estimation of Equation 4b on the plan × drug panel. Each row represents an outcome; the left hand column includes no controls, while the right hand column controls for a price spline.

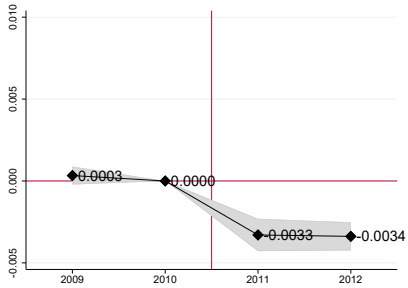
Figure A.5: Event Study: Impact of Subsidy Updates Among Generics, Plan × Drug Analyses



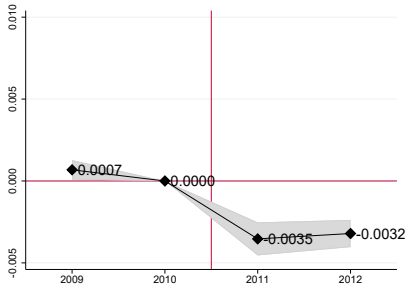
(a) OOP cost ($), no controls



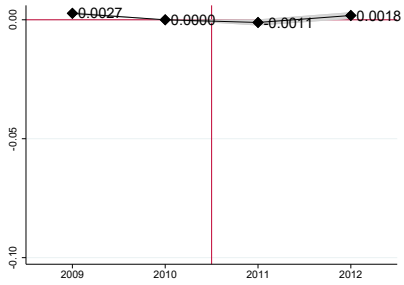
(b) OOP cost ($), price spline



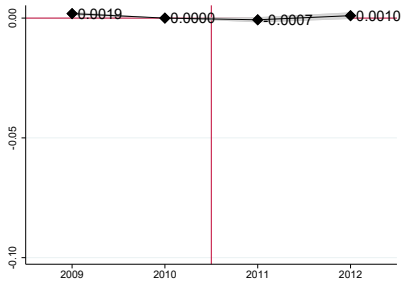
(c) Coverage (p.p.), no controls



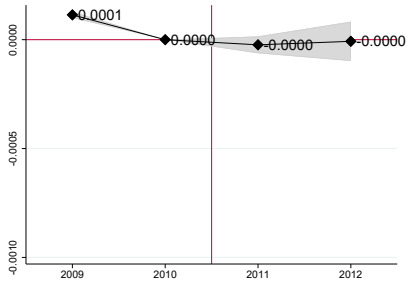
(d) Coverage (p.p.), price spline



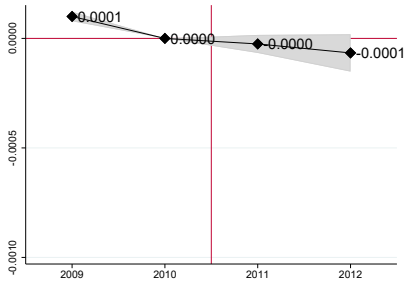
(e) OOP cost | Covered ($), no controls



(f) OOP cost | Covered ($), price spline



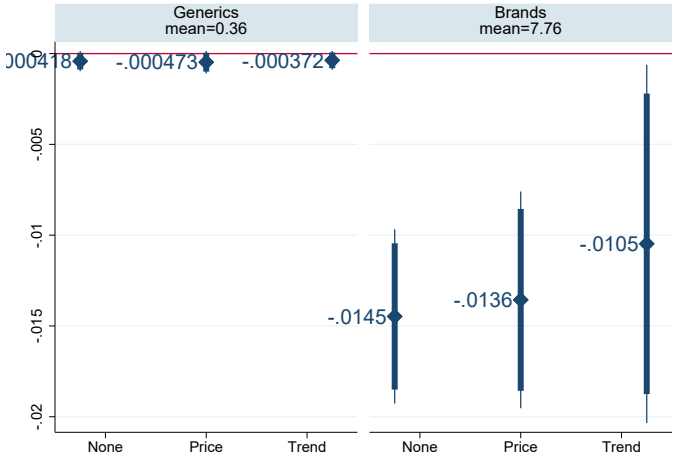
(g) tier, no controls



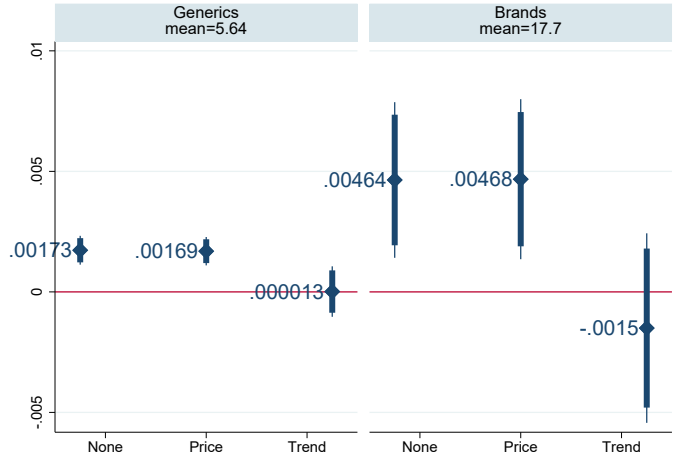
(h) tier, price spline

These figures depict the coefficients from estimation of Equation 4b on the plan × drug panel. Each row represents an outcome; the left-hand column includes no controls, while the right-hand column controls for a price spline.

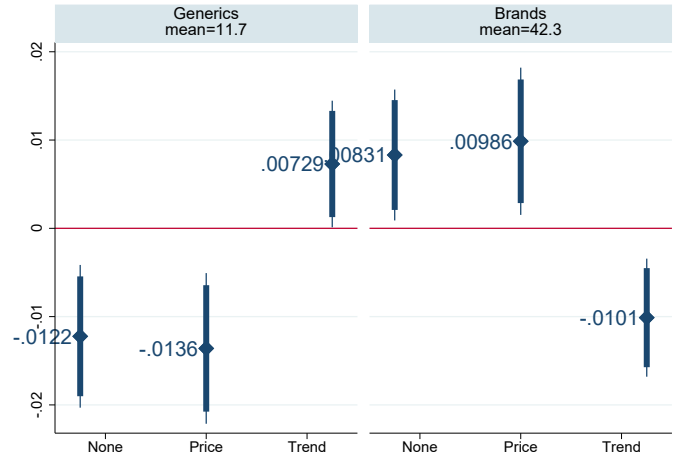
Figure A.6: Impact of Diagnosis-Specific Subsidies on Utilization Management Outcomes by Drug Subsample



(a) Step Therapy



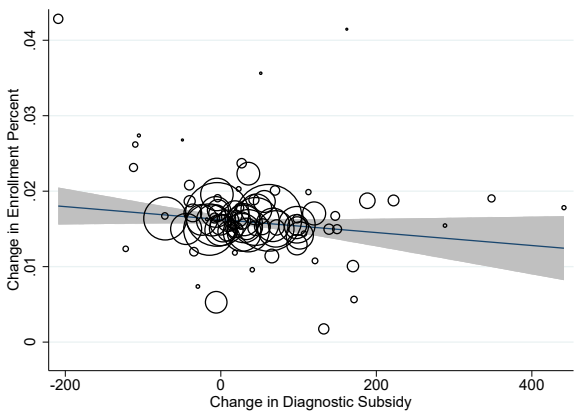
(b) Prior Authorization



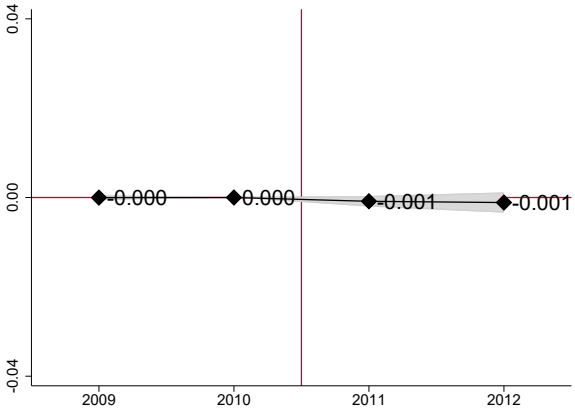
(c) Quantity Limits

This figure represents the β coefficients from Equation 3. Each coefficient represents the effect of a \$1 increase in the subsidy for the diagnosis the drug treats on the drug's utilization management outcomes in a plan $\times$ year. Each panel represents a different utilization management outcome, and the coefficients and mean are expressed as percentage points. Within each panel, the coefficients are estimated separately for generics and brands. Three specifications are reported: no extra controls ("none"), a price spline ("price"), and a drug time trend ("trend").

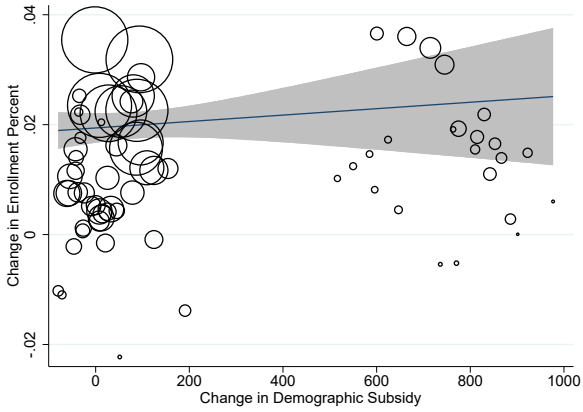
Figure A.7: Response of Enrollment in Part D to Changes in Diagnostic and Demographic Subsidies



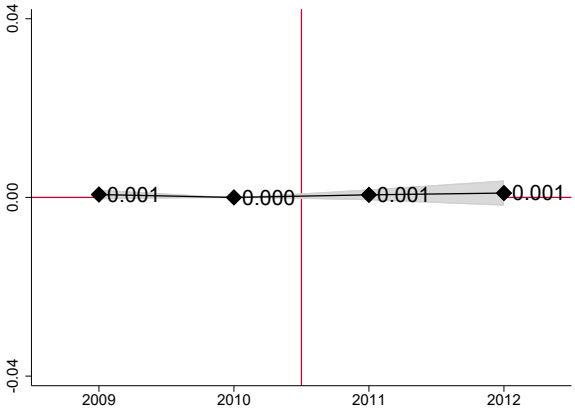
(a) diagnostic subsidy



(b) diagnostic subsidy



(c) demographic subsidy



(d) demographic subsidy

In the left-hand figures, the x-axis measures the change in the subsidy for each diagnosis (panel (a)) or demogaphic bin (panel (c)). The y-axis measures the change in the rate of Part D enrollment among individuals with this diagnosis or in this demographic bin between 2010 and 2011. The size of the marker represents the number of individuals with the diagnosis or in the demographic bin. The right-hand figures represent the estimation of the event study in Equation 4, which predicts enrollment in Part D using the interaction of the change in subsidy for each diagnosis (panel (c)) or demographic bin (panel (d)) with year dummies.