Myopia and Complex Dynamic Incentives: Evidence from Medicare Part D

Christina M. Dalton* Gautam Gowrisankaran† Robert Town ‡

September 7, 2015

Abstract

Standard Medicare Part D drug insurance provides limited coverage in a "doughnut 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 41%. Entirely eliminating the gap would increase insurer spending 31%, compared to 6% for generic-only gap coverage.

JEL Codes: I13, I18, D03, L88
Keywords: non-linear prices, cost sharing, doughnut hole, discontinuity

We have received helpful comments from Jason Abaluck, David Bradford, Chris Conlon, 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 acknowledges research support from the Center for Management Innovations in Health Care at the University of Arizona.

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

1 Introduction

In 2006, the U.S. Medicare system added drug coverage, through a program called Medicare Part D. Part D, which was the largest benefit change to Medicare since its inception, has proven very popular with Medicare enrollees.[1] Perhaps the biggest criticism of Part D is its non-linear price schedule. Enrollees with a standard Part D benefit faced modest out-of-pocket expenditures until their accrued total year-to-date expenditures placed them in the coverage gap—also called the "doughnut hole." Once in the doughnut hole, the enrollee paid the full price of all drugs until reaching the catastrophic zone. In 2008, the year of our data, the gap begins at $2,510 in total expenditures and did not end until $4,050 in out-of-pocket expenditures.

With a non-linear price schedule, a forward-looking rational enrollee must forecast her future expenditures when making a prescription fill decision, i.e., if she forecasts that she will end the year in the donut hole then she might choose cheaper drugs before the donut hole. If enrollees do not act rationally in the presence of non-linear insurance contracts, such contracts can create a welfare loss from "behavioral hazard," or sub-optimal behavior resulting from mistakes or behavioral biases (Baicker et al., 2012). This is relevant because some studies find that Part D enrollees do not act fully rationally in their different but related choice of Part D health plans (see Abaluck and Gruber, 2011, 2013; Ho et al., 2014; Heiss et al., 2010; Schroeder et al., 2014).[2] More generally, because most health insurance plans have non-linear aspects, such as out-of-pocket maxima and deductibles, understanding whether agents in health plans with non-linear prices make rational decisions is important.[3]

This paper has two goals. First, to develop robust tests for whether Part D enrollees act myopically in their choice of drug treatments. Our tests avoid several selection issues

[1] The program enrolled over 38 million (or 68%) of Medicare beneficiaries in 2013 (Medpac, 2014). 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).

[2] Also consistent with myopic behavior, critics of Part D point to the possibility that the doughnut hole may lead to adverse health consequences (Liu et al., 2011). However, some studies find that enrollees are, at least in part, rational in their Part D plan choice (Ketcham et al., 2012).

[3] This point that has been recognized since at least the RAND Health Insurance Experiment, which found that utilization increased once enrollees hit their out-of-pocket maxima (Newhouse, 1993).

that often make such inference challenging. Second, to 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. We test the implications of this model using a discontinuity design and provide evidence that people act myopically in their drug purchase decisions. 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 Part D enrollee faces a weekly 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 within the drug class or 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 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 future utility 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 examines if a set of enrollees, who all face the same future prices, respond when facing different current prices. Enrollees who arrive near the doughnut hole early in the year should all face the same price (full

out-of-pocket cost) by the end of the year with near certainty, which we verify empirically. Without any discounting, these individuals should not change their drug purchase behavior upon crossing into the doughnut hole, because the Part D insurance is essentially a fixed subsidy. Rational geometric discounters (with $0 < \delta < 1$ but $\beta = 1$) will decrease spending inside the doughnut hole only because of the time value of money saved. This implies a near-linear decrease in purchases before the doughnut hole. Finally, myopic individuals – whom we define to be those with either $\beta < 1$ or $\delta = 0$ – may decrease spending at the doughnut hole, but their purchase progression can be flat before entering the doughnut hole.

Using the group of enrollees who arrive near the doughnut hole early in the year, we find that drug purchases drop significantly and sharply upon reaching the doughnut hole: mean total drug expenditures fall by 28% and the number of prescriptions falls by 21%. In contrast, there is no evidence of expenditures falling in the periods leading up to the doughnut hole, with no significant differences between spending in the regions \$2,200 to \$2,399 and \$2,000 to \$2,199. A falsification exercise using a plan with a different doughnut hole reveals no drop in spending upon reaching \$2,510. Thus, 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 price elasticity parameters, fixed effects for each drug, and the discounting parameters β and δ. We estimate these parameters via maximum likelihood. We cannot observe in our data when people have health shocks but substitute to the outside good, a crucial piece for estimation. Thus, we develop methods that allow us to integrate in closed form over the shocks at which the individual chooses a drug, thereby making this estimator computationally tractable.[4] We show that we can identify discount factor parameters because of the across-drug and across-drug-class variation in prices and drug availability upon entry into the doughnut hole.

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

[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).

response that is lower for more expensive drugs with an overall demand elasticity of -0.32. The underlying reason for our results is that the drop in prescriptions upon reaching the doughnut hole is too large to be explained with a positive β, given the level of cross-drug substitution. At $\beta = 0$, δ is not identified. We do not interpret our estimates as implying that individuals are perfectly 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 behavioral and policy counterfactuals for a nationally representative sample. To isolate the extent of behavioral hazard, we examine how prescription purchase behavior changes if individuals are geometric discounters with $\delta = 0.95$ at an annual level. Geometric discounting causes enrollees to reduce their spending by 29%, with total prescription drug spending dropping by 14%. In contrast, moral hazard from insurance changes total prescription drug spending by 31%, implying that behavioral hazard is about one half as important as moral hazard in this market.

Our policy counterfactuals examine eliminating the doughnut hole as mandated by the 2010 Affordable Care Act. We find that eliminating the doughnut 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 the current average of 25% to 36% to implement a revenue neutral insurance scheme without the doughnut hole. Providing doughnut hole coverage for generic drugs only would increase insurer spending by only 6%.

Our work is most closely related to the works of Einav et al. (2015) and Abaluck et al. (2015). We develop complementary tests to Einav et al. (2015): we test for the presence of myopic behavior by quantifying the differences in drug purchase decisions across situations that are very similar except that they have different current prices but the same end-of-year prices (in our case, the same individual right before and after the doughnut hole start); Einav et al. test for the presence of forward-looking behavior by quantifying the differences across situations that are very similar except that they have the same current prices but different end-of-year prices (in their case, different individuals who turn 65 at different points in the

year and who choose plans with deductibles). Einav et al. also estimate a structural, dynamic model and find 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 also builds on Einav et al. by developing a modeling framework for drug choices that is more similar to a standard dynamic multinomial choice models and by providing results on identification for this type of model. Abaluck et al. (2015) use a very different identification strategy based on the assumption that changes in plan benefits are exogenous and do not result in enrollee selection due to plan stickiness. Using this assumption, they develop a simpler structural model of drug choice that abstracts away from the fact that enrollees may not fully know their health shocks requiring drug purchases at the beginning of the year. 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 doughnut hole. Section 5 describes the econometrics of our structural model. Section 6 provides results and counterfactuals, and Section 7 concludes.

2 Model

2.1 Overview

We model the weekly 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 the a given year, and hence who has arrived near the doughnut hole of \$2,510.[5] By focusing on this group, we can credibly assume that individuals in our sample know that they will hit the doughnut hole at some point during the year, which simplifies our tests and estimation.[6] We also assume that people

[5] Our empirical analysis uses the enrollee/week as the unit of observation. A longer time interval, such as a month, would reduce information through aggregation, while a shorter time interval, such as a day, may have noisy outcomes because a typical enrollee will fill zero prescriptions on most days. We chose an interval of a week as a balance between these two constraints.

[6] In reality, 96% of the individuals in this group reach the doughnut hole during the year.

will not leave the doughnut hole during the year. This assumption also generally holds in our data,[7] and, as we discuss in Section 2.3, employing this assumption only weakens our tests of myopia. Finally, we do not explicitly model the fact that there is a benefit "reset" that occurs on January 1 each year and that forward-looking enrollees may optimally stockpile drugs in response to this reset (as in Cabral, 2013). However, our findings are robust to limiting the sample to the part of the year in which this reset is not salient.[8]

We consider the decision-making of an individual enrollee,[9] starting at the first complete week where the individual has reached the $2,000 spending level. 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.* conditional on the individual's type. Because we allow for heterogeneity across health types, our framework is consistent with systematic differences in health shock probabilities across individuals as would occur with chronic diseases.[10]

At any point in time during the week, the individual knows how far she is from the doughnut 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

[7] Only 10% of individuals in this group end up in the catastrophic coverage region at the end of the year.

[8] We estimate the structural model using only data through November 1 and the parameter estimates are qualitatively identical.

[9] Section 5.1 discusses estimation of the model which involves aggregation across individuals.

[10] Our structural estimation stratifies individuals into types based on health risks. One type may be likely to have a chronic disease such as diabetes while another may not. In this case, the first type will be more likely to have a health shock that requires insulin sensitizer medications than will the second type. Thus, the unconditional distribution of health shocks will be serially correlated within individuals, and not *i.i.d.*.

$h = 1, \ldots, H$, there is a set of possible prescription drugs, $j = 1, \ldots, J_h$, that can be used to treat the condition, 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 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 doughnut hole after the cost of the current drug, the flow utility is $\gamma_{hj} - \alpha(oop_{hj}) + \varepsilon_{hj}$, while for individuals in the doughnut hole before the cost of the current drug, the flow utility is $\gamma_{hj} - \alpha(p_{hj}) + \varepsilon_{hj}$, where $\alpha(\cdot)$ maps from price to utility.

In order to flexibly capture the different impacts of price on decisions, we allow $\alpha(p)$ to be a linear spline, with cut points at $c_1 \equiv \$20$, $c_2 \equiv \$50$, and $c_3 \equiv \$100$. 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 doughnut hole may substitute away from drug purchases entirely.

2.2 Dynamics and Hyperbolic Discounting

The enrollee's decision problem is dynamic because filling a prescription brings her closer to the doughnut hole state of no insurance, and filling an expensive prescription brings her closer than filling a cheaper one. 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 case because it is a popular and parsimonious model of time inconsistency.

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 $t' > t$ with the

factor $\beta\delta^{t'}$. A value of $\beta = 1$ would fit into the standard geometric discounting model. 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. The agent would like to tie her future self to not spend too much, but cannot do that in our model. This payoff structure is referred to as "sophisticates" (Baicker et al., 2012), as individuals accurately perceive that they are hyperbolic discounters.[11]

Following Fang and Wang (2013), we formalize behavior using Bellman equations. We first exposit the state space. Let m indicate the monetary distance of the individual from the doughnut hole at the start of a given purchase decision. Since the doughnut 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. 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 doughnut hole is an absorbing state, we can write our dynamic decision problem as an infinite-horizon one, rather than a 52-week problem.

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 $(\varepsilon_{h0}, \ldots, \varepsilon_{hJ_h})$. Let $V^{\delta}(m, n)$ denote the value of future payoffs *gross of the hyperbolic β term* and before it is known whether or which new health shock will occur in the week. Last, define effective price for a consumer as:

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

which covers the consumer being completely in the doughnut hole (the first line), remaining completely below the doughnut hole (the last line), and the intermediate case, where the consumer crosses partially into the doughnut hole with the purchase (the middle line).[12]

[11] We could also model "naifs" who (correctly) recognize that they discount next period with the factor $\beta\delta$ but (incorrectly) perceive that will act as geometric discounters with discount factor δ starting in the next period. 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.

[12] In this case, the individual pays the out-of-pocket price and the remainder after reaching the doughnut

We now write:

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

From (2), the main dynamic effect is that each purchase moves the individual closer to the doughnut hole by p_{hj} dollars.

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:

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

where the $-\ln s(m,n,h,j)$ term accounts for the expectation of $\vec{\varepsilon}$ conditional on choice.[13] 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 sum over drug classes. We cannot use the standard logit expression for utility here: choices are made with $\beta\delta$ discounting of the future, but then future periods are discounted by δ from each other. Mathematically, 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 properties of our model across discount factors. These properties form the basis for our evidence of myopia in Section 4. For simplicity, we specify a linear price disutility, $\overline{\alpha}(p) = \alpha p$, one drug class and one drug in this class, so that $H = 1$ and

hole.

[13] We exclude Euler's constant from this expression as it does not affect choices.

$J_1 = 1$, and prices $p_{11} = \$100.4$ and $oop_{11} = \$25.1$.[14] For now, we only assume that the price coefficient, mean utility, and health shock frequency are in the range where individuals will make some purchases from inside the doughnut hole with probability close to 1. Finally, recall that 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.

First, we consider the case where $\beta = \delta = 1$. To avoid infinite payoffs, 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 end up in the doughnut hole. With these parameters, 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 then equivalent to one where the individual receives a fixed subsidy of \$1,882.5 and does not have any drug insurance. This is then a static multinomial logit problem with the full price. Thus, consumption behavior will not change upon crossing into the doughnut 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 doughnut 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 $75.3(1 - \delta^t)\alpha$ where t is the expected time until payment of this value. With an annual discount factor of 0.95, $(1 - \delta^t)$ is small. Thus, the geometric discounting model can only explain sizable variations in purchase before the doughnut hole with extremely large price elasticities 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

[14] Our example uses these prices as the doughnut hole start of \$2,510 is an even multiple of them.

the 24th and 26th purchase occasions is $75.3(1-\delta^{t'})\alpha$ 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 sum, geometric discounting predicts that purchase probabilities continue to increase the further away they are from the doughnut hole.

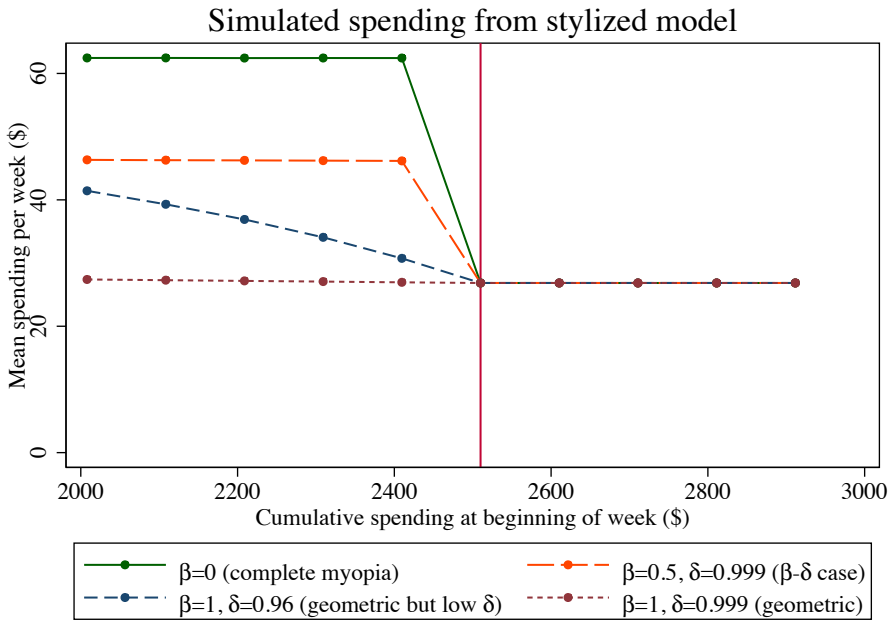
Finally, we consider the myopic case of $\beta < 1$ or $\delta = 0$. Considering again the 25th purchase occasions, the individual discounts the future payoff by $\beta\delta^t$. Thus, the difference in mean utilities between the 25th and 26th purchase occasion is now $75.3(1-\beta\delta^t)\alpha$. Similarly, the difference in mean utilities between the 24th and 26th purchase occasion is now $75.3(1-\beta\delta^{t'})\alpha$. If β is much below 1 and δ is very close to 1 then these two differences in utilities will be very close to each other but far from 0. For the extreme case of $\delta = 0$, individuals react exclusively to the current price, implying that the mean utility for the 24th and 25th 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 before reaching the doughnut hole are similar to each other but very different from those within the doughnut hole.

Figure 1 shows simulated spending for our different models graphically, constructed by solving for the dynamic decisions using the parameters noted above in combination with the assumption that there is exactly one health shock each week (i.e., $Q_0 = 1$, $Q_1 = 0$), a price term of $\alpha = 0.02$, a mean utility of $\gamma_{11} = 1$,[15] and different values of β and δ (as reported in the figure), and then simulating mean spending as a function of cumulative spending.

The figure shows that spending in the doughnut hole is flat and the same across values of β and δ, at about \$27 per week. The $\beta = 0$ case has perfectly flat spending before the doughnut hole and the highest spending of any case, at \$62 per week. With $\beta = 0.5$ and $\delta = 0.999$, mean spending is virtually as flat as with $\beta = 0$ before arriving in the doughnut hole although, at \$46 per week, substantially lower. With geometric discounting and $\delta = 0.999$ (or 0.95 at an annual level), mean spending is virtually flat throughout. Finally, with geometric discounting and a low discount factor of $\delta = 0.96$ (as in Einav et al. (2015)), mean spending drops almost

[15]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



linearly over the pre-doughnut-hole period, from $41 to $27.

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 doughnut hole have a lower purchase probability than previous periods further from the doughnut hole. Moreover, by focusing on the same individual and at similar time periods, our discontinuity-based tests 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. While this complicates the dynamic choice problem, the basic point remains: for individuals in our estimation sample, only discounting can explain changes to spending behavior upon entering the doughnut hole.

Second, one might consider a very simple alternate scenario, where the treatment 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 them surpass the start of the coverage gap (and most by a substantial amount), implying that they have valuations for some drugs that are sufficient for purchase even in the absence of insurance. Given that they know that they will end up purchasing once inside the doughnut 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.[16]

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 doughnut hole and its implications. Because Medicare enrollee drug consumption is principally tied to the treatment of chronic conditions, those who reach the doughnut hole in one year are likely to reach or approach the doughnut hole in the next year. In addition, as we detail below in Section 3, enrollees are provided substantial information regarding the structure of the doughnut hole and receive regular updates on their spending relative to the doughnut 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 doughnut hole. One might worry that these results are due to rational behavior where individuals who reach the doughnut 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 doughnut hole.

[16]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.

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 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.[17]

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 doughnut hole. Figure A1 in Appendix A shows an example of the part of the information that pertains to the distance to the doughnut 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.

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

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

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	A	B	C	D	E	F
Employer	1	1	1	2	2	3
% of employees from employer	26	45	9	79	21	46
Deductible ($)	275	100	100	0	200	0
Doughnut 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 A provides generic coverage in deductible region; Plan F used for falsification exercise only and also provides generic coverage in doughnut 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 doughnut 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 doughnut hole region is substantial. For three of the plans, the doughnut 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' doughnut holes end after the enrollee reaches $4,010 in out-of-pocket expenditures. Generous catastrophic coverage resumes after the doughnut 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 doughnut 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 doughnut hole with very high probability. We define a week as starting on a Sunday.

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 the enrollees in estimation 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 F 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 it.

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.[18] We then calculate the dollars until the doughnut hole (m) for each prescription by tabulating the spending up to this point during the year.[19]

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 shows that the base plans have mean ACG scores which are similar to the over-65 population mean score of 1; the falsification plan has a somewhat lower mean score.

Our analysis classifies each drug into a unique drug class meant to capture the function of the drug. We had the drug class coding performed by a clinically trained research assistant using the pill name, drug type, most common indicator, and national drug code. We classified drugs on the basis of function rather than the diseases they treat 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 drug classes 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 blood pressure medicines, opioids, and antidepressants. Not

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

[19]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.

shown in the table, there are 70 drug classes when we use our full sample.

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 A1 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, the only branded drug on the list. The total prices here range from $8 to $175, with mean out-of-pocket prices in the initial coverage region ranging from $8 to $41.

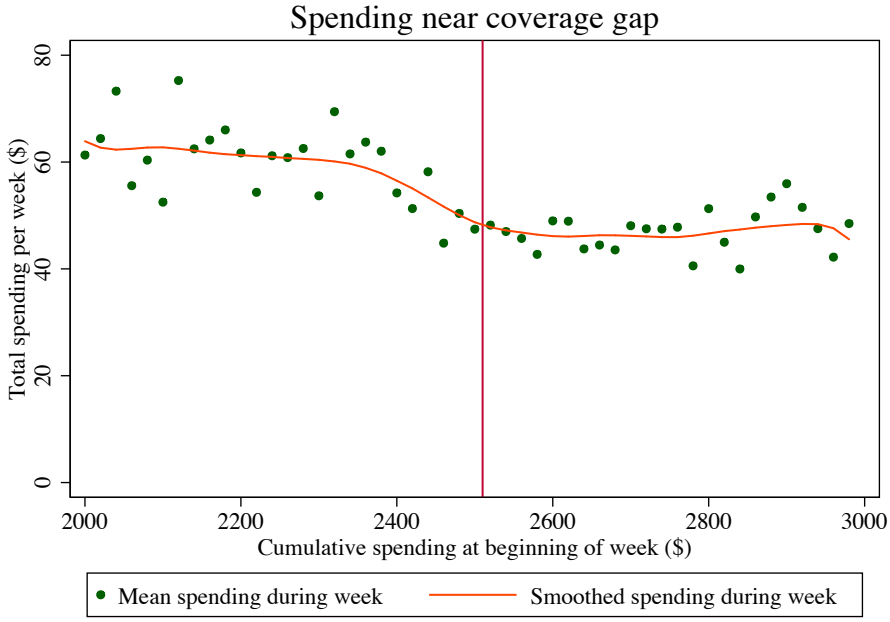
4 Evidence from Discontinuity Near Doughnut Hole

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 discontinuity-based analyses that all use our analysis sample of enrollees who arrived near the doughnut hole in the middle of the year. Our analyses are similar to a standard regression discontinuity framework. However, while regression discontinuity papers typically consider different individuals near a breakpoint, we are able to consider the same individual immediately before and after reaching the coverage gap.

Specifically, the unit of observation for each regression is an enrollee observed over a week. Enrollees are in the estimation sample from the first week with starting expenditures of over $2,000 until the last week with starting expenditures of less than $3,000, or the end of the

year if it comes first.

Figure 2: 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.[20] Figure 2 shows the kernel smoothed regression results as well as the (non-smoothed) mean total drug spending by $20 increments of beginning-of-week cumulative spending. 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 doughnut hole and remains roughly constant until the highest cumulative spending level.

Note that week observations that are near the doughnut hole but not yet in the doughnut hole may move the individual into the doughnut hole, either because of an expensive drug or because of multiple drugs. Thus, the fact that spending starts to drop slightly before the doughnut 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 doughnut hole

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

range is a pattern that is consistent with myopia but not geometric discounting with $\delta > 0$, as in Figure 1.[21]

Figure 3: Mean drug spending on cumulative out-of-pocket spending: falsification plan

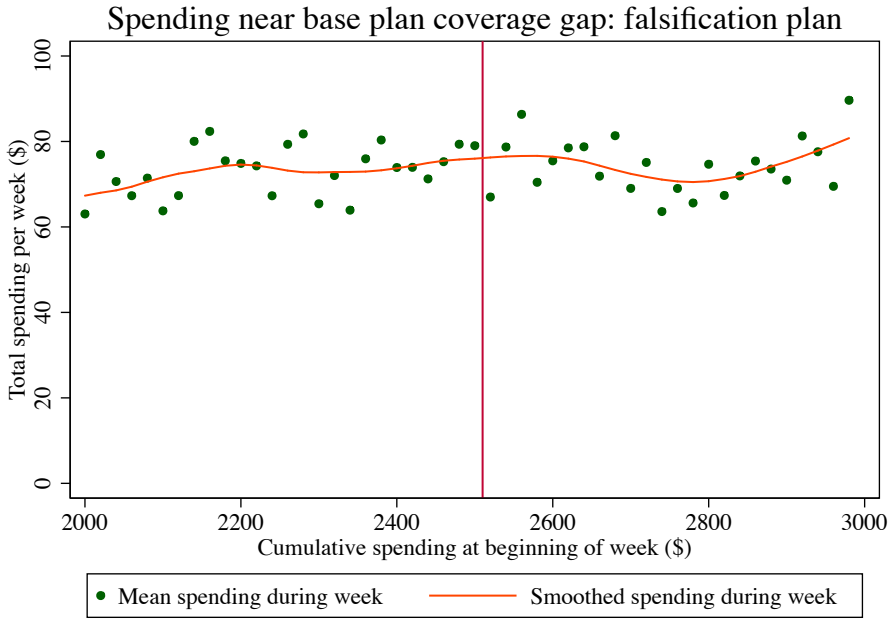


Figure 3 examines a falsification exercise on Plan F, which had a coverage gap that started at $4,000 in total spending. We report the same kernel smoothed regression on this plan as on our base sample. We find very different results: there is no drop in spending upon reaching $2,510 in total spending. 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.

As noted above, for Figure 2 to reflect myopia, it must be that individuals expect to end the year inside the doughnut 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

[21] Interestingly, we see a small increase in mean consumption as enrollees enter the catastrophic phase of the benefit design. Figure A3 in the Appendix displays the analogous figure to Figure 2 for the catastrophic zone. However, the sample sizes are small here and so the estimates of the impact of entering the catastrophic zone on consumption are imprecise.

the out-of-pocket and full prices of the drug, for instance. In order to consider this hypothesis more fully, Figure A2 considers the extent to which there is "bunching" of the end of the year expenditures around the doughnut 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.[22] 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 doughnut 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 doughnut hole, we now turn to linear regressions that quantify the spending drop and examines it 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 m_{it0} is the beginning-of-week spending left until the doughnut hole, 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 doughnut hole) and λ_2 is the coefficient on an indicator for being in the doughnut hole, which implies starting the week with at least \$2,510 in expenditures. We examine a number of different dependent variables Y_{it}, 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 doughnut hole but not yet in the doughnut hole may move the individual into the doughnut hole.

By selecting a small region around the doughnut hole, we are comparing the same individual at similar points in the year but faced with different current prices. This minimizes the possibility that factors other than the presence of the doughnut 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

[22] Einav et al. (2015) find similar evidence of end-of-the-year bunching.

show up more in the region after the doughnut hole.

Our first set of linear regression findings are reported in Table 3.[23] 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 2, 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 prescriptions – 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 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.

These results paint a picture of enrollees who react strongly to being in the doughnut 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 standard economic theories of dynamic decision-making.

While Table 3 presents evidence that enrollees lower spending across most drug categories, it does not indicate whether enrollees are substituting from one category to another in the doughnut hole. For instance, even though generic drugs usage drops overall in the doughnut 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 A2 in Appendix A 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 doughnut 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 doughnut hole.

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

Table 3: 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.

The effect (at .170 prescriptions per week) is a little more than than twice as large as the base drop in medium-priced drugs in the doughnut hole, implying that drug classes with 42% or greater of expensive drugs see net substitution towards medium-priced drugs in the doughnut hole. 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.

Next, Table 4 provides 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 3 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 estimation window. Supporting the results in Figure 2 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 doughnut hole is higher than in the doughnut hole, the increment does not grow as one moves further back. Following Section 2.3, the results are not consistent with geometric discounting with $\delta > 0$ but are consistent with myopia.

Finally, Table A3 provides evidence on the five drug classes which have the largest drops in prescriptions upon entering the doughnut hole and the five with the largest increases

Table 4: 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.

in prescriptions. Here, we perform similar regressions to Table 3 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 doughnut 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 doughnut hole are all drug classes with very few prescriptions. Overall, 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 doughnut 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.

5 Econometrics of the 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 doughnut 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 doughnut 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.

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 doughnut hole, we assume that individuals who are before our estimation window 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 starting after the deductible region (conservatively defined as \$300) until right before the start of our sample.

We enumerate the likelihood for the remaining parameters by conditioning on both the health shock places at which the enrollee could have made her observed drug purchases and the number of health shocks. Conditional on these two variables, the probability of seeing the observed purchases is the same as in equation (6) except that we need to sum over drug classes for the health shocks where the enrollee chooses the outside option. The likelihood, then, is simply the sum of these conditional probabilities taken over the health shock places

and number of health shocks.

We illustrate with an example. Consider an enrollee/week observation with 2 purchased drugs, with A being purchased before B, where the enrollee 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): $((1,2),(1,3),(1,4),(2,3),(2,4),(3,4))$. The total number of health shocks could always be 4, and for some cases, e.g., shocks at (1,2), could be 2 or 3. The likelihood then sums the probability of the observed data conditional on each of the 6 distinct pairs of health shock places and the total number of health shocks.

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. Note that $\mathcal{L}(2,4)$ has six elements as listed above. Then, the log likelihood is:

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

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 doughnut 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: first, 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 second, it combines all health shocks in a week because they are no longer separable given the unknown places and number of 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 enumerating the elements of $\mathcal{L}(\hat{N}_{it}, \overline{N})$. Appendix B 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. 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.[24] Not reported in the paper, we also performed Monte Carlo simulations to verify the accuracy of the code and power of the estimator.

5.2 Identification

The parameters that we seek to identify from our structural likelihood estimation are $\overline{\alpha}$, β, δ, and γ. We show identification by combining insights from two literatures. First, a behavioral economics literature has explained how to identify quasi-hyperbolic discounting models assuming that period utilities are known (Laibson, 1997). The literature shows that one can identify $\beta\delta$ by comparing time t tradeoffs between t and $t+1$ utilities, and δ by comparing time t tradeoffs between $t+1$ and $t+2$ utilities. One can then recover β as the

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

ratio between the recovered $\beta\delta$ and δ.

The analogous identification for our case is illustrated by Figure 1, which uses the simple one-drug case of Section 2.3. The figure shows that, conditional on period utilities, one can identify $\beta\delta^t$ (where t is the expected time between the 25th and 26th purchase) from the purchase probabilities at the 25th and 26th purchase occasions, and $\beta\delta^{t'}$ (where t' is the expected time between the 24th and 26th purchase) from the purchase probabilities at the 24th and 26th purchase occasions. As a caveat, when β or δ are 0, one cannot separately identify β from δ using this argument.

Second, an industrial organization literature has shown that dynamic discrete choice models do not identify both the (geometric) discount factor and static utilities without the presence of exclusion restrictions (Magnac and Thesmar, 2002). The intuition of this result applied to our case is that both very myopic enrollees (with low $\beta\delta$) and very price sensitive enrollees (with high $\overline{\alpha}$) would experience a large drop in their purchase probabilities upon entering the doughnut hole. We need a further restriction to separate $\overline{\alpha}$ from $\beta\delta$.

We achieve identification by using the heterogeneity across drugs in our sample. Specifically, we have cross-drug and cross-time variation in drug prices. For instance, some expensive drugs have more cheap substitutes in their drug class than other expensive drugs. Intuitively then, consider two parameter vectors which generate the same aggregate drop in spending upon reaching the coverage gap region, but where the one has a higher $\overline{\alpha}$ and $\beta\delta$. 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.

Formally, Magnac and Thesmar (2002)'s Proposition 4 shows that identification can be achieved by having two states which have different expected future values for some choices but the same expected future values for at least one choice (other than the outside option). Assuming sufficient heterogeneity across drugs, we observe states $(m, 0)$ and $(m', 0)$ and health shock h for which $m < m' < p_{hj}$ for some drug hj but for which $p_{hk} < m'$ for a different drug hk. Purchase of drug hj leads to being in the doughnut hole at both m and m' but purchase of drug hk does not lead to being in the doughnut hole at m. This implies that the values for choice hj are the same across the two states but that the values for

choice hk are different for the two states, thereby satisfying the conditions of Proposition 4.[25] Thus, we satisfy Magnac and Thesmar (2002)'s Proposition 4 using a restriction that comes directly from the economics of our problem. Assuming that β and δ are positive, we separately identify β and δ by identifying discount factors using two different values of m and m' and combining this identification with the quasi-hyperbolic identification argument above.

Our 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.

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.

We report our estimation results in Table 5. We start with our estimates of the price spline coefficients, which are very similar across the three samples. For all three samples, these coefficients decrease in price. Thus, it appears that people care far less about price for higher-priced drugs than for lower-priced ones.

In order to further understand our price coefficients, we used our estimated parameters to simulate the impact of a 1% increase in all drug prices and out-of-pocket prices on the expected number of drugs purchased by individuals in our sample over the entire year 2008. Using enrollees with middle ACG scores, we find that the 1% price increase would lead to a 0.32% decrease in the number of drugs purchased. This decrease is very heterogeneous across drugs, leading to a 1.0% decrease in the number of expensive drugs (priced at $150 or higher),

[25] Multiple drugs and heterogeneity across drug prices are necessary to reach states $(m, 0)$ and $(m', 0)$.

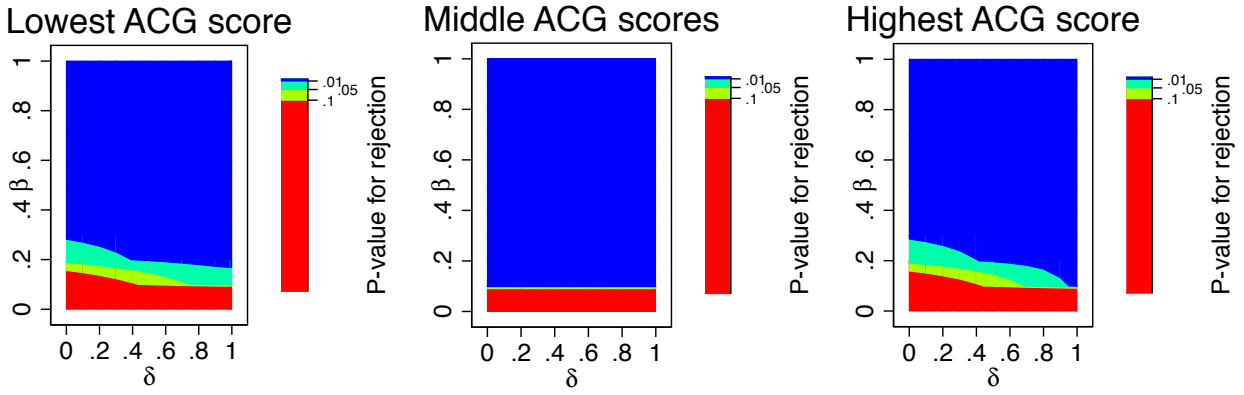
Table 5: Results of structural estimation

Estimation sample:	Lowest ACG score	Middle ACG scores	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 $-\overline{\alpha}_4$, $\geq$ \$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 $\beta = 0.1, \delta = 0.4$	−27,997.5	−98,899.1	−20,674.8
ln L at $\beta = 0.1, \delta = 0.999$	−27,998.8	−98,911.0	−20,677.8
ln L at $\beta = 0.2, \delta = 0.999$	−28,002.8	−98,934.9	−20,683.6
ln L at $\beta = 1, \delta = 0.1$	−28,020.9	−98,979.8	−20,693.5
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 $\geq \$2,000$ and $< \$3,000$.

a 1.1% *increase* in the number of medium-priced drugs ($50-$150), and a 0.4% decrease in the number of inexpensive drugs (priced at less than $50).[26] Comparing our elasticity of −0.32 to analogous numbers from the literature,[27] Abaluck et al. (2015) estimate a Medicare Part D elasticity of −0.09, Einav et al. (2015) estimate −0.50, and Ketcham and Simon (2008) estimate −0.22. Karaca-Mandic et al. (2013) estimate an elasticity of adherence for statin drugs of −0.95. Thus, we believe that our elasticity numbers are roughly in the middle of the range reported by the literature.

Figure 4: Confidence regions for β and δ across estimation samples



Turning to the discount factor parameters, we find 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. While we cannot compute a standard error for β at 0, we report confidence regions for β and δ in Figure 4, with the parameters computed with a grid of 0.1 and with the highest δ being 0.999. Consistent with Table 5, parameters with $\beta = 0$ are never rejected. For the (largest) sample with middle ACG scores, we reject all other values of β and δ.[28] With the other two samples, we find parameter values that are not rejected when $\beta \leq 0.3$. For the middle and highest ACG score

[26]The increase in medium-priced drugs is consistent with the reduced-form evidence in Table A2 of substitution from expensive to medium-priced drugs in the doughnut hole.

[27]Using the nationally representative sample described in Section 6.2 below, we find an elasticity of −0.44.

[28]There is the appearance of a non-reject region in the figure because of the discreteness of the grid points.

samples, $\delta = 0.999$ is always rejected when $\beta > 0$.[29] While we report separate results here only for the three ACG score groups, we estimated the model with age limitations, gender limitations, and using data only through November 1 (to avoid dynamic stockpiling effects) and our parameter estimates are very similar. In particular, our estimate of $\beta = 0$ holds for all the subsamples that we investigated also.

Overall, our estimates of β are consistent with our reduced form findings that there is a large drop in drug spending after the start of the doughnut hole. While they imply complete myopia, we interpret them more broadly. 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 doughnut hole coverage for generics. Other policies, such as eliminating the doughnut hole, result in dynamic sophistication being irrelevant.

6.2 Counterfactuals

We now consider counterfactuals as to enrollee preferences and insurance environments. We use enrollees and estimates with middle ACG scores (the middle column of Table 5), but the results are very similar across ACG scores. Our counterfactuals modify our structural estimation framework in two ways. First, since our estimation sample pertains to a selected set of enrollees who reached a high spending level early in the year, we create a nationally representative sample by taking a convex combination of enrollees in our estimation sample and enrollees in the same plans who are not in our estimation sample. The combination is chosen so that 33% of enrollees reach the coverage gap after 52 weeks, the same as the

[29] Table 5 also provides the log likelihoods for several different values of β and δ.

aggregate figure for 2008.[30] Second, we compute a 52-week model, where we model both the doughnut hole and the catastrophic coverage region, instead of an infinite horizon model with the doughnut hole as an absorbing state. The reason is that individuals here may frequently not reach the doughnut hole, unlike in our base estimation.

We start by examining the relative importance of *behavioral hazard* to *moral hazard* in Medicare Part D. 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. All cases report consumer welfare using $\beta = 1$ and an annualized 95% discount factor.

Table 6: 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	0.56	0.55	0.43
Number of branded Rxs	0.16	0.13	0.12
Number of generic Rxs	0.33	0.36	0.26
Expensive Rxs	0.10	0.07	0.06
Medium Rxs	0.15	0.14	0.12
Inexpensive Rxs	0.31	0.33	0.25
Enrollee spending ($)	16.33	11.56	29.47
Insurer spending ($)	26.44	25.29	0.00
Total spending ($)	42.78	36.85	29.47
Consumer welfare	1.10	1.16	0.67

Note: simulations use estimated parameters from Table 5 column 2. 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 same plans so that 33% reach the doughnut hole in the base case. Geometric discounting case uses an annualized discount factor of 95%.

The results, in Table 6, show that geometric discounting (Case 2) would cause a 29% drop

[30] Our counterfactual sample draws 29.6% from our estimation sample, with the remainder from other enrollees. For each ACG score, we predict the arrival of disease shocks and drug types (Q and P respectively) separately for the non-estimation sample.

in weekly enrollee prescription drug spending and a 14% drop in total drug expenditures relative to our estimated baseline with $\beta = 0$ (Case 1). 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 30% 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. 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, consumer surplus, as evaluated by the rational agent, increases by 5.5% by moving from the extremely myopic case of $\beta = 0$ to the time-consistent, geometric discounting case.

The no-insurance counterfactual allows us to compare the relative importance of moral hazard to behavioral hazard. We find that without any insurance (Case 3), mean total spending drops by 31% and consumer welfare drops by 39% relative to the baseline. Thus, moral hazard has about twice as large an impact on total spending as does behavioral hazard. Overall, Part D insurance induces two types on inefficiencies: time-inconsistent consumption and moral hazard. We find that both distortions are important.

We now examine the implications of counterfactual policies regarding eliminating the doughnut hole. Table 7 presents the results of the baseline (Policy 1) and three counterfactual policies. Policy 2 removes the doughnut hole by implementing the same prices during the doughnut hole region as existed before the doughnut hole. Policy 3 removes the doughnut hole but would leave insurance spending constant by setting the coinsurance to a constant fraction of the total price of the drug. Finally, Policy 4 removes the doughnut hole for generics only.

We find that without the doughnut hole (Policy 2) the total number of prescriptions increases 5% and total drug spending increases 10%. Insurer drug spending would increase 31%. Enrollees consume more drugs and more expensive drugs. Einav et al. (2015) also

Table 7: Impact of filling the doughnut hole

Statistic (per week)	Baseline	No doughnut hole	No doughnut hole with constant insurer spending	No doughnut 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	36%	from data

Note: simulations use estimated parameters from Table 5 column 2. 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 same plans so that 33% reach the doughnut hole in the base case.

estimate that removing the doughnut 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.

It is important to evaluate what might be the overall health consequences of removing the doughnut hole. We can provide some back-of-the-envelope calculations using Chandra et al. (2010), who estimate substitution between drug utilization and inpatient hospitalization. Applying these estimates to the increase in drug consumption under Policy 2 implies that inpatient hospital admissions would decrease by 1.4% by eliminating the doughnut hole.[31]

Under a linear contract with the same insurer cost (Policy 3), enrollees face a 36% coinsurance rate. This is significantly higher than the current average 25% coinsurance. Enrollees consume more drugs but fewer expensive ones than in the baseline as their myopia is no longer relevant since the contract is not dynamic. This contract is significantly welfare improving

[31]Chandra et al. (2010) find that a drug use drop of 18.2% leads to an increase in hospitalizations of 5.4%. We derive our result by applying the resulting elasticity of 0.27 to our 5% increase in drug spending. Our calculation assumes that all the offset in Chandra et al. (2010) is attributable to the decline in drug consumption and not the decline in outpatient visits.

as consumer surplus increases 7% relative to the baseline. This policy also shows that the nonlinear benefit design leads to reductions in welfare because of the myopic consumption behavior of consumers.

Finally, Policy 4, removing the doughnut hole for generics, yields a 10% decrease in enrollee spending and a 6% increase in insurer spending relative to the baseline. The end effect is that total spending is very slightly lower than the baseline, as enrollees substitute to generic drugs and away from branded drugs. Yet, consumer surplus increases 5% relative to the baseline. Thus, by causing individuals to substitute to generic drugs, this policy is welfare increasing while not raising total prescription spending.

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 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 doughnut hole early in the year change their behavior upon reaching the doughnut 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 doughnut hole, with a disproportionate drop for drugs that cost over \$150 and branded drugs. Moreover, the data can reject the presence of geometric

discounters with a low discount factor in favor of myopia.

Having established evidence in favor of myopia, we 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 find that consumers have significant price elasticities and complete myopia.

Our structural estimation approach has several limitations. We do not allow for any medical dynamics to treatment, we do not measure substitute therapies to drugs, we do not model imperfect physician agency, and 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 discontinuity evidence.

Last, we examine the impact of counterfactual preferences and policies. We find that closing the doughnut hole would raise total spending 10% or necessitate a 36% coinsurance for budget balancing. Doughnut hole coverage for generics only would be much less expensive.

References

Abaluck, J. and Gruber, J. (2011). Choice inconsistencies among 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. NBER Working Paper 19163.

Abaluck, J., Gruber, J., and Swanson, A. (2015). Prescription drug utilization under Medicare Part D: A dynamic perspective. NBER Working Working Paper 20976.

Baicker, K., Mullainathan, S., and Schwartzstein, J. (2012). Behavioral hazard in health insurance. NBER Working Paper 18468.

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

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. *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. Mimeo, Columbia University.

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. and Simon, K. (2008). Medicare Part D's effects on elderly patients' drug costs and utilization. *American Journal of Managed Care*, 14(11 Suppl):SP14–22.

Ketcham, J. D., Lucarelli, C., Miravete, E. J., and Roebuck, M. C. (2012). Sinking, swimming, or learning to swim in Medicare Part D. *American Economic Review*, 102(6):2639–2673.

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.

Magnac, T. and Thesmar, D. (2002). Identifying dynamic discrete decision processes. *Econometrica*, 70(2):801–816.

Medpac (2014). Report to the Congress: Medicare payment policy. Technical report, Medicare Payment Advisory Commission.

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

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. Mimeo, University of Iowa.

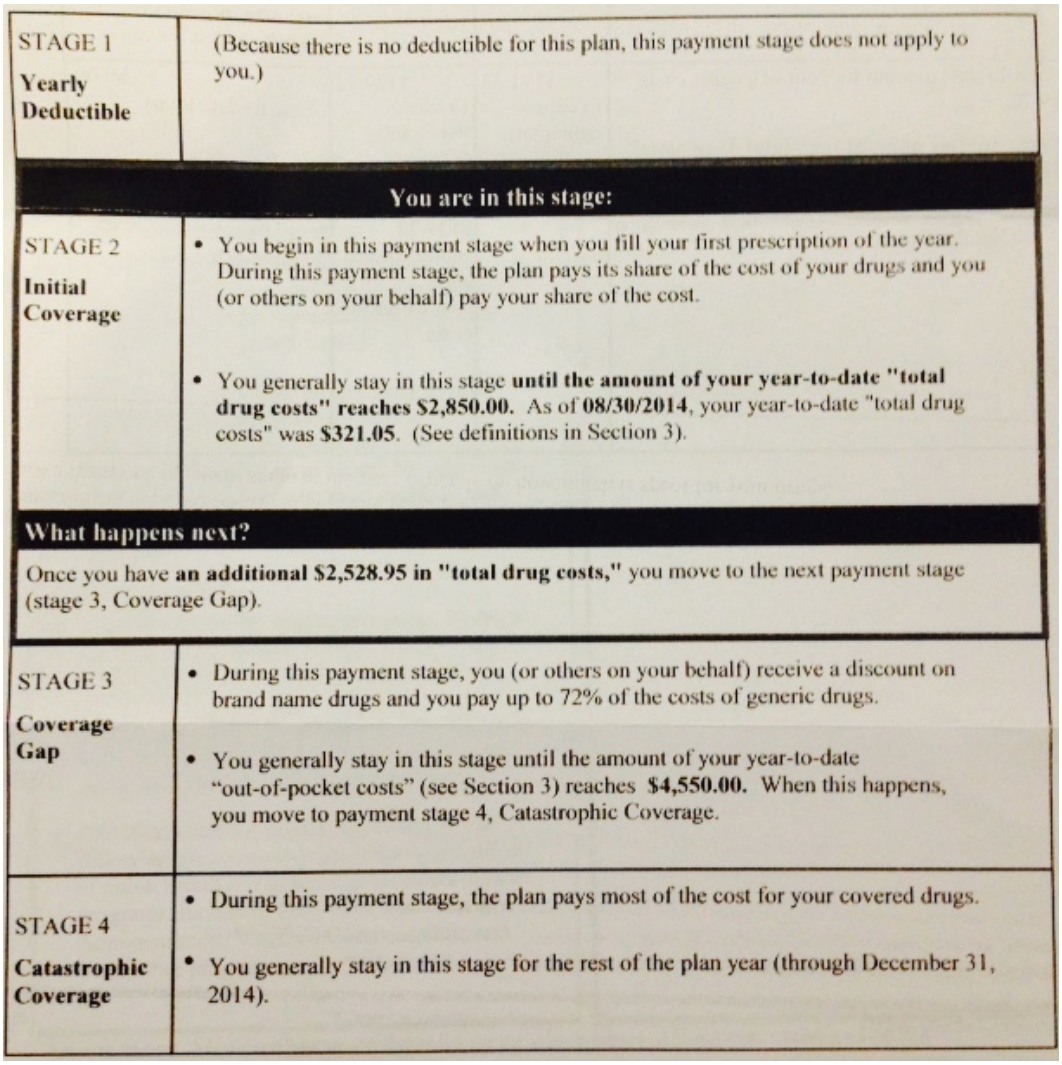
Strotz, R. H. (1956). Myopia and inconsistency in dynamic utility maximization. *Review of Economic Studies*, 23:165–80.

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: Extra Figures and Tables

Figure A1: Information provided to Part D enrollees on distance to doughnut 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).

Figure A2: Histogram of end-of-year drug spending for estimation and full sample

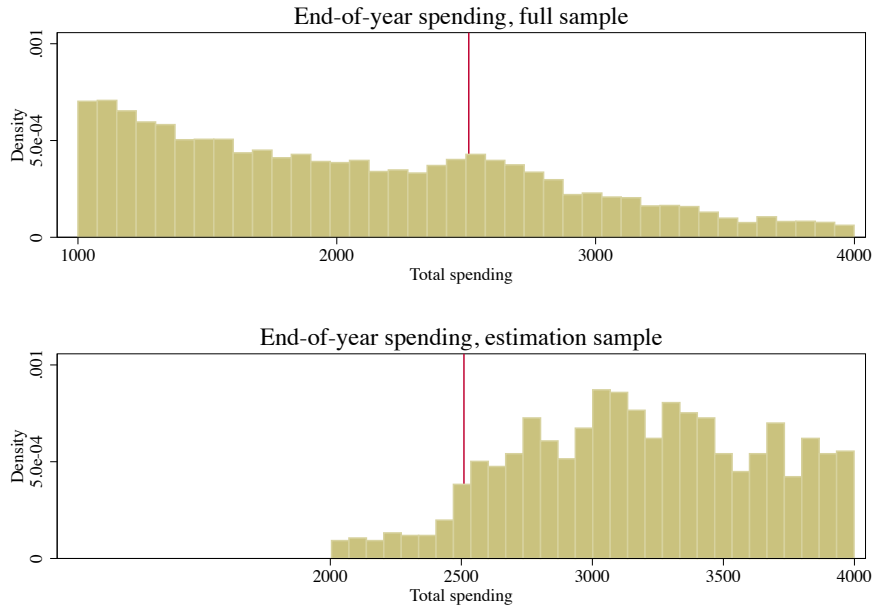


Figure A3: Mean drug spending on cumulative spending near catastrophic coverage start

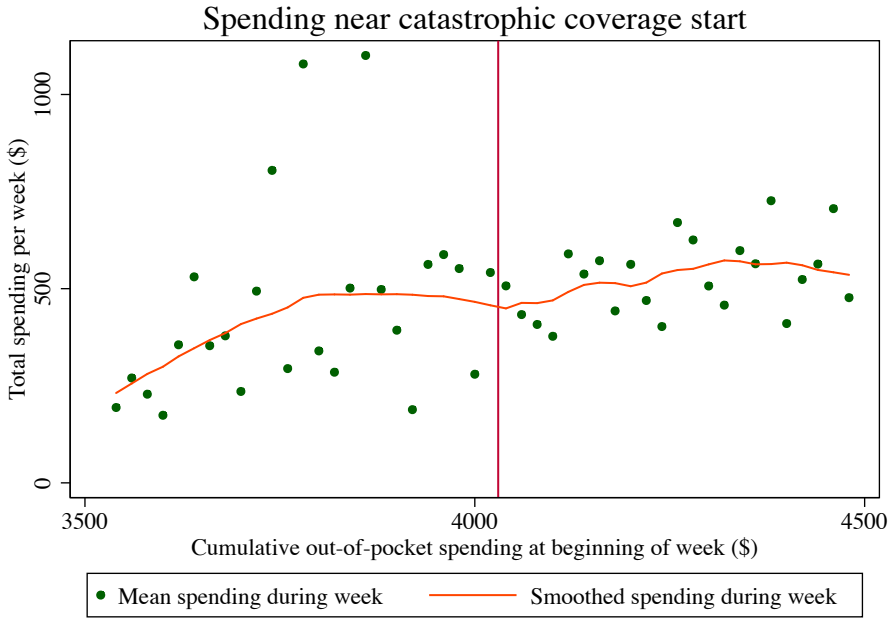


Table A1: Most common drugs in estimation sample

Drug name	Indication	Branded	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.

Table A2: 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	–	–	–
× doughnut hole	(.042)	(.060)			
Frac expensive	–	–	1.47**	.168*	−1.79**
			(.051)	(.065)	(.078)
Frac expensive	–	–	−.065	.170**	.056
× doughnut hole			(.042)	(.054)	(.084)
Doughnut 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.

Table A3: 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.

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

This section describes our method for enumerating all the elements in $\mathcal{L}(\hat{N}, \overline{N})$.[32] 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

[32] 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 here also.