Myopia and Complex Dynamic Incentives: Evidence from Medicare Part D

Christina M. Dalton * Gautam Gowrisankaran† Robert Town ‡

March 30, 2015

Abstract

Standard Medicare Part D drug insurance provides no coverage in a "donut hole" region, making the purchase problem dynamic. We develop a discontinuity based test for myopia using enrollees who arrived near the coverage gap early in the year. We find that there are fewer and cheaper purchases immediately after reaching the gap, providing evidence in favor of myopia. We structurally estimate a dynamic drug purchase model and find complete myopia. For a nationally representative sample, "behavioral hazard" increases enrollee spending by 44%. Entirely eliminating the gap would increase insurer spending 31%, with a 6% increase for generic gap coverage.

Keywords: Medicare Part D, cost sharing, donut hole, regression discontinuity

We have received helpful comments from Jason Abaluck, David Bradford, Juan Esteban Carranza, Pierre Dubois, Liran Einav, Dave Frisvold, Guido Lorenzoni, Carlos Noton, Matthew Perri, Asaf Plan, Mary Schroeder, Marciano Siniscalchi, Ashley Swanson, William Vogt, Glen Weyl, and seminar participants at numerous institutions. We are especially indebted to Doug Mager at Express Scripts for data provision. Nora Becker and Mike Kofoed provided excellent research assistance. Gowrisankaran acknowledge research support from the Center for Management Innovations in Health Care at the University of Arizona.

*University of Georgia
†University of Arizona, HEC Montreal and NBER
‡University of Pennsylvania and NBER

1 Introduction

The introduction of Medicare Part D in 2006 was the most important benefit change to Medicare since its inception. Part D added drug coverage to Medicare's hospital and physician/outpatient benefits. Evidence indicates that Part D lowered Medicare beneficiaries' out-of-pocket costs while increasing prescription drug consumption (Yin et al., 2008; Zhang et al., 2009; Lichtenberg and Sun, 2007; Ketcham and Simon, 2008). As a consequence, the program is popular, enrolling over 35 million Medicare beneficiaries in Medicare sponsored plans and an additional 3.5 million in employer provided, Medicare subsidized plans in 2013 (Medpac, 2014).

Medicare Part D nonetheless has its critics. Perhaps the biggest criticism is concern over Part D's "donut hole" benefit structure. Enrollees in plans with a standard benefit structure face a nonlinear price schedule where out-of-pocket expenditures are modest until their accrued total year-to-date expenditures place them in the coverage gap – the so-called donut hole. Once in the coverage gap, the enrollee pays the full price of all drugs until she spends through the gap into the "catastrophic zone." In 2008, the year of our data, the donut hole cannot begin before $2,510 in total expenditures or end after $4,050 in out-of-pocket expenditures[1] and approximately one-third of Part D enrollees entered the coverage gap.[2]

Importantly, the coverage gap creates a complex dynamic decision problem for enrollees. To optimally decide on a drug purchase, an enrollee must forecast her future expenditures over the remaining portion of the year. While a donut hole is relatively unusual, most health insurance plans have dynamic aspects, in the form of out-of-pocket maximums and deductibles. These dynamic aspects create a complex decision problem where optimal decision-making depends on rationally forecasting future healthcare consumption, a point that has been recognized since at least the RAND Health Insurance Experiment.[3]

Critics of Part D point to the possibility that enrollees may reduce drug consumption upon

[1]Some plans offer partial coverage in the donut hole, though few offer complete coverage (Neuman and Cubanski, 2009).

[2]Source: http://www.medpac.gov/documents/reports/mar11_ch13.pdf?sfvrsn=0.

[3]The Rand Health Insurance Experiment found that utilization increased once enrollees hit their out-of-pocket maxima (Newhouse, 1993). The study did not disentangle the reasons for this increase.

entering the donut hole, with adverse consequences for health (Liu et al., 2011). Concern regarding the coverage gap has also led to legislation that will phase out the donut hole over time, enacted as part of the 2010 Affordable Care Act. However, economic models of forward-looking rational consumers predict that the donut hole, *per se*, should have little real effect on behavior, as what matters most is the expected end-of-year price.

Given this dichotomy, a relevant concern is whether the rational model accurately reflects Part D enrollee behavior. Are the mostly senior-citizen enrollees accurately making the tradeoffs required to dynamically optimize their choice of prescription drugs? This concern is underscored by the fact that a number of studies have found that Part D enrollees act myopically, or at least not fully rationally, when faced with the plan choice part of the Part D decision (see Abaluck and Gruber, 2011, 2013; Abaluck et al., 2015; Ho et al., 2014; Heiss et al., 2010; Schroeder et al., 2014). In contrast, other studies find that Medicare Part D enrollees are at least partly rational in their choices of plans or drug purchases (see Ketcham et al., 2012; Einav et al., 2015; Ketcham et al., 2014).

Dynamic health insurance contracts can lead to "behavioral hazard" (Baicker et al., 2012), where non-rational consumers make suboptimal decisions. A donut hole structure is particularly problematic in the presence of myopia, because it causes individuals to over-consume before the coverage gap, compounding the moral hazard from health insurance, which also causes overconsumption before the gap. In contrast, with myopia, deductibles may counterbalance the effects of moral hazard.

This paper has two goals. First, we develop robust tests for whether Part D enrollees act myopically in their choice of drug treatments. Our tests avoid several selection issues that often makes such inference challenging. Second, we quantify the relative extents of moral and behavioral hazards and empirically evaluate how these hazards affect the outcomes and welfare of policies such as eliminating the coverage gap. We proceed by specifying a dynamic model of drug choice which allows for non-rational, time inconsistent behavior. Using the model, we first provide evidence – using a regression discontinuity – that people do not act rationally in their drug purchase decision. We then structurally estimate the parameters of the model and examine the implications of counterfactual policies regarding the coverage

gap. We believe that both our tests of myopia and our estimation framework may be more broadly useful in analyzing the implications of different insurance schemes.

Our model and estimation strategy focus on enrollees' spending behavior as they transition from the initial coverage region (before \$2,510 in spending) to the coverage gap region (starting at \$2,510). In our model, each week, each Part D enrollee has a distribution of possible health shocks. Each health shock is characterized by a class of drugs that can be used to treat the shock. Given a health shock, the individual makes a multinomial logit choice from among the class of drugs and an outside option of not filling a prescription. Each drug has three characteristics: the (full) price, the out-of-pocket price, and the mean utility from consumption, which we model with fixed effects. The outside option has utility normalized to 0. The model is dynamic because purchasing a drug in the initial coverage region moves an enrollee closer to the coverage gap. We allow for non-rationality in the form of *quasi-hyperbolic discounting* (Laibson, 1997; Phelps and Pollak, 1968; Strotz, 1956): individuals discount their future health shocks in the current week with the factor β, in one week with the factor $\beta\delta$, in two weeks with the factor $\beta\delta^2$, etc. An individual with $\beta < 1$ is partly myopic regarding future actions: she would make different tradeoffs at time t between utility at times $t+1$ and $t+2$ than she would make upon reaching time $t+1$. An individual with either β or δ of 0 does not value any utility in future periods and hence is completely myopic.

We empirically test for myopia using 2008 Medicare Part D administrative claims data from a large pharmacy benefit manager. Our empirical strategy is to identify a set of enrollees who face the same end-of-year prices but different current prices, and examine if they respond to the current prices. We identify such a group: enrollees who arrive *near* the donut hole early in the year. We postulate (and empirically verify) that these individuals will almost certainly reach the donut hole before the end of the year. Without any discounting, these individuals should not change their drug purchase behavior upon crossing into the donut hole, because the Part D insurance is essentially just a fixed subsidy. Rational geometric discounters (with $0 < \delta < 1$ but $\beta = 1$) will consume less inside the donut hole than before only because of the time value of money saved. This implies a near-linear progression of

increases in purchases before the donut hole. Finally, myopic individuals – whom we define to be those with either $\beta < 1$ or $\delta = 0$ – may have more drug purchases before the donut hole than in the donut hole, but their purchases can be flat in the periods before the donut hole.

Using the group of enrollees who arrive near the donut hole early in the year, we find that drug purchases drop significantly and sharply upon reaching the donut hole: mean total drug expenditures falls by 28%; the number of prescriptions falls by 21%; expensive prescriptions with a price of over \$150 fall by 33%; and branded prescriptions fall by 27%. In contrast, there is no evidence of expenditures falling in the periods leading up to the donut hole, with no significant differences between spending in the regions \$2,200 to \$2,399 and \$2,000 to \$2,199. A falsification exercise on enrollees in a plan with a different donut hole reveals no drop in spending upon reaching \$2,510. Thus, our discontinuity-based evidence points to myopic rather than rational individuals.

Having found evidence of myopia, we structurally estimate our dynamic model using the same group of enrollees. The parameters of the structural model are fixed effects for each drug, price elasticity parameters, and the discounting parameters β and δ. We estimate these parameters via maximum likelihood. An estimation challenge is that we do not observe when people have health shocks but consume the outside good, and yet substitution to the outside good in the donut hole is a crucial feature of the data. We develop methods that allow us to integrate over the number of shocks and time periods of the shocks where the individual chooses a drug in closed form, thereby making this estimator computationally tractable.[4]

As with the discontinuity-based evidence, conditional on price elasticity parameters, the discounting parameters β and δ are identified by the difference in purchase patterns between the initial coverage region and coverage gap regions. Our price elasticity parameters are identified by the fact that we have across-drug and across-drug-class variation in prices upon entry into the donut hole. For instance, some drugs will have many substitutes and we observe variation in the availability of substitutes for drugs within their drug class and in the

[4]The unobserved nature of the health shocks also prevents us from using the computationally advantageous conditional choice probability estimators initially developed by Hotz and Miller (1993).

coverage-gap prices relative to the initial coverage region prices, and in substitution based on these attributes.

Our structural estimation splits our sample into subsamples based on their ex-ante health risk score. For each of three subsamples, we find that $\beta = 0$. We find a substantial price response that is lower for more expensive drugs. At $\beta = 0$, δ is not identified. For the smaller subsamples, we cannot reject a low β and high δ. However, we can always reject $\beta = 1$ with low values of δ. The underlying reason for our results is that the drop in prescriptions upon reaching the donut hole is too large to be explained with a positive β, given the level of cross-drug substitution, which identifies the price elasticity. We do not interpret our estimates as implying that individuals are extremely impatient and not forward-looking. Rather, we believe that our estimates imply that Medicare Part D enrollees are not making dynamically sophisticated decisions in their choices of drugs and hence that their behavior is much more myopic than rational models would typically imply.

Using our structural estimates, we examine counterfactuals as to preferences and policies. To increase external validity, we use a linear combination of enrollees in our estimation sample and other enrollees in the same plans, chosen so that the fraction hitting the coverage gap during the year is 33%, the same as the mean fraction who hit the coverage gap in 2008. To examine the extent of behavioral hazard, we examine how prescription purchase behavior would change if individuals were geometric discounters with $\delta = 0.95$ at an annual level. We find that geometric discounting would cause enrollees to reduce spending by 31%. Much of this reduction in spending would help them avoid the coverage gap implying that, while insurers would spend less, the drop in insurer spending from geometric discounting would be only 4%, with total prescription drug spending dropping by 15%. In contrast, without moral hazard from insurance, total prescription drug spending would drop 22%, implying that behavioral hazard is almost as important as moral hazard in this market.

Our policy counterfactuals regard eliminating the donut hole as mandated by the Affordable Care Act. We find that eliminating the donut hole would increase total spending by 10% and insurer spending by 31%, implying a substantial cost to the government. Coinsurance would have to increase from an average of 25% to 42% to implement a revenue neutral

insurance scheme without the donut hole. Providing donut hole coverage for generic drugs only would increase insurer spending by 6% relative to the baseline and actually decrease total spending very slightly as consumers would substitute to generics.

Our work is most closely related to the works of Einav et al. (2015) and Abaluck et al. (2015). Our theoretical model builds on Einav et al., who test for the presence of dynamic behavior by considering individuals faced with the same current prices but different end-of-year prices. Our tests are complementary, as we test for the presence of myopic behavior by considering individuals faced with different current prices but the same end-of-year prices. Einav et al. also estimates a structural, dynamic model and finds that the weekly discount factor is $\delta = 0.96$, implying an annualized discount factor of $\delta = 0.12$. Our test also allows us to distinguish myopia from a low discount factor. Our structural estimation builds on Einav et al. by developing a modeling framework for drug choices that is more similar to standard dynamic multinomial choice models and by identifying our parameters using actual price data and the same sample of enrollees on whom our tests for myopia are based. Abaluck et al. (2015) use a very different identification strategy of relying on plan stickiness of enrollees following changes in underlying benefit structures. Similar to our findings, their estimates indicate substantial myopia. Finally, our structural model of quasi-hyperbolic discounting builds on Fang and Wang (2013) and Chung et al. (2013).

The paper proceeds as follows. Section 2 provides our model. Section 3 describes our data. Section 4 presents evidence based on the discontinuity near the donut hole. Section 5 describes the econometrics of estimating our model. Section 6 provides results and counterfactuals, and Section 7 concludes.

2 Model

2.1 Overview

We model the decisions of a Medicare Part D enrollee who has reached $2,000 in spending between the end of March and the end of July of a given year, and hence who has arrived near

the donut hole of $2,510. By focusing on this group, we can credibly assume that individuals in our sample know that they will hit the donut hole at some point during the year, which simplifies our tests and estimation.[5] We also assume that people will not leave the donut hole during the year. This assumption also generally holds in our data,[6] but, as we discuss in Section 2.3, the existence of the catastrophic coverage region will only tend to weaken our empirical findings.

We consider the decision-making of an individual enrollee,[7] starting at the first complete week where the individual has reached the $2,000 spending level.[8] At the start of the week, the individual is faced with a distribution of health shocks, each of which might benefit from drug treatment. The number of health shocks is distributed multinomial, with a minimum of 0 and a maximum of $\overline{N}$. Each health shock is *i.i.d.* At any point in time during the week, the individual knows how far she is from the donut hole as well as how many health shocks she has already had in the current week, but does not know her future health shock realizations for the week (or future weeks). The individual can calculate the distribution of future health shock realizations during the week based on the number that have already occurred. Let $Q_n, n = 0, \ldots \overline{N}$ denote the conditional probability of having another health shock given that n have already occurred this week. Note that $Q_{\overline{N}} = 0$.

We now consider an individual who is faced with a health shock. Health shocks can take on different values, $h = 1, \ldots, H$, each corresponding to a particular drug class (e.g., calcium channel blockers). Health shock h occurs with probability P_h. For each health shock $h = 1, \ldots, H$, there are a set of possible prescription drugs that can be used to treat the condition $j = 1, \ldots, J_h$, plus the outside option $j = 0$ which consists of no drug treatment.

The flow utility from any drug treatment j for health shock h is a function of the perceived fixed utility from treatment, γ_{hj}, which is a parameter to estimate; the total price of the drug, p_{hj}; the out-of-pocket price oop_{hj}; and an unobservable component ε_{hj}, which we assume is distributed type 1 extreme value, *i.i.d.* across health shock occurrences and individuals. We

[5] In reality, 96% of the individuals in this group reach the donut hole during the year.
[6] Only 10% of individuals in this group end up in the catastrophic coverage region at the end of the year.
[7] Section 5.1 we will discuss estimation of the model which involves aggregation across individuals.
[8] We define a week starting on Sunday and ending on Saturday.

assume that current, but not future values of ε_{hj} are known to the individual when making her choice decision. For individuals not in the donut hole after the cost of the current drug, the flow utility is $\gamma_{hj} - \alpha(oop_{hj}) + \varepsilon_{hj}$, while for individuals in the donut hole before the cost of the current drug, the flow utility is $\gamma_{hj} - \alpha(p_{hj}) + \varepsilon_{hj}$, where $\alpha(\cdot)$ is the function that maps from price to utility.

In order to flexibly capture the different impact of price on decisions, we allow $\alpha(p)$ to be a linear spline, with cut points at $c_1 \equiv \$20$, $c_2 = \$50$, and $c_3 = \$100$. We let $\overline{\alpha} \equiv (\overline{\alpha}_1, \ldots, \overline{\alpha}_4)$ denote the four parameters of the spline. E.g., if $c_1 < p < c_2$, then $\alpha(p) = \overline{\alpha}_1 c_1 + \overline{\alpha}_2(p - c_1)$. Note that the linear spline nests the simple linear case, which has $\overline{\alpha}_1 = \overline{\alpha}_2 = \overline{\alpha}_3 = \overline{\alpha}_4$.

The outside option costs nothing (so $oop_{h0} = p_{h0} = 0$) and has $\gamma_{h0} = 0$ so that its flow utility is ε_{h0}. It is important to model this option because individuals in or near the donut hole may substitute away from drug purchases.

2.2 Dynamics and Hyperbolic Discounting

The enrollee's decision problem is dynamic because filling a prescription brings the enrollee closer to the donut hole state of no insurance, and filling a more expensive prescription brings the individual relatively closer than filling a cheaper prescription. We develop a dynamic quasi-hyperbolic discounting model of individual decision-making over the choice of drug purchases. While there are potentially several different models of time-inconsistency we could specify, we focus on the hyperbolic model as it is the most common model of time inconsistency in the literature, it is parsimonious, it lends itself to straightforward counterfactuals and welfare analysis, and its parameters are identified in our empirical setting.

At any week t, an individual discounts the utility from treatments of a future health shock in the current week with the factor β and the utility of a health shock at week $t+1$ with the factor $\beta\delta$. More generally, she discounts health shocks at the week $\hat{t} > t$ with the factor $\beta\delta^{\hat{t}-t}$. A value of $\beta = 1$ would fit into the standard geometric discounting model. With either β or δ substantially less than 1, we would expect the individuals to fill fewer prescriptions and cheaper ones as they approach the donut hole.

In our model, an agent at each time period solves for optimal decisions knowing that her future self will value different time periods differently than in the present. Thus, for example, the agent knows that she values a dollar at time $t+1$ equivalently to $1/\delta$ dollars at time $t+2$, but that, upon reaching time $t+1$, her future self will value a time $t+1$ dollar equivalently to a smaller value of $1/\beta\delta$ at time $t+2$. The agent would like to tie her future self to not spend too much, but cannot do that in our model. The agent thus essentially plays a game with future versions of herself, optimizing at each point in time taking as given the actions of her future selves, and with the knowledge that her future selves are choosing actions that are different than she would currently like them to do. In the hyperbolic discounting literature, this payoff structure is referred to as "sophisticates" (Baicker et al., 2012), as individuals accurately perceive that they are hyperbolic discounters.[9]

Following Fang and Wang (2013), we formalize the behavior of sophisticates using Bellman equations. We first exposit the state space and use it to define the Bellman equation and choice decision problem under quasi-hyperbolic discounting. Let m indicate the monetary distance of the individual from the donut hole at the start of a given purchase decision. Since the donut hole occurs at \$2,510, if the individual has already spent \$2,350 then $m = 160$. Let n denote the number of health shocks that the individual has already incurred this week. Note that $n \in \{0, 1, \ldots, \overline{N}\}$. The state space prior to the realization of a particular health shock can be written as (m, n). Importantly, because of our assumption that the donut hole is an absorbing state, we can write our dynamic decision problem as an infinite-horizon problem, rather than a 52-week problem.[10]

Let $V^{\delta}(m, n)$ denote the value of future payoffs *gross of the hyperbolic β term*. This value pertains to an individual with m dollars till the donut hole who has already experienced n health shocks this week, and before it is known whether another health shock will occur in the

[9]We could also model "naifs" who (incorrectly) perceive that they are geometric discounters with discount factor $\beta\delta$. For naifs, behavior is governed by $\beta\delta$ and hence one cannot separately identify β from δ. Tests for naiveté typically consider whether individuals are willing to sign binding contracts to commit future versions of themselves (DellaVigna, 2009). Such contracts do not exist in our data.

[10]Our counterfactuals specify the Bellman equation as a 52-week problem with a catastrophic coverage region because the assumption that the donut hole is the absorbing state is less accurate at different points in the year and with counterfactual policies and preferences.

week or the nature of that shock. Let $V(m, n, \vec{\varepsilon}, h)$ be the value at the time of realization of a health shock, net of the β term, where $\vec{\varepsilon}$ refers to the vector $(\varepsilon_{h0}, \ldots, \varepsilon_{hJ_h})$. It is necessary to define V and V^δ separately because of the presence of quasi-hyperbolic discounting.

To exposit the agent's optimal choice, we first define her effective price as:

$$p^{ef}(m, p_{hj}, oop_{hj}) = \begin{cases} p_{hj}, \text{ if } 0 \leq m < oop_{hj} \\ oop_{hj} + p_{hj} - m, \text{ if } oop_{hj} < m \leq p_{hj} \\ oop_{hj}, \text{ if } m > p_{hj}. \end{cases} \tag{1}$$

Note that p^{ef} covers the two cases mentioned above, of being in the donut hole and being out of the donut hole, as well as the intermediate case where the drug purchase would move the individual into the donut hole, where the individual pays the out-of-pocket price and the difference left after reaching the donut hole, if this is less than the full price. We now exposit the value function V as:

$$V(m, n, h, \vec{\varepsilon}) = \max_{j=0,\ldots J_h} \left\{ \gamma_{hj} - \alpha(p^{ef}(m, p_{hj}, oop_{hj})) + \beta V^\delta \left(\max\{m - p_{hj}, 0\}, n + 1 \right) + \varepsilon_{hj} \right\}. \tag{2}$$

From (2), the main dynamic effect is that each purchase moves the individual closer to the donut hole by p_{hj} dollars. Given quasi-hyperbolic discounting, the V^δ value function must be further discounted by β in the current decision problem. Also, note that the problem is only dynamic before the donut hole: when the individual is in the donut hole, $m = 0$ and the continuation value is the same across options.

Using (2) we can define market shares using the standard logit formulas as:

$$s(m, n, h, j) = \frac{\exp(\gamma_{hj} - \alpha(p^{ef}(m, p_{hj}, oop_{hj})) + \beta V^\delta(\max\{m - p_{hj}, 0\}, n + 1))}{\sum_{k=0}^{J_h} \exp\left(\gamma_{hk} - \alpha(p^{ef}(m, p_{hk}, oop_{hk})) + \beta V^\delta(\max\{m - p_{hk}, 0\}, n + 1)\right)}. \tag{3}$$

Finally, we define V^δ as:

$$V^\delta(m,n) = (1-Q_n)\delta V^\delta(m,0) + Q_n \sum_{h=1}^{H} P_h \sum_{j=0}^{J_h} s(m,n,h,j)\times$$
$$\left[\gamma_{hj} - \alpha(p^{ef}(m,p_{hj},oop_{hj})) + V^\delta(\max\{m-p_{hj},0\},n+1) - \ln s(m,n,h,j)\right], \quad (4)$$

where the $-\ln s(m,n,h,j)$ term accounts for the expectation of $\vec{\varepsilon}$ conditional on choice.[11]

From (4), there are two possibilities ex-ante to the health shock realization: either there are no more health shock in this week (which occurs with probability $1-Q_n$ or there are more health shocks, in which case we need to sum over the different drug classes. Importantly, we cannot use the standard logit expression for utility in (4): essentially choices are made with $\beta\delta$ discounting of the future, but then different periods in the future are discounted only by δ from each other. Mathematically, then, it is as though the individual is making "wrong" choices but faced with geometric discounting.

2.3 Testable Implications of the Model

This subsection discusses some properties of our model with regards to enrollee behavior across different discount factors. These properties in turn form the basis for our evidence of myopia in Section 4. For simplicity, we specify that there is only one drug class and one drug in this class, so that $H = 1$ and $J_1 = 1$, and that $p_{11} = \$100.4$ and $oop_{11} = \$25.1$.[12] We do not need to specify the price coefficients, mean utility, or health shock frequency, but we need to assume that they are in the range where individuals will make some purchases from inside the donut hole at the end of the year with probability close to 1.

First, we consider the case where $\beta = \delta = 1$. It is difficult to analyze this case as the value is infinite since there is no discounting. Hence we modify this case slightly to assume that the decision process ends at some finite time sufficiently far in the future that the enrollee will surely be in the donut hole at that point. Since the enrollee knows that she will reach

[11] Note that we exclude Euler's constant from this expression as it does not affect choices.
[12] Our example uses these prices as the donut hole start of \$2,510 is an even multiple of them.

the donut hole, consumption of an extra drug in the insured region will still cost her the full price of the drug. Even though she will only pay the out-of-pocket price at the time of consumption, this consumption will move her total expenditures up by the full price which will then push her last previously-insured drug into the donut hole. In this setup, then, the first 25 drugs consumed are insured and the enrollee receives a 75% per-drug subsidy on them for a total subsidy of \$1,882.5. The choice problem is equivalent to one where the individual receives a fixed subsidy of \$1,882.5 and does not have any drug insurance. This it is then a static multinomial logit problem with the full price. Thus, consumption behavior will not change upon crossing into the donut hole.

Now consider the geometric discounting case, where $\beta = 1$ but where $0 < \delta < 1$. We start by analyzing a time point when the enrollee has already bought 24 drugs, and is deciding on the last insured purchase. Purchasing the drug at this point costs the enrollee the \$25.1 immediately. But, it also moves the next purchase into the donut hole, adding \$75.3 to the cost of the next purchase. Thus, the only increase in mean utility from this 25th purchase relative to the 26th purchase is that the \$75.3 is not paid immediately but instead is paid at a later date. One can write this difference in mean utility as $75.3(1 - \delta^t)$ where t is the expected time until payment of this value. Using a standard annual discount factor of 0.95 implies a discounting of 0.001 over one week, implying that this loss in mean utility is small. Thus, the rational model can only explain sizable variations in purchase before the donut hole with extremely large price elasticities (which would amplify small mean utility differences in choice) or very low values of δ.

Moreover, for the same rational case, consider an individual who has previously bought only 23 drugs and is now at her 24th purchase occasion. As with the 25th purchase occasion, the only increase in mean utility from this 24th purchase relative to the 26th purchase is that the \$75.3 is not paid immediately but instead is paid at a later date. However, it is now paid two purchase dates in the future, not one. Thus, the difference in mean utilities between the 24th and 26th purchase occasions would be $75.3(1 - \delta^{t'})$ where t' is the expected time between the 24th and 26th purchase occasions. If t' is twice t, this implies a mean utility difference that is close to twice as large between the 24th and 26th purchase occasions than

between the 25th and 26th purchase occasions. In other words, if the enrollee purchased more frequently at the 25th purchase occasion than at the 26th purchase occasion then she would purchase even more frequently at the 24th purchase occasion – roughly as much more often as between the 25th to the 26th purchase occasion. Essentially then, geometric discounting predicts that purchase probabilities continue to increase the further away they are from the donut hole.

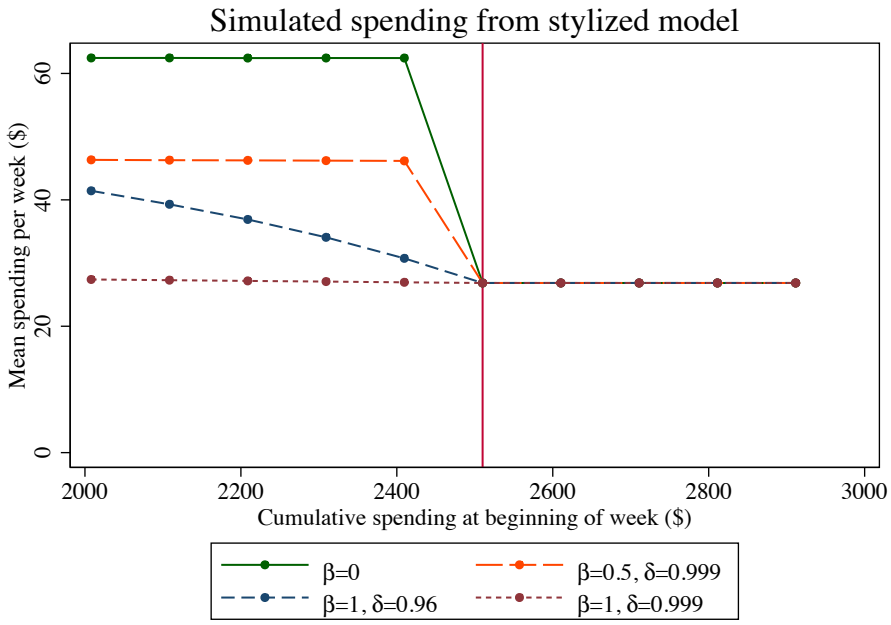
Finally, we consider the myopic case of $\beta < 1$ or $\delta = 0$. Considering again the 25th purchase occasions, the individual discounts the future payoff by $\beta\delta^t$. Thus, the difference in mean utilities between the 25th and 26th purchase occasion is now $75.3(1-\beta\delta^t)$. Similarly, the difference in mean utilities between the 24th and 26th purchase occasion is now $75.3(1-\delta^{t'})$. If β is much below 1 and δ is very close to 1 then these two values will be very close to each other but far from 0, implying similar purchase probabilities at the 24th and 25th purchase occasions that are significantly higher than at the 26th purchase occasion. Note that for the extreme case of $\delta = 0$, individuals react exclusively to the current price, implying that the mean utility for the 24th and 25tth purchase occasions would be exactly the same, and again more than at the 26th purchase occasion. The implication is that myopia can predict that purchase probabilities from before the donut hole are similar to each other but very different from within the donut hole.

Figure 1 shows simulated spending for our different models graphically. To construct this figure, we first solved for the dynamic decisions using the parameters noted above in combination with a linear price term of $\alpha(p_{11}) = 0.02p_{11}$, a mean utility of $\gamma_{11} = 1$,[13] and different values of β and δ (as reported in the figure). We then simulated mean spending in each week enough times so that the simulation variance is negligible and report the mean spending per week given that the individual has spent \$2,008, \$2,108.4, \$2,208.8, etc., at the start of the week.

As predicted by theory, spending is flat and the same across values of β and δ within the donut hole, at about \$27 per week. We find that the $\beta = 0$ case has the biggest mean spending per week before the donut hole, at about \$62 per week. Also as predicted by theory,

[13]These parameters are chosen to roughly calibrate the model to observed spending patterns.

Figure 1: Mean simulated drug spending on cumulative spending for stylized model



spending before the donut hole is completely flat for this case. With $\beta = 0.5$ and $\delta = 0.999$, mean spending is virtually as flat as with $\beta = 0$ before arriving in the donut hole although, at \$46 per week, it is substantially lower than with $\beta = 0$. With geometric discounting and $\delta = 0.999$ (or 0.95 at an annual level), mean spending before and in the donut hole are approximately the same. Finally, with geometric discounting and a low discount factor of $\delta = 0.96$ (as in Einav et al. (2015)), mean spending drops almost linearly over the pre-donut-hole period, from \$41 at \$2,008 in beginning-of-week spending to \$27 at \$2,510 in beginning-of-week spending.

Overall then, we can test for whether $\beta\delta$ is substantially less than 1 by examining whether there is a substantial change in purchase probability between immediately before the coverage gap and immediately after the start of the coverage gap. We can test for whether there is myopia by examining whether periods immediately before the donut hole have a lower purchase probability than periods slightly further away from the donut hole. This regression discontinuity design provides a test of rationality that compares behavior for the same indi-

vidual at different time periods that are near each other in time. By focusing on the same individual and at similar time periods, we avoid many issues of selection that potentially confound other tests of myopia.

These implications of the model are robust to alternative assumptions. First, suppose that there are multiple drugs with different coinsurance rates. In this case, evaluating the dynamic choice problem becomes more involved. As an extreme case, suppose we add a new drug without a copay, so that $oop_{12} = 0$. In this case, the enrollee would like to choose the second drug relatively more often when she is in the initial coverage region. While the analysis is more involved here, the basic point remains: in our model, the only substantial reason to change spending behavior upon entering the donut hole is discounting.

Second, one might consider a very simple alternate scenario, where the value of the drug always lies somewhere between the value of \$25.1 and \$100.4. In this case, the individual would rationally purchase the drug only in the initial coverage region and stop purchasing the drug in the coverage gap. Yet, this scenario does not fit enrollees in our estimation sample since 96% of enrollees in our estimation sample surpass the coverage gap start (and most by a substantial amount), implying that individuals in the sample have valuations for some drugs that are sufficient for purchase even in the absence of insurance. Given that individuals in the sample know that they will end up purchasing once inside the donut hole, curtailing a purchase in the initial coverage region will save them the same amount as curtailing a purchase in the coverage gap region, absent any discounting.[14]

Third, one might believe that our results are due to enrollees simply being misinformed regarding the benefits. Yet, because our data are from the third year of the program, it is unlikely that our results on myopia are driven by a lack of understanding about the presence of the donut hole and its implications. Because Medicare enrollee drug consumption is principally tied to the treatment of chronic conditions, those who reach the donut hole in one year are likely to reach or approach the donut hole in the next year. In addition, as we detail below in Section 3, enrollees are provided substantial information regarding the structure of

[14]Note that in our Figure 1 simulation, the value of the drug is $1 + \varepsilon_{it}$ while the value of $[\$25.1, \$100, 4]$ is $[.052, 2.008]$. Thus, the ε term ensures that the value of the drug is not always between the value of \$25.1 and \$100.4.

the donut hole and receive regular updates on their spending relative to the donut hole.

Finally, we consider the biases of our tests based on the fact that our model omits the presence of the catastrophic coverage region. We will find evidence of myopia if individuals curtail purchases upon reaching the donut hole. One might worry that these results are due to rational behavior where individuals who reach the donut hole due to unexpected health shocks recognize that these health shocks increase the probability that they will reach the catastrophic coverage zone. But, spending should rationally be higher in the catastrophic region. Since a higher probability of reaching the catastrophic region should increase spending, the presence of a catastrophic coverage region will bias us towards a finding of no impact upon entry into the donut hole.

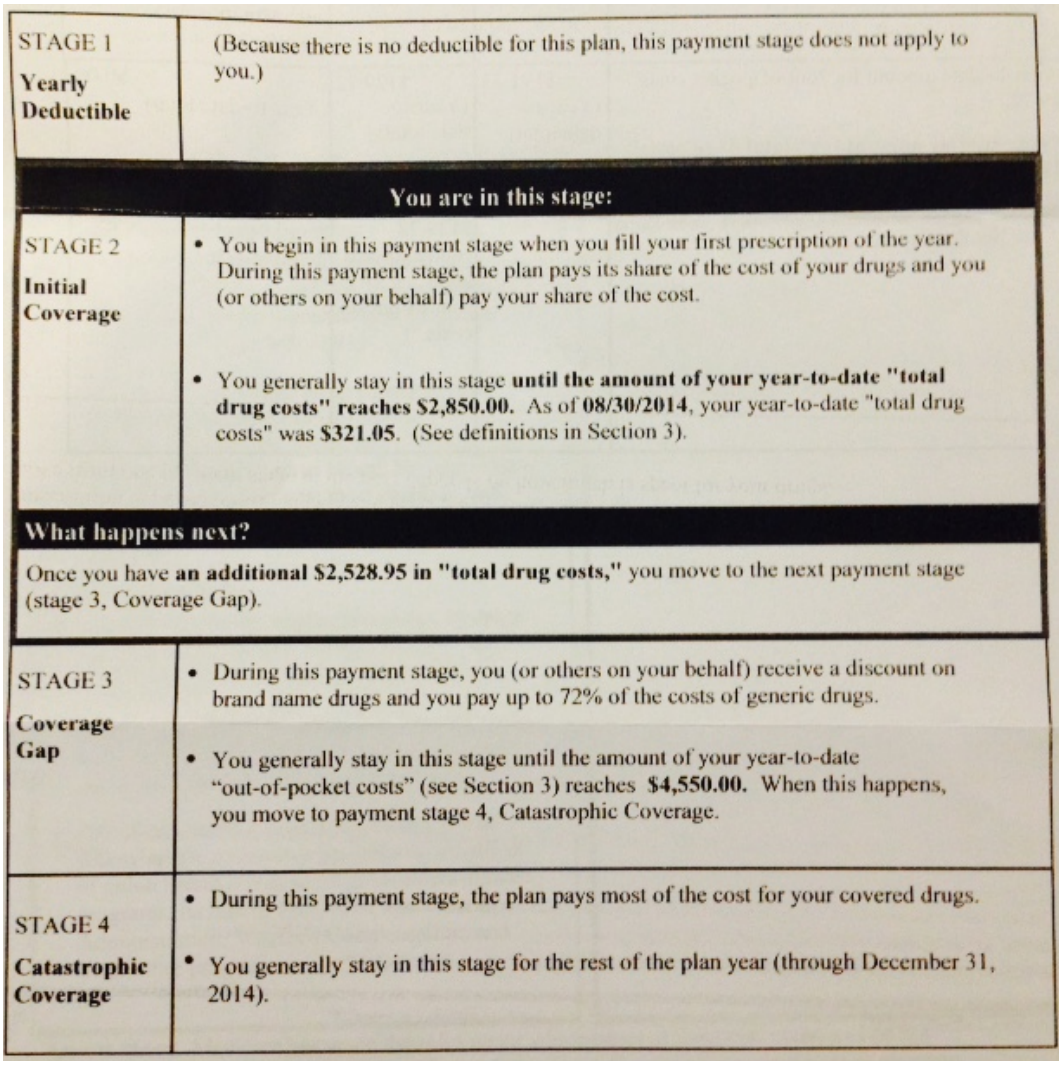
3 Data

For our analysis, we rely on a proprietary claims-level dataset of employer-sponsored Part D plans in the year 2008, the third year of the program. The data come from the pharmacy benefits manger Express Scripts, which managed Medicare Part D benefits for approximately 30 different employer-sponsored Medicare Part D plans with a total of 100,000 enrollees. The plans were offered to eligible employees and retirees as part of their retirement benefits. Employers receive subsidies from Medicare in exchange for providing these plans to their employees. Enrollees in employer-sponsored Part D plans have, on average, higher income than typical Part D enrollees. The employer-sponsored Part D market constituted nearly 7 million enrollees or 15 percent of Part D enrollment in 2008.[15]

An interesting feature of the Part D program is that enrollees are mailed detailed monthly information that lists their out-of-pocket and total costs for the month, the cost of their drugs to the plan as well as the out-of-pocket costs and explains how far they are from the donut hole. Figure 2 shows an example of the part of the information that pertains to the distance to the donut hole. In our view, the frequency and detail of the information suggests that rational enrollees have the opportunity to be informed about the coverage gap.

[15]Source: http://www.medpac.gov/documents/reports/mar09_ch04.pdf?sfvrsn=0, p. 282).

Figure 2: Information provided to Part D enrollees on distance to donut hole



STAGE 1 **Yearly Deductible**	(Because there is no deductible for this plan, this payment stage does not apply to you.)
You are in this stage:	
STAGE 2 **Initial Coverage**	• You begin in this payment stage when you fill your first prescription of the year. During this payment stage, the plan pays its share of the cost of your drugs and you (or others on your behalf) pay your share of the cost. • You generally stay in this stage **until the amount of your year-to-date "total drug costs" reaches $2,850.00.** As of **08/30/2014**, your year-to-date "total drug costs" was **$321.05**. (See definitions in Section 3).
What happens next?	
Once you have **an additional $2,528.95 in "total drug costs,"** you move to the next payment stage (stage 3, Coverage Gap).	
STAGE 3 **Coverage Gap**	• During this payment stage, you (or others on your behalf) receive a discount on brand name drugs and you pay up to 72% of the costs of generic drugs. • You generally stay in this stage until the amount of your year-to-date "out-of-pocket costs" (see Section 3) reaches **$4,550.00.** When this happens, you move to payment stage 4, Catastrophic Coverage.
STAGE 4 **Catastrophic Coverage**	• During this payment stage, the plan pays most of the cost for your covered drugs. • You generally stay in this stage for the rest of the plan year (through December 31, 2014).

The data contain all claims made by an enrollee in the year 2008 for each plan. For each claim, we have plan and patient identifiers, the age (at the fill date) and gender of the patient, the date the prescription was filled, the total price of the drug, the amount paid by the patient, the national drug code (a unique identifier for each drug), the pill name, the drug type (e.g., tablet, cream, etc.), the most common indicator of the drug (e.g., skin conditions, diabetes, infections, etc.), the dispensed quantity of the drug, and an indicator for whether the drug is generic or branded. We keep only individuals who are 65 or older at the time that they fill their first prescription.

Each of the employers offered multiple plans, each with different coverage structures. Our base analysis uses data from five Express Scripts plans. We chose these plans because (1) they have a coverage gap that starts at exactly $2,510 in total expenditures and ends at greater than $4,000 in out-of-pocket expenditures; and (2) the employers that offer these plans allowed us to use their data. We also include falsification evidence from a sixth plan with a higher coverage gap phase initiation.

Table 1: Plan characteristics and enrollment

Plan	CL1A	CL1D	CL1E	CL2A	CL2B	CL3A
Employer	1	1	1	2	2	3
% of employees from employer	26	45	9	79	21	46
Deductible ($)	275	100	100	0	200	0
Donut hole start (total $)	2,510	2,510	2,510	2,510	2,510	4,000
Catastrophic start (out-of-pocket $)	4,050	4,050	4,050	4,010	4,010	4,050
Total enrollment	7,541	12,858	2,431	4,062	1,058	35,395
% hitting $2,510	20	13	16	16	13	20
% hitting catastrophic coverage	2	1	1	1	1	0
Estimation sample:						
Enrollment	672	717	149	326	52	3,341
% hitting $2,510	96	95	96	97	94	98
% hitting catastrophic	12	8	9	10	12	0
Mean total spending ($)	4,543	4,135	4,082	4,232	3,982	4,150
Mean out-of-pocket ($)	2,398	2,038	2,160	2,032	2,068	1,032
Mean age	74	73	72	75	75	78
Percent female	62	57	54	62	56	64
Mean ACG score	1.07	1.18	1.22	0.94	1.14	0.67

Note: plan CL1A provides generic coverage in deductible region; plan CL3A used for falsification exercise only and also provides generic coverage in donut hole.

Table 1 displays the characteristics of the six plans that we consider. The plans represent three different employers; plan and employer identities are masked. We consider all covered employees at employer 2 and the majority of covered employees at employer 1 (with the other employees at this employer choosing plans with different coverage gap regions or some insurance in the coverage gap). Importantly, the fact that each employer offered similar plans to all their employees minimizes the selection issues across plans that one might observe in non-employer-sponsored Part D coverage.

Four of the five plans in our base analysis have a deductible. All deductibles take relatively low values of $275 or less. Each plan features a tiered drug copayment structure, with higher copays for brand and specialty drugs, and reduced copays for the use of mail-order pharmacies. By our inclusion criterion, all five base plans have a donut hole region starting at $2,510. After an enrollee's drug spending surpasses $2,510, all plans drop coverage and the enrollee's expenses are completely out of pocket. The donut hole region is substantial. For three of the plans, the donut hole ends when the enrollee reaches $4,050 in out-of-pocket expenditures (and hence much higher in total expenditures), while the remaining two plans' donut hole ends after the enrollee reaches $4,010 in out-of-pocket expenditures. Generous catastrophic coverage resumes after the donut hole for all six plans, with low copays of $2.15 for generic drugs and $5.60 for branded drugs.

Table 1 also lists summary statistics on plan enrollment. The five base plans cover a total of 27,950 individuals. Our base estimation sample consists of all enrollees who reach the donut hole between Sunday, March 30 and Saturday, July 26, 2008. We chose these dates to be in the part of the year where people should perceive that they will end the year in the donut hole with very high probability. We define a week as starting on a Sunday. Our main estimation is performed on these enrollees starting at the first complete week of the year that they start with $2,000 or more in total expenditures and ending in the last complete week that they start with less than $3,000 in total expenditures.

Our base estimation sample contains 1,916 enrollees distributed across the five plans in our sample. The mean age across the plans ranges from 72-75 and the mean percent female ranges from 54-62%. Between 94 and 97 percent of these enrollees hit the coverage gap during

the year, reflected in mean total spending levels of approximately $4,000 across the plans. The mean percent hitting the much higher catastrophic coverage region ranges from 8 to 12 percent, reflected in mean out-of-pocket spending levels of approximately $2,200 across plans, or about 55 percent of the value necessary to hit the start of catastrophic coverage.

The falsification plan CL3A has the coverage gap start at $4,000 in total spending, a much higher level than for the base plans. Its enrollees are older and disproportionately female relative to the plans in our base analysis sample. It also provides generic drug coverage during its coverage gap. Very few of its enrollees hit the catastrophic coverage region, due to the fact that they will require much higher total spending to reach a given out-of-pocket spending level.

Using our database of claims, we first drop claims for drugs which we believe are not in the formulary. Drugs that are not in the formulary are sometimes reported to the insurance company by the enrollee but do not count towards spending for purposes of determining if the enrollee is in the coverage gap or catastrophic coverage regions. We assume that any claim in the initial coverage region for which the total price is $100 or higher and the out-of-pocket price is the same as the total price reflects a drug that is not in the formulary.[16] We then calculate the dollars until the donut hole (m) for each prescription by tabulating the spending up to this point during the year.[17]

We merge our claims data with data on the expected pharmacy claims cost for each patient, based on their claims from before our sample period. Specifically, we use claims from Jan. 1, 2008 to Mar. 29, 2008 to construct the Johns Hopkins Adjusted Clinical Group (ACG) Version 10.0 score for each enrollee. The ACG score is meant to predict the drug expenditures over the following one-year period. We use the ACG scores to define groups for the structural analysis and then estimate separate coefficients for each group. ACG scores have been widely used to predict future health expenditures in the health economics and health services literature (see, e.g., Handel, 2013; Gowrisankaran et al., 2013). Table 1 also

[16]We also drop one claim with a quantity-filled entry of over 1 million.

[17]There is some ambiguity of the order of claims if there are multiple claims filled on the same date for a given enrollee. For such multiple claims, we assume that the claims are filled in increasing order of out-of-pocket price. For multiple claims for an enrollee on a given date with the same out-of-pocket price, we use the order specified in the database that we received from Express Scripts.

provides the mean ACG scores for each plan in our sample. The scores are normalized so that the population mean score for individuals aged 65 or over is 1. We find that the mean ACG scores are similar to the population mean and each other across our five base plans, and lower for the falsification plan CL3A.

Our analysis classifies each drug into a drug class meant to capture the function of the drug. We define a unique drug class for each pill name, drug type, and most common indicator. We had the drug class coding performed by a research assistant with clinical training using these three fields and the national drug code. We made this classification on the basis of drug function rather than as a function of the underlying disease because we believe that drug function is the relevant attribute for a choice model. Thus, even though both calcium channel blockers and renin-angiotensin system blockers are used to treat hypertension, we treat them as separate classifications because their mechanisms are separate. We lump together drug classes with fewer than 100 prescriptions filled for the estimation sample over the entire year and in a class called "Other." We also lump together drugs within a drug class as "Other" until such point as every drug has at least 50 prescriptions filled over the entire year.

Table 2 lists the drug classes with the most claims in the dataset. Approximately 9 percent of the claims were for cholesterol-lowering (antihyperlipidemic) drugs. The next most common categories include renin-angiotensin system blockers, opioids (for pain relief), and antidepressants. Not shown in the table, there are 70 drug classes when we use our full sample.

Table 3 provides details on the most common drugs consumed. Nine of the ten most common drugs are generic. Not surprisingly, these drugs are all cheaper than the anti-platelet drug Plavix, which is the only branded drug on the list. The total prices of the drugs range from \$8 to \$175, with mean out-of-pocket prices in the initial coverage region ranging from \$8 to \$41.

Table 2: Most common drug classes in estimation sample

Indicator	Number Rx	% of obs.	Most common Rx
Cholesterol Lowering	2,201	9.4	Simvastatin
Renin-Angiotensin System Blocker	1,851	7.9	Lisinopril
Beta-Blocker	1,288	5.5	Metoprolol
Opioid	1,233	5.2	Hydrocodon
Antidepressant	1,220	5.2	Sertraline
Diuretic	1,214	5.2	Furosemide
Calcium Channel Blocker	957	4.1	Amlodipine
Insulin Sensitizer	815	3.5	Metformin
Gastroesophageal Reflux & Peptic Ulcer	799	3.4	Omeprazole
Hypothyroidism	798	3.4	Levothyroxine

Table 3: Most common drugs in estimation sample

Drug name	Indication	Branred	Total price ($)	Out of pocket price ($)	Number of Rxs	% of obs.
Lisinopril	Renin-Angiotensin System Blocker	N	18.84	9.90	727	3.1
Metoprolol	Beta-Blocker	N	30.73	10.27	649	2.8
Simvastatin	Cholesterol Lowering	N	33.52	11.51	640	2.7
Hydrocodon	Opioid	N	21.48	7.92	625	2.7
Plavix	Antiplatelet	Y	174.60	41.43	603	2.6
Furosemide	Diuretic	N	8.23	6.85	589	2.5
Levothyroxine	Hypothyroidism	N	11.52	9.26	562	2.4
Metformin	Insulin Sensitizer	N	24.59	9.68	526	2.2
Amlodipine	Calcium Channel Blocker	N	52.35	11.07	508	2.2
Warfarin	Anticoagulant	N	16.20	8.46	346	1.5

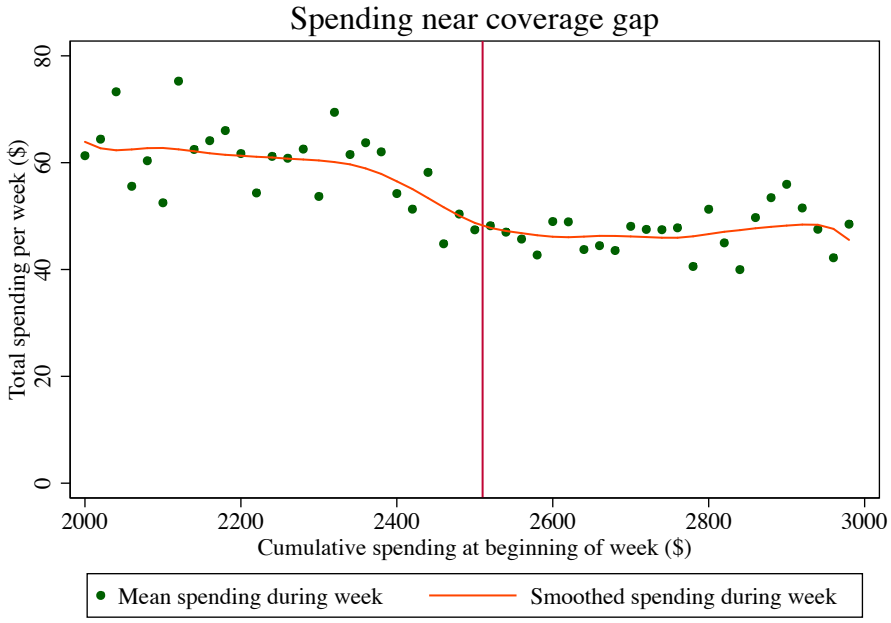
Note: reported total prices and out-of-pocket prices derived from authors' calculations.

4 Evidence from Discontinuity Near Donut Hole

This section presents evidence on whether individuals act in a way that is consistent with rational forward-looking behavior or with myopia. We base our evidence on the testable implications of the model developed in Section 2.3.

We perform a series of regression discontinuity analyses that all use our analysis sample of enrollees who arrived near the donut hole in the middle of the year. The unit of observation for each regression is an enrollee observed over a week. We allow these enrollees to be in the estimation sample for the complete weeks after their entry into the donut hole until either the last complete week of the year or the last week where their expenditures at the beginning of the week are less than $3,000.

Figure 3: Mean drug spending on cumulative spending: plans in base sample



We start with non-parametric kernel smoothed "lowess" regressions of mean total drug spending on beginning-of-week cumulative spending.[18] Figure 3 shows the kernel smoothed

[18] We use a bandwidth of 0.3 for these regressions.

regression results as well as the (non-smoothed) mean total drug spending by $20 increments of beginning-of-week cumulative spending, both for our base sample. The mean total drug spending shows little change in spending over the range $2,000-2,380 in beginning-of-week cumulative spending. Mean spending then drops until the donut hole and remains roughly constant until the highest cumulative spending level.

Note that week observations that are near the donut hole but not yet in the donut hole may move the individual into the donut hole, either because of an expensive drug or because of multiple drugs. This will result in an effective current price that is higher than before the donut hole but lower than in the donut hole (see equation 1). Thus, the fact that spending starts to drop slightly before the donut hole does not indicate that individuals are forward-looking. In contrast, the flat spending in the $2,000-2,380 range and the flat but lower spending in the donut hole range is a pattern that is consistent with myopia but not geometric discounting, as in Figure 1.

Figure 4: Mean drug spending on cumulative out-of-pocket spending: falsification plan

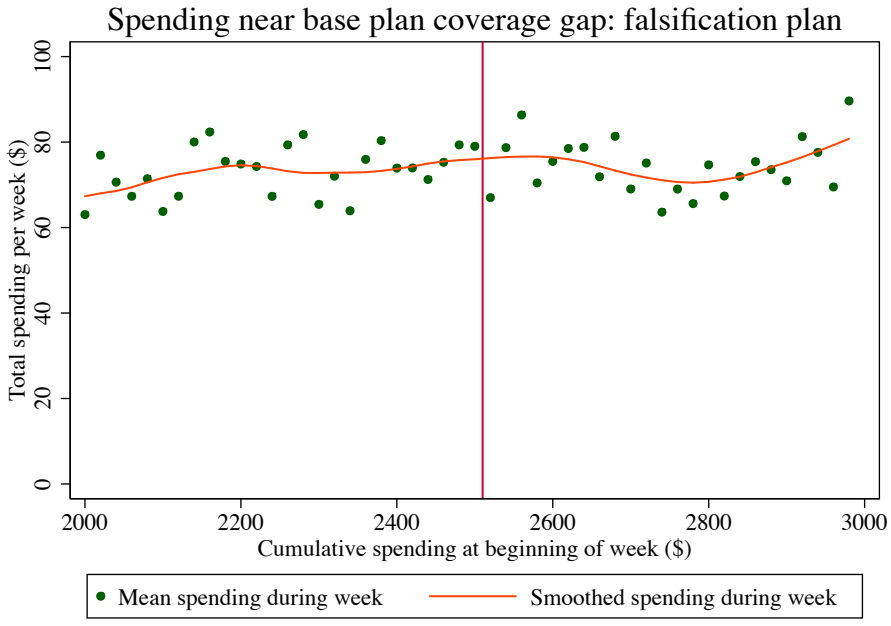


Figure 4 examines a falsification exercise on Plan CL3A, which had a coverage gap that

started at $4,000 in total spending. We run the same kernel smoothed regression on this plan as on our base sample of five plans, using again enrollees who reach $2,000 in spending early in the year. We find very different results: there is no drop in spending upon reaching $2,510 in total spending. Thus, this figure supports the conclusion that there is nothing medically significant regarding $2,510 that would cause drug expenditures to fall upon reaching this level, but rather that it is due to the coverage gap itself.

Figure 5: Mean drug spending on cumulative spending: plans in base sample

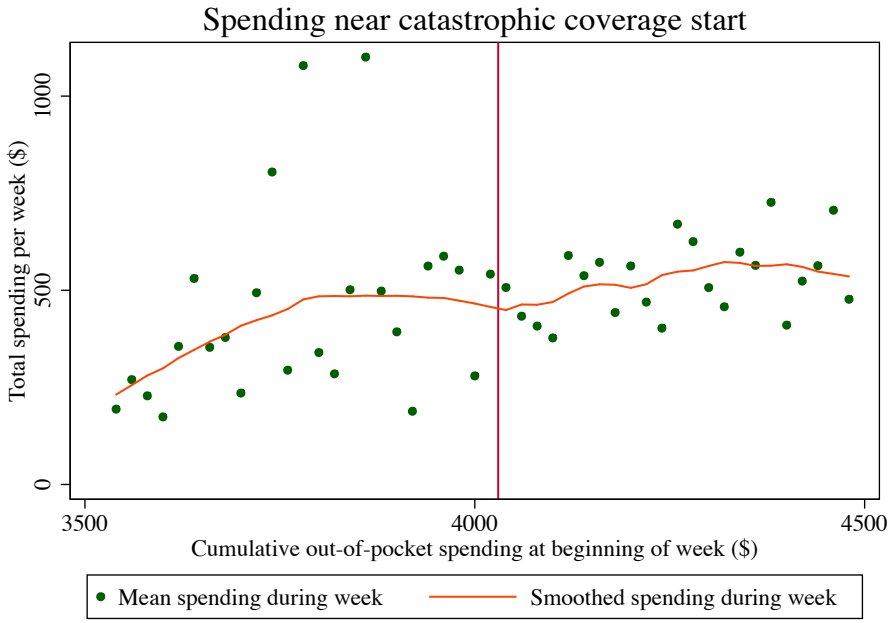


Figure 5 performs the final non-parametric regression. It again examines the base sample, but now considering the region around $4,030 in total (not out-of-pocket) spending, which is the mean of the start of catastrophic coverage across our five base plans. For this figure, we use the sample of enrollees who reach $3,530 in total spending between March 30 and July 26, 2008. The figure shows much weaker but somewhat suggestive evidence that spending rises until near the catastrophic coverage region and is flat from this point on. Any evidence here will likely be much weaker due to the much smaller sample size of people reaching the catastrophic coverage region than reaching the donut hole. Note that this figure shows a

ramp-up in spending in the pre-catastrophic-coverage period, implying preferences here that may be consistent with geometric discounting.

As noted above, for Figure 3 to reflect myopia, it must be that individuals expect to end the year inside the donut hole and not at the boundary between the initial coverage and coverage gap regions, as would have occurred if valuations for drugs lie strictly in-between the out-of-pocket and full prices of the drug, for instance. In order to consider this hypothesis more fully, Figure 6 considers the extent to which there is "bunching" of the end of the year expenditures around the donut hole. For the full sample of enrollees in the plans we consider, we find evidence of a small amount of bunching, consistent with both myopia and geometric discounting.[19] In contrast, we see no evidence of bunching for our estimation sample, which is consistent with the fact that enrollees in the estimation sample almost always end the year well within the donut hole. The lack of bunching also implies that individuals are not stockpiling drugs at the end of the year to exhaust the insurance from the initial coverage region, as has been found in other medical contexts (Cabral, 2013).

Having shown with non-parametric regressions that there is a discontinuity in spending near the donut hole, we now turn to linear regressions that quantify the spending drops and examine them in more detail. Our linear regression specifications follow the form:

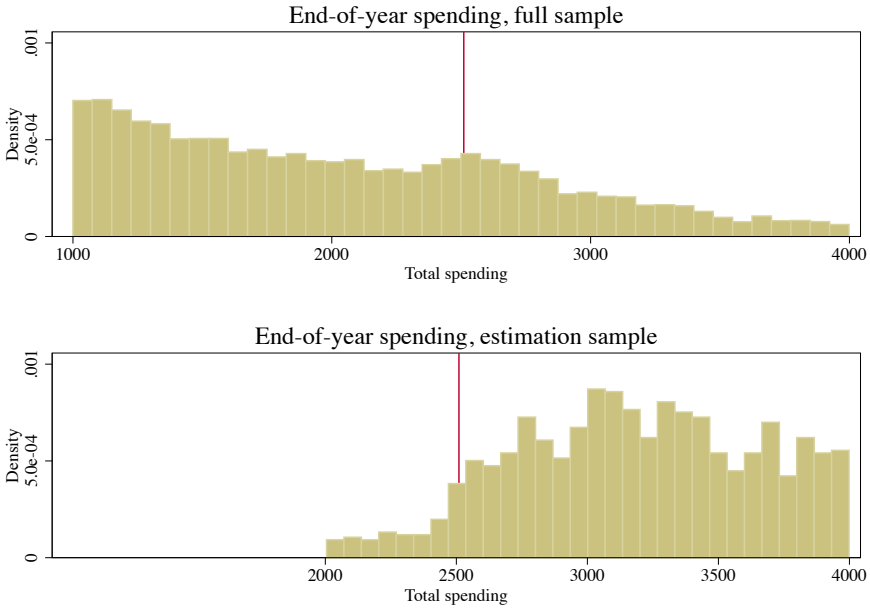
$$Y_{it} = FE_i + \lambda_1 1\{0 < m_{it0} \leq \$110\} + \lambda_2 1\{m_{it0} = 0\} + v_{it}, \tag{5}$$

where Y_{it} is the dependent variable, FE_i are enrollee fixed effects, λ_1 is the coefficient on an indicator for being above \$2,400 in spending (within \$110 of the donut hole) and λ_2 is the coefficient on an indicator for being in the donut hole, which implies starting the week with at least \$2,510 in expenditures. We examine a number of different dependent variables, including total prescription drug expenditures, branded drug expenditures, and number of prescriptions filled. The λ_1 coefficient captures the fact that observations that are near the donut hole but not yet in the donut hole may move the individual into the donut hole.

By selecting a small region around the donut hole, we are comparing the same individual

[19] Einav et al. (2015) also find evidence of end-of-the-year bunching for the set of all Part D enrollees in their data.

Figure 6: Histogram of end-of-year drug spending for estimation and full sample



at similar points in the year but faced with different current prices. This minimizes the possibility that factors other than the presence of the donut hole might be influencing our findings. By including individual fixed effects, we are further controlling for individual differences at different points in our sample, i.e. the possibility that more severely ill individuals show up more in the region after the donut hole.

Our first set of linear regression findings are reported in Table 4.[20] All our results cluster standard errors at the level of the enrollee. We find sharp drops in every measure of prescription drug use. Supporting the results in Figure 3, total drug spending dropped by $18 from a baseline of $62. The number of prescriptions fell by 21% from a baseline mean of 0.84 per week. Branded prescriptions fell more than generic prescriptions: 27% versus 19%. Similarly, expensive prescription – those with a total price of $150 or more – fell by 27% while inexpensive ones – those under $50 – had no significant drop. The mean total price of a prescription fell by 12% from a baseline level of $80. All effects, except for those

[20]In the interest of brevity, we do not report either the enrollee fixed effects or λ_1 values in our tables.

on the number of inexpensive prescriptions, are statistically significant. Not reported in the table, the indicators for weeks that start with $2,400 to $2,509 in total spending are generally significantly negative and much smaller than the reported coverage gap indicators, as would be consistent with some, but not all, of these observations moving enrollees into the donut hole.

These results paint a picture of enrollees who react strongly to being in the donut hole. These are enrollees faced with roughly the same end-of-year price and yet, their spending drops when faced with a different current price. As discussed in Section 2.3, the interpretation of this result is that individuals have either a β or a δ that is substantially less than one: they are not discounting with a factor suggested by rational economic theories at standard discount rates. Note that this interpretation does not account for the fact that a non-negligible subset of these enrollees will end up in the catastrophic coverage region. As in Section 2.3, to the extent that this effect is biasing our results, it should imply that individuals *increase* their spending as they cross through the coverage gap region. Thus, if anything, accounting for the catastrophic coverage effect would make our coefficient estimates more strongly negative.

Table 4: Behavior for sample arriving near coverage gap

Dependent variable:	Mean value before $2,400	Beginning of week spending in: $2,510 - 2,999	N
Total drug spending in week	61.83	-17.62^{**} (1.39)	30,305
Mean total price per Rx	79.47	-9.83^{**} (1.37)	11,197
Number of Rxs	0.84	-0.18^{**} (0.02)	30,305
Number of branded Rxs	0.30	-0.08^{**} (0.01)	30,305
Number of generic Rxs	0.54	-0.10^{**} (0.01)	30,305
Expensive Rxs	0.12	-0.04^{**} (0.00)	30,305
Medium Rxs	0.23	-0.06^{**} (0.01)	30,305
Inexpensive Rxs	1.10	-0.01 (0.01)	30,305

Note: Standard errors in parentheses. '**' denotes significance at the 1% level and '*' at the 5% level. Each row represents one regression. All regressions also include enrollee fixed effects and an indicator for beginning-of-week spending between $2,400 and $2,509, and cluster standard errors at the enrollee level. Sample consists of enrollees who reach $2,000 in spending between Mar. 30 and Jul. 26, 2008. An observation is an enrollee/week and includes weeks with beginning-of-week spending $\geq \$2,000$ and $< \$3,000$. Inexpensive Rxs are less than $50 and expensive ones are $150 or more.

While Table 4 presents evidence that enrollees lower spending across spending and branded categories, it does not indicate whether enrollees are substituting from one category to an-

other in the donut hole. For instance, even though generic drugs usage drops overall in the donut hole, it is possible (but not certain) that people substitute away from branded drugs and towards generic drugs, but that this is more than offset by substitution from generic drugs to the outside option. Table B1 in the Appendix addresses this point further by examining the extent to which branded or expensive drug availability in the period before our sample (with less than $2,000 in total spending) influences substitution towards generic or cheaper drugs in the donut hole. We find that having a high fraction of expensive drugs ex ante in a drug class predicts a significant increase in medium-priced drugs in the donut hole. The effect (at .170 prescriptions per week) is a little more than than twice as large the base drop in medium-priced drugs in the donut hole, implying that drug classes with 42% or greater expensive drugs see an increase in medium-priced drugs in the donut hole, implying substitution from expensive drugs to medium-priced drugs. We find similar effects of substitution towards generic drugs when there are many branded drugs in a class, but the effect is not statistically significant, nor is the effect of substitution from expensive drugs towards inexpensive drugs.

Our next set of results seek to provide evidence on whether consumers act myopically or as though they are geometric discounters with a low discount factor. Here, we perform the same regressions as in Table 5 but with the addition of an extra regressor, which measures the change in spending in the region $2,200 to $2,399. Thus, the excluded region is now $2,000 to $2,199, the beginning of the sample for any individual. Supporting the results in Figure 3 again, there is no significant effect of total spending in the $2,200 to $2,399 range. The implication is that, while spending before the donut hole is higher than in the donut hole, the increment does not grow as one moves further back before the start of the donut hole. Only one of the seven other regressions yields significant coefficients in the $2,200 to $2,399 range, and this coefficient is much smaller than the coefficients in the donut hole. Following Section 2.3, the results are not consistent with geometric discounting with $\delta > 0$ but are consistent with myopia.

Finally, Table 6 provides evidence on the five drug classes which have the largest drops in prescriptions upon entering the donut hole and the five with the largest increases in prescrip-

Table 5: Behavior near coverage gap with variation in pre-coverage gap region

	Mean value	Beginning of week spending in:		
Dependent variable:	before \$2,400	\$2,510 - 2,999	\$2,200 - 2,399	N
Total drug spending in week	61.83	-18.03^{**} (1.78)	-0.83 (2.26)	30,305
Mean total price per Rx	79.47	-9.01^{**} (1.72)	1.68 (2.13)	11,197
Number of Rxs	0.84	-0.20^{**} (0.02)	-0.04 (0.03)	30,305
Number of branded Rxs	0.30	-0.08^{**} (0.01)	0.01 (0.01)	30,305
Number of generic Rxs	0.54	-0.12^{**} (0.02)	-0.04^{*} (0.02)	30,305
Expensive Rxs	0.12	-0.04^{**} (0.01)	-0.00 (0.01)	30,305
Medium Rxs	0.23	-0.06^{**} (0.01)	0.00 (0.01)	30,305
Inexpensive Rxs	1.10	-0.02^{*} (0.01)	-0.02 (0.02)	30,305

Note: Standard errors in parentheses. '**' denotes significance at the 1% level and '*' at the 5% level. Each row represents one regression. All regressions also include enrollee fixed effects and an indicator for beginning-of-week spending between \$2,400 and \$2,509, and cluster standard errors at the enrollee level. Sample consists of enrollees who reach \$2,000 in spending between Mar. 30 and Jul. 26, 2008. An observation is an enrollee/week and includes weeks with beginning-of-week spending $\geq \$2,000$ and $< \$3,000$. Inexpensive Rxs are less than \$50 and expensive ones are \$150 or more.

tions. Here, we perform similar regressions to Table 4 but with the number of prescriptions in the drug class as the dependent variable. We then report the drug classes with the biggest and smallest coefficients on the spending drop in the donut hole region. The five drug classes with the biggest drops in prescriptions are also among the ten most common drug classes, as reported in Table 2. Indeed, the only one of the top five drug classes that does not have a drop that is also in the top five is opioids. The five drug classes with the biggest increases in prescriptions upon entering the donut hole are all drug classes with very few prescriptions. The magnitudes of the positive effects are very small and none are statistically significant.

We believe that, more than anything, this table shows that the percentage drops in prescriptions is similar across drug classes other than pain relief medications. The fact that individuals reduce spending on opioids less upon entry into the donut hole may reflect an interaction between myopia and the greater immediacy of the benefit from this class of drugs relative to the other most common drugs. This finding is also consistent with Chandra et al. (2010) who find similar demand responses to increased cost-sharing across drug categories.

Table 6: Largest and smallest changes in drug classes near coverage gap

Dependent variable Number of Rxs for:	Mean value before \$2,400	Beginning of week spending in: \$2,510 - 2,999	N
Cholesterol Lowering	0.080	-0.0176^{**} (0.0034)	30,305
Beta-Blocker	0.046	-0.0136^{**} (0.0023)	30,305
Gastroesophageal Reflux & Peptic Ulcer	0.031	-0.0129^{**} (0.0023)	30,305
Renin-Angiotensin System Blocker	0.065	-0.0124^{**} (0.0029)	30,305
Antidepressant	0.045	-0.0105^{**} (0.0024)	30,305
Opthalmic Antibiotic	0.002	-0.0000 (0.0006)	30,305
Antidiarrheal	0.001	0.0002 (0.0004)	30,305
Folic Acid Antagonist Antibiotic	0.003	0.0003 (0.0008)	30,305
Diuretic & Renin-Angiotensin System Blocker	0.002	0.0003 (0.0005)	30,305
Antiarrhythmic	0.002	0.0007 (0.0005)	30,305

Note: Standard errors in parentheses. '**' denotes significance at the 1% level and '*' at the 5% level. Each row represents one regression. All regressions also include enrollee fixed effects and an indicator for beginning-of-week spending between \$2,400 and \$2,509, and cluster standard errors at the enrollee level. Sample consists of enrollees who reach \$2,000 in spending between Mar. 30 and Jul. 26, 2008. An observation is an enrollee/week and includes weeks with beginning-of-week spending $\geq \$2,000$ and $< \$3,000$. Inexpensive Rxs are less than \$50 and expensive ones are \$150 or more.

5 Econometrics of Structural Model

5.1 Estimation

Let the individuals in our sample be denoted $i = 1, \ldots, I$. We group the individuals into groups $g = 1, \ldots, G$ based on their ACG score. Let $g(i)$ denote the group of individual i. We assume that Q (the distribution of health shocks), $\overline{N}$ (the maximum number of health shocks), and P (the probability of each health shock) vary across groups, implying that s (the market share of each drug) also varies across groups. Let these terms now be indexed by g also, so that we have Q_{gn}, $\overline{N}_g$, P_{gh}, and $s(g, m, n, h, j)$ respectively.

We now define some additional necessary notation. For each person/week observation it, let N_{it} denote the number of health shocks. Now, for $n = 1, \ldots, N_{it}$, let m_{itn} denote the value of m, the dollars till the donut hole; $h_{itn} \in \{1, \ldots, H\}$ denote the realization of the health shock; and $j_{itn} \in \{0, \ldots, J_h\}$ denote the drug chosen.

Our basic approach to estimation is maximum likelihood with a nested fixed point estimation. Essentially, the model boils down to an optimal stopping problem (where stopping indicates a drug purchase) together with a discrete choice of many options (where an option is a particular drug). In this way, the problem is similar to Rust (1987)'s classic paper on optimal stopping and also to more recent work that combines optimal stopping decisions with a multinomial choice (see, for instance, Melnikov, 2013; Hendel and Nevo, 2006; Gowrisankaran and Rysman, 2012).

Our framework differs from these models in that we do not observe all health shocks. Indeed, we only observe health shocks when the individual chooses to purchase a drug rather than the outside option. Moreover, we cannot condition on purchasing the inside option since a large part of our identification will come from people choosing not to purchase drugs as they approach or are in the donut hole.

We start by explaining what our likelihood would be if we observed the choice of the outside option, and then explain how the likelihood is different based on not observing the outside option. If all health events were observable, then N_{it}, h_{itn}, j_{itn}, and m_{itn} would all

be observable. We could then write the log likelihood for individual i at week t as:

$$\ln L_{it} = \sum_{n=0}^{N_{it}} \ln\left(1\{n = N_{it}\}(1 - Q_{g(i)n}) + 1\{n < N_{it}\}Q_{g(i)n}P_{g(i)h_{itn+1}}s(g(i), m_{itn+1}, n+1, h_{itn+1}, j_{itn+1})\right). \tag{6}$$

In words, the log likelihood for an observation can be broken down into a sum across health shocks n. For each n (starting at 0), there are two possibilities: an additional health shock occurrence or none. If there is an additional health shock what matters is the probability of seeing the additional shock multiplied by the conditional probability of the observed shock (given that one is observed) and the conditional probability of the drug chosen for that shock (given the observed shock). If there is no additional shock, then the likelihood is simply the probability of seeing no more shocks. The likelihood in this case would be a function of the values of P and Q, and of the parameters $(\overline{\alpha}, \beta, \delta, \gamma)$ which enter through s.

Having described the likelihood for the model with the outside option observed, we now consider the likelihood given our actual data in which we only observe health shocks when the individual purchases a drug instead of the outside option. Rather than attempting to identify P and Q from our sample of observations near the donut hole, we assume that individuals who are sufficiently below the donut hole to be before our sample have a low enough price sensitivity that they will always choose an inside drug. Given this assumption, we estimate P and Q non-parametrically for each group g using the enrollees observed in the period from after the deductible region (we start at \$300 to be conservative) until right before the start of our sample. Thus, these parameters are no longer part of the likelihood function.

Conceptually, we can compute the likelihood for an enrollee/week observation by specifying all possibilities for the places (in terms of at which health shocks) the health shocks with a drug chosen could have occurred, and then summing the probabilities of the observed data given all these possibilities. We illustrate this idea with an example. Consider an enrollee/week observation with 2 purchased drugs, with A being purchased before B, and has

a maximum of 4 health shocks in a period. The drug purchases could occur at the following health shock places (with A being before B always): $((1,2),(1,3),(1,4),(2,3),(2,4),(3,4))$. Thus, we sum the probabilities over each distinct set of health shock times. For each of the above six cases, we also need to sum over the possible number of total health shocks. For instance, if the drug purchases occur at shocks $(1,2)$, the total number of health shocks can vary between 2 and 4, while if the drug purchases occur at shocks $(2,4)$, there must be exactly four health shocks.

Given an instance for when the health shocks with purchases occur and the number of health shocks, the likelihood for a weeklong observation is relatively straightforward to define: it is simply the probability of observing the exact sequence of events, e.g., if there are health shocks at times 1 and 3 and four health shocks total, it is the probability of observing the particular choices at those four periods, including the outside option choices made at periods 2 and 4. Our likelihood estimator then simply sums over the cases of when the health shocks with purchases occur and the number of health shocks.

To illustrate our likelihood function more formally, let $\hat{N}_{it}$ denote the number of health shocks where the purchase included an inside product. Let $l_{itn}, n = 1, \ldots \hat{N}_{it}$ denote the places of each health shock, with $1 \leq l_{it1} < \cdots < l_{it\hat{N}_{it}} \leq \overline{N}_{g(i)}$. Let $\mathcal{L}(\hat{N}, \overline{N})$ denote the set of possible vectors of places when there are $\hat{N}$ health shocks with an inside good purchase and $\overline{N}$ possible purchase occasions. Thus, the above example shows that $\mathcal{L}(2,4)$ has six elements, all listed above. Then, the log likelihood is:

$$\begin{aligned}\ln L_{it} = \log \Bigg(& \sum_{l_1,\ldots,l_{\hat{N}_{it}} \in \mathcal{L}(\hat{N}_{it},\overline{N})} \sum_{N_{it}=l_{\hat{N}_{it}}}^{\overline{N}_{g(i)}} \left(\prod_{n=0}^{N_{it}-1} Q_{g(i)n} \right) (1 - Q_{g(i)N_{it}}) \\ & \prod_{n=1}^{\hat{N}_{it}} P_{g(i)h_{itn}} s\left(g(i), m_{itn}, l_n, h_{itn}, j_{itn}\right) \Big) \\ & \prod_{n=1, n\neq l_1,\ldots,n\neq l_{\hat{N}_{it}}}^{N_{it}} \left(\sum_{h=1}^{H} P_{g(i)h} s\left(g(i), \min_{\tilde{n} \text{ s.t. } l_{\tilde{n}}<n} m_{it\tilde{n}}, n, h, 0\right)\right)\Bigg). \quad (7)\end{aligned}$$

In words, the first line of (7) represents the double sum over the possible places of each health

shock (l) and the number of health shocks ($\overline{N}_{g(i)}$), and, for each case, lists the probability of observing that many health shocks. The second line provides the probabilities of seeing the drugs chosen for the health shocks with observed drug choices, where the places of the drug shocks show up through l_n. The third line is the probability of seeing an outside option chosen at each place without a drug purchase, where the dollar amount until the donut hole m is simply the dollar amount from the most recent drug purchase (which is also the minimum dollar amount across previous purchases). Note that equation (7) is similar to the earlier likelihood in equation (6) but with two main differences: (1) it integrates over the places of each observed shock, the total number of health shocks, and the drug class for health shocks with the outside option chosen; and (2) it combines all observations in a week because these observations are no longer separable given the unknown places and number of health shocks.

The advantage of our formulation in (7) is that it derives the likelihood in closed form conditional on any set of health shock occurrences $\mathcal{L}(\hat{N}_{it}, \overline{N})$. By solving for the likelihood in closed form, we eliminate the need for simulation which improves the efficiency and computational time required to estimate our model.

The remaining challenge is in iterating over the elements of (7). Here we adapt the methods of Gowrisankaran (1999), who derived an algorithm for enumerating all possible states for a number of different exchangeable structures. Our enumeration of the elements of $\mathcal{L}(\hat{N}_{it}, \overline{N})$ is similar, though not identical, to any of the cases described in Gowrisankaran. Appendix A describes our method in more detail, which follows Gowrisankaran (1999) closely.

Finally, note that our estimation algorithm has over 100 parameters, with a fixed effect for every drug. It can be difficult to estimate structural, dynamic models with this many parameters. Fortunately, with the exception of β and δ our estimation is similar to a multinomial logit model, which is quasiconcave and hence can be estimated with standard derivative-based search methods. Hence, we estimate the model by performing a grid search over β and δ and then using a derivative-based search for every value of β and δ.[21] Not reported in the paper, we also performed Monte Carlo simulations to verify the accuracy

[21] We also sped up computation by using parallel computation methods and by using the structure of the problem, where the donut hole is an absorbing state without any dynamic behavior, to simplify the computation. This was particularly important when δ was close to 1.

of the code and also that the estimator was identified and empirically well-behaved and of sufficient power to identify the parameters of interest.

5.2 Identification

The parameters that we seek to identify from our structural likelihood estimation are $\overline{\alpha}$, β, δ, and γ. To understand the identification, let us first condition on the price elasticity terms $\overline{\alpha}$. Following the arguments in Section 2.3, variation in the aggregate purchase probability immediately before and further before the donut hole region will help identify β and δ. The γ parameters are fixed effects on each drug and can be identified purely from a cross-section. The data after the start of the coverage gap form a repeated cross-section without dynamics that can identify these fixed effects.

We have seen evidence from Section 4 that, provided δ is positive, the data cannot be consistent with $\beta = 1$. However, beyond this, the identification of the discount factor parameters depends crucially on the price elasticity parameters. For instance, both a low β and a high $\overline{\alpha}$ would lead to a large drop in drug purchases following the donut hole. Thus, we cannot separately identify β and $\overline{\alpha}$ using the variation in aggregate purchase probabilities. Thus, the simple one-drug model as presented in Section 2.3 will not separately identify these parameters.

Fortunately, there is another source of variation in our analysis, which is the cross-drug and cross-time variation in drug prices. Because drugs are organized into drug classes and because there is different substitutability across different drug classes, the variation in the donut hole induced by the discount factors will be different from $\overline{\alpha}$. For instance, some expensive drugs have more cheap substitutes in their drug class than other expensive drugs. Expensive sole-source drugs typically have a larger portion paid by the insurance in the initial coverage region than do cheap drugs.

As an example, consider two parameter vectors with the same aggregate drop in spending upon reaching the coverage gap region. Suppose that the first parameter vector has a higher $\overline{\alpha}$ and β and the second one has a lower $\overline{\alpha}$ and β. Both could generate the same aggregate

drop in spending upon reaching the donut hole. However, under the first parameter vector, we would expect to see more substitution away from drugs with many substitutes in their drug class and less substitution away from drugs with few substitutes. This variation will then separately identify $\overline{\alpha}$ from β.

Our overall takeaway is that to identify discount factors from administrative data such as ours, it is necessary to concurrently identify price elasticity parameters. An accurate characterization of the drug purchase choice is, in turn, necessary to identify price elasticity parameters. We believe that our choice model for drugs, which follows closely from the standard multinomial choice models widely used in the industrial organization literature, works reasonably well in this regard. Within this class of models, defining drug classes appropriately is also very important.

6 Structural Estimation Results and Counterfactuals

6.1 Estimation Results

This subsection describes the results of our structural estimation. We stratify the sample of patients in Section 4 by ACG score. We perform three separate estimations: one for the lowest ACG score, one for the highest ACG core, and a third for everyone with a score that lies in between. Note that the ACG score takes on discrete values so there are substantial mass points of people with the lowest and highest scores.

We report our estimation results in Table 7. We start with our estimation of the price elasticity coefficients. We estimate a spline function here, and find that, for all three samples, the price coefficients follow a decreasing pattern in drug costs. While the price coefficients are all significant, it appears that people care far less about prices for higher-priced drugs than for lower-priced ones. The price coefficients are also very similar across the three samples, suggesting that the individuals across these groups have very similar preferences.

Translating the price coefficient into elasticities yields average own price elasticities between $-.20$ and $-.95$ depending on the price of the drug and the ACG category of the enrollee.

Table 7: Results of structural estimation

Estimation sample:	Lowest ACG score	Middle ACG score	Highest ACG score
Price spline $-\overline{\alpha}_1$, < \$20	-0.110^{**} (0.011)	-0.120^{**} (0.006)	-0.127^{**} (0.014)
Price spline $-\overline{\alpha}_2$, < \$50	-0.011^{**} (0.004)	-0.007^{**} (0.002)	-0.017^{**} (0.005)
Price spline $-\overline{\alpha}_3$, < \$100	-0.019^{**} (0.003)	-0.016^{**} (0.01)	-0.022^{**} (0.003)
Price spline $-\overline{\alpha}_4$, ≥ \$100	-0.006^{**} (0.002)	-0.007^{**} (0.0007)	-0.005^{**} (0.0016)
Hyperbolic discounting β	0 (–)	0 (–)	0 (–)
Geometric discounting δ	–	–	–
ln L	−27,995.3	−98,888.8	−20,672.6
ln L at $\beta = 0.1, \delta = 0.4$	−27,997.8	−98,899.1	−20,674.8
ln L at $\beta = 0.1, \delta = 0.999$	−27,999.1	−98,911.1	−20,677.8
Number of drug fixed effects	125	254	142
N	4,898	20,091	8,592

Note: Standard errors in parentheses. '**' denotes significance at the 1% level and '*' at the 5% level. Sample consists of enrollees who reach \$2,000 in spending between Mar. 30 and Jul. 26, 2008 and is stratified by ACG score. Each row displays the results from one maximum likelihood estimation. All specifications include fixed effects γ for each drug. An observation is an enrollee/week and includes weeks with beginning-of-week spending ≥ \$2,000 and < \$3,000.

These elasticities are very similar to the typical estimates from the literature on Medicare enrollee price elasticities.[22] Einav et al. (2015) estimate Medicare Part D arc elasticities between −.50 and −.75, Karaca-Mandic et al. (2013) estimates arc elasticities of adherence for statin drugs in Part D of −.95, and Simon and Ketcham (2008) estimate an arc-elasticity of −.22.[23]

Turning to the discount factor parameters, these estimates imply complete myopia, that $\beta = 0$ for all three samples. At this level of β, δ is not identified, since the future does not impact current decisions at all. Hence, we do not report a value for δ. We report the log likelihood values with a few different values of β and δ, optimizing across other parameters. For the sample with the lowest ACG score, a likelihood ratio test would not cannot reject $\beta = 0.1$ and $\delta = 0.4$ ($\chi^2(2) = 4.4, P = 0.11$). However, we can reject most of the other parameter values that are presented. These structural estimates are reconciling our reduced

[22]See Goldman and Joyce (2012) for a review of the early literature.

[23]There is a segment of the literature that finds lower elasticities. Using data prior to the implementation of Medicare Part D, Chandra et al. (2010) and find arc own price elasticities for prescription drug consumption of −.08 to −.15 while Abaluck et al. (2015) find an elasticity of −.09.

form findings that there is a large drop in drug spending after the start of the donut hole. Following our discussion in Section 5.2, the drop in drug spending here is too large to be explained by price elastic demand and a β that is strictly positive. Thus, the model with $\beta = 0$ fits the data the best.

While our parameter estimates imply complete myopia, we interpret these results somewhat more broadly. That is, we do not believe that Part D enrollees are not partially forward looking. Rather, we believe that the parameter estimates capture the fact that these consumers are not sophisticated, dynamic decision makers. This interpretation squares with the large behavioral economics literature that finds that individuals often behave in time inconsistent ways (DellaVigna, 2009). Given that we use our model to predict the impact of counterfactual policies, the important issue is the extent to which the preference parameters that we estimate remain constant across such policies. Here, we believe that the lack of dynamic sophistication will remain true if, for example, the government mandated donut hole coverage for generics. Other policies, such as eliminating the donut hole, result in dynamic sophistication being irrelevant. However, we also view believe that one should view geometric discounting for these enrollees not so much as a change in preferences, but rather, as an increase in dynamic sophistication.

6.2 Counterfactuals

In this subsection, we consider counterfactuals as to enrollee preferences and insurance environments. Our counterfactuals build on our structural estimation framework in two ways.

First, we use a different sample for the counterfactuals. Our estimation sample for the structural estimation pertains to a selected set of enrollees who reached a high spending level of \$2,000 early in the year. In order to consider more externally useful counterfactuals, we take a convex combination of enrollees in our estimation sample and enrollees in the same plans who are not in our estimation sample. We use enrollees and estimates with middle ACG scores (the middle column of Table 7), but the results are very similar across ACG scores. The combination is chosen so that 33% of enrollees reach the coverage gap, the same as the

aggregate figure for 2008. We find that using a counterfactual sample that is 32.1% from our estimation sample, with the remainder from other enrollees, generates this 33% figure with the estimated parameters. For each ACG score, we then predict the arrival of disease shocks and drug types (Q and P respectively) separately for the non-estimation sample, using the time period after the deductible and before \$2,000, as with the estimation sample.

Second, we compute a 52-week model, where we model both the donut hole and the catastrophic coverage region, instead of an infinite horizon model with the donut hole as an absorbing state. The reason we make this choice is that it is quite probable that some individuals will not reach the donut hole with high probability in the counterfactual analysis. This includes both individuals who are not in the estimation sample and also individuals in the estimation sample when we endow them with counterfactual preferences such as geometric discounting. Thus, we no longer compute an infinite-horizon model where the donut hole is the absorbing state.

Using our new sample and finite-horizon environment, we start by examining the relative importance of *behavioral hazard* to *moral hazard* that we estimate for our enrollees. Behavioral hazard is the extent to which enrollees' behavior is affected by the fact that they are myopic. To compute the extent of behavioral hazard, we compare the baseline Part D program to outcomes with geometric discounting and $\delta = 0.999$ at the weekly level (or 0.95 at an annual level). Moral hazard in healthcare refers to the extent to which health insurance raises enrollees' spending relative to environments without insurance (Pauly, 1968). We examine the extent of moral hazard by comparing the baseline to the case without insurance.

Our results on the relative extent of behavioral and moral hazard are reported in Table 8. We report the results from three cases: (1) the baseline, (2) the case where the enrollees are geometric discounters, and (3) the case without insurance. For each case, we examine outcomes, in terms of spending by the government and enrollees. We also examine consumer welfare, here calculated as the level of welfare from the point of view of a planner who geometrically discounts with a weekly factor of $\delta = 0.999$ and who knows when individuals are myopic.

Comparing the estimated $\beta = 0$ case we estimate to the standard, time-consistent case, we

Table 8: Relative impact of behavioral and moral hazard

Statistic (per week)	Baseline	Geometric discounting (no behavioral hazard)	No insurance (no moral hazard)
	Case 1	Case 2	Case 3
Number of Rxs	0.70	0.68	0.58
Number of branded Rxs	0.14	0.11	0.11
Number of generic Rxs	0.40	0.43	0.33
Expensive Rxs	0.06	0.04	0.04
Medium Rxs	0.25	0.22	0.21
Inexpensive Rxs	0.39	0.41	0.33
Enrollee spending ($)	17.69	12.28	32.99
Insurer spending ($)	24.68	23.80	0.00
Total spending ($)	42.37	36.08	32.99
Consumer welfare	1.50	1.55	0.97

Note: Counterfactual calculations based on estimated parameters from Table 7. Simulations are performed for 52 weeks starting enrollees at $0 in expenditures and use a mix of the estimation sample and other enrollees in the same plans so that 33% reach the donut hole in the base simulation. Geometric discounting case uses an annualized discount factor of 5%.

find that geometric discounting would cause a 31% drop in weekly enrollee prescription drug spending and a 15% drop in total drug expenditures. However, there is little difference in the number of prescriptions drugs between the two scenarios. Instead, there is a significant change is the composition of drugs consumed. There is a 33% drop in prescriptions for expensive drugs with substitution towards the most inexpensive. This substitution effect is also apparent in the relative increase in the number of generic drugs. Interestingly, there is a small decrease in insurer expenditures in moving to geometric discounting, as, enrollees substitute to drugs which are cheaper for themselves and also for the insurers.

In our model, the fact that $\beta = 0$ in combination with the nonlinear benefit design causes enrollees to over-consume prescription drugs relative to the consumption path they would follow if they could pre-commit to optimal state-contingent behavior. Because of this, the consumer surplus, as evaluated by the rational agent, increases by 3.3% by moving from the extremely myopic case of $\beta = 0$ to the time-consistent, geometric discounting scenario.

The no-insurance counterfactual allows us to compare the relative importance of moral hazard to behavioral hazard in affecting prescription purchase behavior. We find that without any insurance, mean total spending drops by 22% and consumer welfare drops by 35%. Thus,

the effect of behavioral hazard on spending is almost as large as the moral hazard effect from eliminating prescription drug insurance entirely.

Note that enrollees will have the same consumption behavior without insurance regardless of their discounting, as dynamics no longer impact the consumption decision when there is no insurance. Thus, comparing Cases 1 and 2 against Case 3, we find that insurance has a larger impact on myopic enrollees than it does on geometric discounters. Insurance induces two types on inefficiencies: time-inconsistent consumption and the standard moral hazard effect. Here we document that both distortions are important.

Table 9: Impact of filling the donut hole

Statistic (per week)	Baseline	No donut hole	No donut hole with constant insurer spending	No donut hole for generics only
	Policy 1	Policy 2	Policy 3	Policy 4
Number of Rxs	0.56	0.59	0.58	0.57
Number of branded Rxs	0.16	0.17	0.16	0.15
Number of generic Rxs	0.33	0.35	0.35	0.36
Expensive Rxs	0.10	0.11	0.09	0.09
Medium Rxs	0.15	0.15	0.14	0.15
Inexpensive Rxs	0.31	0.32	0.35	0.33
Enrollee spending ($)	16.33	12.57	15.11	14.73
Insurer spending ($)	26.44	34.54	26.44	28.01
Total spending ($)	42.78	47.10	41.56	42.74
Consumer welfare	1.10	1.22	1.18	1.16
Coinsurance rate	from data	from data	42%	from data

Note: Counterfactual calculations based on estimated parameters from Table 7. Inexpensive Rxs are less than $50 and expensive ones are $150 or more. Simulations are performed for 52 weeks starting enrollees at $0 in expenditures and use a mix of the estimation sample and other enrollees in the same plans so that 33% reach the donut hole in the base simulation.

We now examine the implications of counterfactual policies regarding eliminating the donut hole. Table 9 presents the results of the baseline (Policy 1) and three counterfactual policies. Policy 2 removes the donut hole and implements the same prices during the donut hole as existed before the donut hole. Policy 3 removes the donut hole but would leave insurance spending constant by setting the coinsurance to a constant fraction of the total price of the drug that we expected would be higher than the average 25% fraction in our

sample. Finally, Policy 4 removes the donut hole for generics only. We take the same sample of patients as in Table 8, meant to approximate a nationally representative sample of Medicare Part D enrollees.

We find that without the donut hole (Policy 2) the total number of prescriptions increases 4% and total drug spending increases 7%. Insurer drug spending would increase 29%. Enrollees consume more drugs and more expensive drugs. Interestingly, this impact is sandwiched between the estimates of Einav et al. (2015) and Abaluck et al. (2015). Einav et al. (2015) estimate that removing the donut hole will increase pharmaceutical spending 10% while Abaluck et al. (2015) estimate that figure to be 6%. Not surprisingly, we find that consumer welfare increases significantly under the more generous benefit structure.

An obvious questions to raise is what are the health consequences of different drug benefit structures. We can provide some rough estimates using the results in Chandra et al. (2010). They find an inpatient cost offset of increasing prescription drug out-of-pocket expenditures for the Medicare population. Specifically, increases in co-payments lead to increases in inpatient utilization. Using their estimates and applying to the increase in drug consumption under Policy 2 implies that the probability of an inpatient admission would decrease by a modest .4%.[24]

Policy 3, removing the donut hole for generics, yields an increase in consumer surplus and a small decrease in total spending. That is, this policy is welfare increasing relative to the baseline policy. Individuals decrease spending on branded drugs and substitute to cheaper generics and slightly increase the total number of drugs they consume.

Under a linear contract with the same insurer financial exposure (Policy 4), enrollees face a 42% coinsurance rate, up from a current average of roughly 25%. Enrollees consume slightly more drugs as they no longer have a large, time-inconsistent decline in consumption once they hit the donut hole. This contract is modestly welfare improving as consumer surplus increases 1.5%.

[24]This calculation uses the Chandra et al. (2010) 2002 policy change translating our weekly change in drug consumption to a monthly impact and then scaling the inpatient offset by the relative difference in the change in consumption. We assume that the average Medicare Part D enrollee spends 14.8% of the year in the donut hole – the average implied by our data. For this calculation we assume all the offset in Chandra et al. (2010) is attributable to the decline in drug consumption and not the decline in outpatient visits.

The message of the counterfactuals is that the nonlinear benefit design leads to reductions in welfare because of the myopic consumption behavior of consumers. Our results indicate that policies that reduces the need for enrollees to be sophisticated dynamic consumers improve welfare.

7 Conclusion

The Medicare Part D program established an important prescription drug benefit, but one that required enrollees interested in optimizing their drug purchases to calculate an inherently dynamic problem, due to the coverage gap. This paper considers the extent to which consumers follow the complex dynamic incentives created by Part D, and examines the implications of the myopia that we find.

We develop a dynamic modeling framework for complex insurance contracts which allows for quasi-hyperbolic discounting. Using the framework we provide a regression-discontinuity based test for myopia. A central challenge of estimating the impact of dynamic incentives on consumer behavior is selection: individuals compared across different settings may be different in dimensions that are often unobservable. Our test is based on examining how individuals who arrive near the donut hole early in the year change their behavior upon reaching the donut hole. It avoids selection issues by considering how a given individual changes her behavior within a relatively small time period.

We find strong evidence in favor of myopia, defined as hyperbolic discounting, $\beta < 1$, or a complete absence of forward-looking behavior, $\delta = 0$. Enrollees lower their prescription drug purchases upon reaching the donut hole, with a disproportionate drop for drugs that cost over \$150 and branded drugs. Moreover, the data can reject the presence of rational individuals with a low discount factor in favor of myopia as there is no progressive drop as enrollees approach the donut hole.

Having established evidence in favor of myopia, we then turn to structurally estimating the parameters of our model. Our modeling framework builds on standard industrial organization choice models, with a multinomial choice problem where enrollees face random disease shocks

that require treatment by a particular drug class and then choose to purchase one of a number of drugs in that class or the outside option. Notably, the price elasticity parameters are separately identified from the discount factor parameters by the panel data variation in drugs within different drug classes. We perform our structural estimation with a nested fixed point algorithm. The difficulty in estimation is that we do not observe enrollee decisions to purchase the outside option of no drugs, and yet we know that substitution to the outside good is a central response to entry into the donut hole. We develop computationally attractive methods to estimate this model with maximum likelihood. Here we find that consumers have significant price elasticities but that we cannot reject complete myopia.

Our structural estimation approach has several limitations. These include (1) that we do not allow for any medical dynamics to treatment, only financial dynamics as the enrollee approaches the donut hole; (2) that we do not measure substitute therapies to drugs; (3) that we do not model imperfect physician agency; and (4) that our arrival process for diseases is relatively simple. Nonetheless, we believe that our structural results are reasonable, given the large drop in spending shown by the regression-discontinuity evidence.

Using our parameter estimates, we then examine the impact of counterfactual preferences and policies. We find that myopia in this market is almost as important than moral hazard. That is, if we could transform all Medicare Part D enrollees into geometric discounters with standard discount factors, it would lower prescription drug usage almost as much as if we eliminated prescription drug insurance altogether. We also examine policies to close the donut hole, as mandated by the 2010 Affordable Care Act. We find that they would either increase insurer spending significantly (in the case of no change in coinsurance rates) or increase coinsurance rates significantly (in the case of revenue neutrality).

References

Abaluck, J. and Gruber, J. (2011). Choice inconsistencies amount the elderly: Evidence from plan choice in the Medicare Part D program. *American Economic Review*, 101(4):1180–1210.

Abaluck, J. and Gruber, J. (2013). Evolving choice inconsistencies in choice of prescription drug insurance.

Abaluck, J., Gruber, J., and Swanson, A. (2015). Prescription drug utilization under Medicare Part D: A dynamic perspective.

Baicker, K., Mullainathan, S., and Schwartzstein, J. (2012). Behavioral hazard in health insurance. Technical report.

Cabral, M. (2013). Claim timing and ex post adverse selection. Technical report, Mimeo, University of Texas.

Chandra, A., Gruber, J., and McKnight, R. (2010). Patient cost-sharing and hospitalization offsets in the elderly. *American Economic Review*, 100:193–213.

Chung, D. J., Steenburgh, T., and Sudhir, K. (2013). Do bonuses enhance sales productivity? a dynamic structural analysis of bonus-based compensation plans. *Marketing Science*, 33(2):165–187.

DellaVigna, S. (2009). Psychology and economics: Evidence from the field. *Journal of Economic Literature*, 47:315–372.

Einav, L., Finkelstein, A., and Schrimpf, P. (2015). The response of drug expenditures to contract design in Medicare Part D. *Quarterly Journal of Economics*, forthcoming.

Fang, H. and Wang, Y. (2013). Estimating dynamic discrete choice models with hyperbolic discounting, with an application to mammography decisions.

Goldman, D. P. and Joyce, G. F. (2012). Consumer demand and health effects of cost-sharing. *The Oxford Handbook of the Economics of the Biopharmaceutical Industry*, page 365.

Gowrisankaran, G. (1999). Efficient representation of state spaces for some dynamic models. *Journal of Economic Dynamics and Control*, 23:1077–1099.

Gowrisankaran, G., Norberg, K., Kymes, S., Chernew, M. E., Stwalley, D., Kemper, L., and Peck, W. (2013). A hospital systems wellness program linked to health plan enrollment cut hospitalizations but not overall costs. *Health Affairs*, 32(3):477–485.

Gowrisankaran, G. and Rysman, M. (2012). Dynamics of consumer demand for new durable goods. *Journal of Political Economy*, 120:1173–1219.

Handel, B. R. (2013). Adverse selection and inertia in health insurance markets: When nudging hurts. *The American Economic Review*, 103(7):2643–2682.

Heiss, F., McFadden, D., and Winter, J. (2010). ?mind the gap! consumer perceptions and choices in Medicare Part D prescription drug plans. In Wise, D., editor, *Research Findings in the Economics of Aging*. University of Chicago Press.

Hendel, I. and Nevo, A. (2006). Measuring the implications of sales and consumer stockpiling behavior. *Econometrica*, 74:1637–1673.

Ho, K., Hogan, J., and Morton, F. S. (2014). The impact of consumer inattention on insurer pricing in the Medicare Part D program.

Hotz, J. V. and Miller, R. J. (1993). Conditional choice probabilities and the estimation of dynamic models. *Review of Economic Studies*, 60:497–529.

Karaca-Mandic, P., Swenson, T., Abraham, J. M., and Kane, R. L. (2013). Association of medicare part d medication out-of-pocket costs with utilization of statin medications. *Health services research*, 48(4):1311–1333.

Ketcham, J., Lucarelli, C., Miravete, E., and Roebuck, C. (2012). Sinking, swimming, or learning to swim in Medicare Part D. *American Economic Review*.

Ketcham, J., Lucarelli, C., and Powers, C. A. (2014). Paying attention or paying too much in Medicare Part D?

Ketcham, J. and Simon, K. (2008). Medicare Part D's effects on elderly drug costs and utilization. *American Journal of Managed Care*, pages 14–22.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics*, 112(2):443–478.

Lichtenberg, F. and Sun, S. (2007). The impact of Medicare Part D on prescription drug use by the elderly. *Health Affairs*, 16:1735–1744.

Liu, F., Alexander, C., Crawford, S., Hedeker, D., Pickard, S., and Walton, S. (2011). The impact of Medicare Part D on out-of-pocket costs for prescription drugs, medication utilization, health resource utilization, and preference-based health utility. *Health Services Research*, 46(4):1104–1123.

Melnikov, O. (2013). Demand for differentiated products: The case of the U.S. computer printer market. *Economic Inquiry*, 51:1277–1298.

Neuman, P. and Cubanski, J. (2009). Medicare Part D update - lessons learned and unfinished business. *New England Journal of Medicine*, 361(4):406–414.

Newhouse, J. P. (1993). *Free for all?: lessons from the RAND health insurance experiment*. Harvard University Press.

Pauly, M. (1968). The economics of moral hazard: Comment. *American Economic Review*, 58(3):531–537.

Phelps, E. and Pollak, R. (1968). On second-best national saving and game-equilibrium growth. *Review of Economic Studies*, 35:201–208.

Rust, J. (1987). Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher. *Econometrica*, 55:999–1033.

Schroeder, M. C., Brooks, J. M., and Cook, E. A. (2014). Heart attacks and subsequent medication choices: the effect of medicare part d benefit phase and time of year.

Simon, K. and Ketcham, J. (2008). Medicare part d's effects on elderly drug costs and utilization. Technical report.

Strotz, R. H. (1956). Myopia and inconsistency in dynamic utility maximization. *Review of Economic Studies*, 23:165–80.

Yin, W., Basu, A., Zhang, J. X., Rabbani, A., Meltzer, D. O., and Alexander, C. (2008). The effect of the Medicare Part D prescription benefit on drug utilization and expenditures. *Annals of Internal Medicine*, 148:169–177.

Zhang, Y., Donohue, J. M., Lave, J. R., O'Donnell, G., and Newhouse, J. P. (2009). The effect of Medicare Part D on drug and medical spending. *New England Journal of Medicine*, 361:52–61.

Appendix A: Method to Enumerate Elements in $\mathcal{L}(\hat{N}, \overline{N})$

This section describes our method for enumerating all the elements in $\mathcal{L}(\hat{N}, \overline{N})$.[25] Recall that each element in $\mathcal{L}(\hat{N}, \overline{N})$ corresponds to one vector of places for the health shocks with inside good purchases when there are $\hat{N}$ health shocks with inside good purchases and $\overline{N}$ is the maximum number of health shocks. For instance, if $\overline{N} = 8$ and $\hat{N} = 3$, an element of $\mathcal{L}(\hat{N}, \overline{N})$ is $(1, 5, 8)$.

As in Gowrisankaran (1999), let $o(\cdot)$ denote the number of elements in a set. Using a similar proof structure to Gowrisankaran (1999) Theorem 1, we offer the following:

Proposition 1. *Using induction, the number of elements in $\mathcal{L}(\hat{N}, \overline{N})$ can be described as follows:*

Base case 1: $\hat{N} = 1$. $o(\mathcal{L}(1, \overline{N})) = \overline{N}$.

Base case 2: $\hat{N} = \overline{N}$. $o(\mathcal{L}(\overline{N}, \overline{N})) = 1$.

Inductive case: $1 < \hat{N} < \overline{N}$. $o(\mathcal{L}(\hat{N}, \overline{N})) = o(\mathcal{L}(\hat{N}, \overline{N} - 1)) + o(\mathcal{L}(\hat{N} - 1, \overline{N} - 1))$.

Proof I split the proof into assertions of the base cases and the inductive case.

Base case 1: $\mathcal{L}(1, \overline{N})$ enumerates all possible places for the single health shock with an inside good purchase. This single health shock can occur at any of the purchase occasions between 1 and $\overline{N}$. There are thus $\overline{N}$ possible places.

Base case 2: Here $\mathcal{L}(\overline{N}, \overline{N})$ represents all possible place vectors for the inside good purchases when the number of inside good purchases is equal to the maximum number of purchase occasions. Here, each purchase occasion must be used for an inside good purchase. Thus, the unique place vector is $(1, \ldots, \overline{N})$, which gives $o(\mathcal{L}(1, \overline{N})) = 1$.

Inductive case: Assume by induction that the theorem hold for all cases with maximum number of purchase occasions less than $\overline{N}$ and also for the $(\overline{N}, \overline{N})$ case. We now prove that it holds for the $(\hat{N}, \overline{N})$ case by induction, where $1 < \hat{N} < \overline{N}$.

We divide the possible place vectors into two exhaustive and mutually exclusive cases. Either the $\overline{N}$th health shock place has no inside good purchase or it has one. Suppose first

[25] For brevity of notation, this section suppresses the dependence of variables on individual i, group g, or time t.

that it has none. Then, all the $\hat{N}$ inside good health shocks must occur at the first $\overline{N}-1$ places. By the inductive assumption, there are $o(\mathcal{L}(\hat{N},\overline{N}-1))$ possible place vectors that satisfy this criterion. Now suppose that the last place contains the last inside good purchase. Then the $\hat{N}-1$ earlier inside good purchases must occur sometime during the first $\overline{N}-1$ places. Again by the inductive assumption, there are $o(\mathcal{L}(\hat{N}-1,\overline{N}-1))$ possible place vectors that satisfy this vector. Adding up the number of elements in both cases, we have proven the inductive case. ∎

Note that the inductive formula in Proposition 1 is the same as the inductive formula that defines binomial coefficients. Hence, we could also write $\mathcal{L}(\hat{N},\overline{N}) = \text{Binom}(\overline{N}+1,\hat{N}) \equiv \frac{(\overline{N}+1)!}{(\hat{N})!(\overline{N}+1-\hat{N})!}$. Finally, note that Gowrisankaran (1999) Theorem 2 provides a computationally efficient method for enumerating and accessing individual elements of $\mathcal{L}(\hat{N},\overline{N})$. The analogous method works here and hence we use the method from that paper to enumerate the elements here.

Appendix B: Extra Tables

Table B1: Behavior near coverage gap with variation across drug classes

	Dependent variable: number of Rxs in category during a week				
Regressor	Branded Rx	Generic Rx	Expensive Rx	Medium Rx	Inexpensive Rx
Frac branded	1.45**	−1.50**	–	–	–
	(.038)	(.054)			
Frac branded	−.065	.085	–	–	–
× donut hole	(.042)	(.060)			
Frac expensive	–	–	1.47**	.168*	−1.79**
			(.051)	(.065)	(.078)
Frac expensive	–	–	−.065	.170**	.056
× donut hole			(.042)	(.054)	(.084)
Donut hole	−.045*	−.041	−.031**	−.072**	−.002
($2,510 - 2,999)	(.019)	(.036)	(.011)	(.018)	(.027)
N	11,197	11,197	11,197	11,197	11,197

Note: Standard errors in parentheses. '**' denotes significance at the 1% level and '*' at the 5% level. Each column represents one regression. The regressor "frac branded" indicates the mean fraction of branded prescriptions filled for the drug classes chosen in a week prior to $2,000 in spending; similarly for "frac expensive." All regressions also include enrollee fixed effects and an indicator for beginning-of-week spending between $2,400 and $2,509, and cluster standard errors at the enrollee level. Sample consists of enrollees who reach $2,000 in spending between Mar. 30 and Jul. 26, 2008. An observation is an enrollee/week with at least one prescription filled and includes weeks with beginning-of-week spending $\geq \$2,000$ and $< \$3,000$. Inexpensive Rxs are less than $50 and expensive ones are $150 or more.