THE HEALTH COSTS OF COST-SHARING*

Amitabh Chandra

Evan Flack

Ziad Obermeyer†

Abstract

The design of health insurance can cause large, abrupt, and arbitrary changes in drug prices. We use this fact to study what happens to patients' health when they stop taking their drugs. Thanks to a quirk in Medicare's drug coverage, 65-year-olds are as-if-randomly assigned a budget—a spending limit on drugs, beyond which they are responsible for 100% of the cost—as a function of their birth month. We find that an exogenous $100 per month decrease in this budget (a 24.4% change) causes mortality to increase by 0.0164 percentage points per month (13.9%). This estimate is robust to a range of falsification checks, and in the 97.4th percentile of 541 'placebo effects' formed in settings that are observably similar, but lack the policy quirk linking birth month to drug budgets. We make sense of this large effect in three ways. First, patients stop taking drugs that not only appear 'high-value' (e.g., blood pressure medications), but are also known to have withdrawal or 'rebound' effects. Harm from abruptly stopping these drugs can be large, dwarfing any foregone benefits of not taking the drug. Second, using machine learning, we identify patients at the highest risk of drug-preventable adverse events (e.g., heart attack). Contrary to the predictions of some economic models of behavior, high-risk patients cut back *more* than low-risk patients on precisely those drugs that would benefit them the most (e.g., statins). Third, patients appear largely unaware of the risks. In a survey, we find only one-third believe that missing their drugs for up to a month could have serious health consequences. We conclude that, far from curbing waste and moral hazard, cost-sharing causes patients to miss opportunities to purchase health at low cost ($11,321 per life-year).

Word Count (entire manuscript): 10,556 **JEL Codes**: I12, I13

*We acknowledge funding from the NSF Graduate Research Fellowship (Flack) and National Institute on Aging grant P01AG005842 (Obermeyer). We thank, without implicating, Amy Finkelstein, Ben Handel, Peter Hull, Jon Kolstad, Jon Gruber, Sendhil Mullainathan, Mohan Ramanujan, Julian Reif, Jon Skinner, and Josh Schwartzstein, for many comments that have improved our paper.

†Obermeyer (corresponding author): 2121 Berkeley Way, Berkeley, CA 94720, zobermeyer@berkeley.edu

I Introduction

Few patients take their drugs as prescribed (World Health Organization, 2003). Even in carefully controlled clinical trials, only 43–78% of doses are ever taken (Osterberg and Blaschke, 2005). Health insurance design often encourages this inconsistency: the non-linear contract structures of many drug plans create variation in out-of-pocket price over time, causing patients to stop then restart prescriptions abruptly (Einav et al., 2015; Einav and Finkelstein, 2018; Einav et al., 2018).

In general, price-driven interruptions in drug consumption have not been taken as cause for concern, at least by economists, because their welfare effects are theoretically ambiguous. Patients are known to stop taking apparently high-value medicines, like statins and beta-blockers (Brot-Goldberg et al., 2017; Choudhry et al., 2011; Einav et al., 2018), but there is no clear-cut evidence that these decisions harm their health.[1] So patients may be deciding, rationally, that costs outweigh benefits: they may judge that financial savings outweigh health gains, or have private information about their own treatment benefits. Indeed, patients themselves seem unbothered by drug interruptions. In a survey of Medicare-age patients taking medication (Figure I), we find that only one-third (33.5%) believe that missing their drugs for up to a month could have serious consequences, like hospitalization or death. A majority (53%) do not predict *any* negative health consequences from missing their drugs for up to a week—even simply feeling worse symptomatically.

The medical literature, by contrast, gives more cause for concern. An obvious reason is that interrupting a drug means forgoing potentially large health benefits shown in clinical trials. A more surprising reason is the existence of drug withdrawal or 'rebound' effects, which can cause harms that dwarf any foregone benefits. Consider a classic study that randomly alternated different doses of the beta-blocker propranolol (a drug for high blood pressure) with placebo, to find an optimal dose (Miller et al., 1975). Of 20 patients participating in what should have been a fairly mundane trial, 10 had adverse cardiac events, 6 of which were heart attack or death; *all* adverse events occurred in the placebo periods (4 of 44 total weeks), when the drug was abruptly withdrawn. In another study, researchers noted that for idiosyncratic reasons, patients taking statins as outpatients sometimes failed to get them durings hospitalizations (Heeschen et al., 2002). In the setting of heart attack, patients whose statins were interrupted had three-fold higher risk of death or repeat heart

[1] Previous studies show only effects on *proxies* for health, like hospitalizations or spending. For example, Chandra et al. (2010) find that price increases in costs for drugs led to increased hospitalizations, but did not study mortality.

attack than similar patients whose statins were continued—and 69% higher risk than patients who *never* took statins. Clinical guidelines now mandate initiating a statin immediately on diagnosis of heart attack (Cannon and Freeman, 2023). Rebound effects have also been documented for diuretics (Walma et al., 1997), corticosteroids (Jarad et al., 1999), insulin (Czosnowski et al., 2009), inhaled bronchodilators (Hancox, 2006), antithrombotics (Broderick et al., 2011), hormone replacement (Waddell et al., 1999), antihypertensives (Psaty et al., 1990), and most psychotropics (Horowitz et al., 2021). Guidelines thus suggest carefully tapering these drugs under medical supervision, rather than simply stopping them (Steinman and Reeve, 2023; Bain et al., 2008).

Despite these concerns, little is known about the consequences of drug interruptions in the real world, because formal study is difficult and potentially unethical. So outside of clever studies like the ones above, the literature consists largely of non-randomized comparisons—consistent vs. inconsistent patients—which are confounded. Practicing physicians are unlikely to notice even large effects, because adverse events are rare. Detecting a 35% increase in mortality from a baseline of 1% would require a physician with perfect recall and 30,000 patients with randomly assigned interruptions (vs. a typical panel of 2000 patients (Raffoul et al., 2016)).

We study the mortality effects of drug interruptions using a quirk in Medicare coverage that, abruptly and as-good-as randomly, assigns patients to higher or lower drug prices. This strategy, pioneered by Aron-Dine et al. (2015) and Kaplan and Zhang (2017), leverages variation linked to birth month. Every January, beneficiaries start by paying only 25% of drug costs out-of-pocket. If they exceed an annual budget cap of $2500, however, they enter the "donut hole," where they pay 100% of costs. Whether or not a beneficiary enters the donut hole is not random: it depends on prior consumption. Critically, however, annual plan thresholds are not pro-rated, if enrollment happens part-way through the year. And because eligibility for coverage begins in the month someone turns 65, birth month generates exogenous variation in drug budgets in the first year of enrollment. Early and late enrollees get the same "pre-donut budget" of $2500, before cost-sharing abruptly rises to 100%. But by the time a late enrollee spends her first dollar, an earlier enrollee has been spending down their budget for months, and approaching the donut hole. These early vs. late differences in donut hole exposure grow as spending accumulates over the course of the first year. Reassuringly though, "pre-treatment" characteristics are similar: we find good balance on demographics, drug consumption, drug spending, and mortality in the first 90 days of enrollment.

Variation in drug budget will affect drug consumption only for those whose spending is high enough to approach the budget cap. So our analysis focuses on a very specific set of "compliers": those with the highest drug spending in their first 90 days. Higher initial spending puts them on track to approach the budget cap if they enroll early in the year, but not if they enroll later.[2] By contrast, the budget cap should not affect lower-spending "never-takers," who will end the year far below $2500, even if they enroll early. Like Aron-Dine et al. (2015), we further focus on the month of December, when we expect budget differences across enrollment months to affect consumption the most: as spending accumulates for early enrollees, more begin to approach or enter the donut hole, while later enrollees are protected. Empirically, we find that by December, early-enrolling compliers are far more likely to enter the donut hole (5.2 p.p. for every month of earlier enrollment) and more likely to interrupt their drugs as a result (3.65 fewer drug-days for every month).

These exogenous budget changes cause large mortality differences. For each $100/month decrease in the pre-donut budget caused by enrollment month (on average, a 24.4% change in our sample), mortality increases by 0.0164 p.p. per month (13.9%).[3] We make sense of this large effect in two ways. First, like Einav et al. (2018); Brot-Goldberg et al. (2017), we find that patients interrupt drugs that are both "high-value," and known to have dangerous rebound effects: statins and antihypertensives, glucose-lowering agents, and inhalers and steroids for pulmonary disease. Second, we document a *positive* correlation between a beneficiary's benefit from a drug and their likelihood of interrupting that drug. For example, those at the highest predicted risk of heart attack and stroke interrupt their cardiovascular drugs (e.g., statins, antihypertensives) four times more than lower-risk patients (2.46 vs. 0.60 drug-days per $100 change in monthly pre-donut coverage).[4] Similar patterns exist for diabetes and respiratory drugs. These differences are unlikely to be explained solely by income: we see similar cutbacks in both high- and low-income zip codes. This is surprising,

[2] As noted above, initial spending is uncorrelated with enrollment month, meaning forward-looking behavior does not induce differential selection. This allows us to identify a similar group of compliers across enrollment months. A small handful (3%) of the highest spenders not only enter the donut hole, but also approach catastrophic coverage; we exclude them from the complier group, and discuss them in detail below and in Section II.B.

[3] We present our estimates in terms of this monthly budget, rather than two-stage least squares, because enrollment month affects mortality through a complex network of variables. Prices and quantities consumed in month t affect prices and quantities in $t+1$, and prices in $t+1$ affect quantities in t (see Appendix Figure A.1). Any of these could affect health outcomes at year-end, so choosing just one as the endogenous variable is difficult to defend. By contrast, the pre-donut budget is both interpretable, and a mechanical function of enrollment month: e.g., a $2500 budget gives a February enrollee $228/month and a September enrollee $628/month.

[4] There is strong evidence from the medical literature that such risk is a good proxy for treatment benefit, particularly for cardiovascular drugs, an assumption we discuss in detail in Section III.D.

at least from the point of view of economic models. Moral hazard predicts that the marginal drugs patients interrupt should be low-value; and well-informed patients with private information about their own health should not be interrupting drugs that benefit them.

To build confidence in our estimate, we begin by ruling out two obvious confounders: while earlier enrollees are of course a few months older than later enrollees, mortality differences are 10 times larger than those implied by age differences. Likewise, well-known health differences across birth seasons are far too small to explain our results (Doblhammer and Vaupel, 2001). We then conduct a range of additional analyses and falsification checks, all of which replicate our analysis in closely-related settings, and ensure mortality effects are only present when enrollment month affects drug budgets. Our first check is to compare the effect of enrollment month on mortality in the compliers, as above, to the effect in those with lower and higher initial drug spending in the population. We find that the lowest-spending 70% of our sample (never-takers), who remain far below the donut hole budget cap in all enrollment months, and show no enrollment month gradient in either drug consumption or mortality. By contrast, the highest-spending 3% spend *through* the donut hole and approach the "catastrophic coverage" (around $6000), where cost-sharing drops from 100% to near zero. Here, the effect of enrollment month reverses: earlier enrollees now pay less for drugs, thanks to the catastrophic coverage, while later enrollees pay more because they are still stuck in the donut hole. Among these "defiers," we find large, significant, and opposite-sign effects of enrollment month on drug consumption and mortality, just as we would expect. It is hard to imagine a confounder that so perfectly matches this interaction of enrollment month, initial spending level, and Medicare policy.

We further isolate the mortality effect temporally, in two ways. First, we track the effect of enrollment month on drug consumption and mortality over the months before and after Deccember of the first enrollment year. We find that the only significant difference in mortality is in December. This is exactly when differences in drug consumption across enrollment months peak, abruptly and discontinuously, driven by early enrollees holding out until prices reset in January.[5] We also take advantage of reforms to Medicare that began to close the donut hole, starting in 2011. Replicating our analysis before vs. after this change, we find significant negative mortality effects only before.

[5]These results complement our balance checks, which as noted above show no differences in mortality in the first three months of enrollment.

In our final, most comprehensive falsification test, we construct a large set of "placebo estimates," of the effect of enrollment month on mortality, in observably similar Medicare populations for whom enrollment month does not affect drug budgets. This includes, for example, 66 year-olds, who are not affected by the enrollment month quirk we exploit at age 65, dual eligibles who face no cost-sharing, and disabled beneficiaries whose enrollment timing is not driven by birth month. Replicating our analysis in each of these samples, across a range of calendar months, we find that our main estimate is in the 97.4th percentile of mortality effects—larger than 527 of 541 placebo estimates.

The large mortality effect we find fits with a growing literature in economics showing that health insurance design matters for health. Most closely related, Abaluck et al. (2021) find that switching into a plan with at partial (as opposed to no) donut hole coverage reduces mortality by 0.46 p.p. per year, or .038 p.p. per month. A rough calculation based on our estimates implies that full (as opposed to no) donut hole coverage would reduce mortality by 0.063 p.p. per month; considering we are comparing partial to full coverage, these effect sizes are quite similar.[6] More broadly, Miller et al. (2019) show Medicaid expansion reduced mortality (0.089 p.p./year), and Goldin et al. (2020) find similar effects from the Obamacare insurance tax incentive (0.09 p.p./year).

Cost-sharing, via coinsurance or deductibles, has been a cornerstone of health insurance design for decades, driven by worries about wasteful spending and moral hazard. Our results build on a growing body of literature suggesting that, far from reining in low-value care, cost-sharing causes patients to systematically miss opportunities to purchase health at very low cost (Baicker et al., 2015). We estimate that eliminating cost-driven drug interruptions would extend life at a cost of around $11,321 per life-year (vs. commonly-used thresholds for cost-effectiveness: $100-200,000 (Neumann et al., 2014))—and given that increased inpatient spending can offset any savings on drug costs, this number is likely to be an over-estimate (Chandra et al., 2010). We view these results optimistically: improving the design of prescription drug insurance gives policy makers a unique opportunity to achieve large health gains at little to no cost.

[6] To generate this, we instrument for donut hole entry using enrollment month, then regress mortality on predicted donut hole entry (all in December). Because this happens in compliers only, while Abaluck et al. (2021) deal with all enrollees, we scale our estimate by the fraction of compliers (0.27). For context, we estimate the mean change in coinsurance in our data at 45 p.p. (initial vs. donut hole), vs. 20 p.p. in Abaluck et al. (2021) (partial vs. no donut hole coverage). In other words, both the price difference and mortality effect are twice as large in our setting.

II Empirical Strategy

Studying interruptions in drug consumption presents a number of empirical challenges, because interruptions are not randomly assigned. Patients who interrupt consumption are quite different from those who do not. Perhaps the best evidence of this comes from randomized trials, in which every drug dose is carefully tracked. Patients with more drug interruptions are often found to have worse outcomes by the end of the trial—whether they have been receiving the active treatment or the placebo (Osterberg and Blaschke, 2005). This elegantly shows how simply comparing the health of patients with vs. without interruptions is confounded by the myriad unmeasured links between drug-taking behavior and health.

An idiosyncratic feature of Medicare's prescription plan design allows us to isolate a setting where interruptions are as-good-as-randomly assigned, due to abrupt changes in the price patients must pay for their drugs. Starting with the RAND health insurance experiment, and extending through work in the modern economics and medical literature, random assignment of higher drug prices is known to reduce drug consumption (Newhouse and Group, 1993; Choudhry et al., 2011; Finkelstein et al., 2012; Brot-Goldberg et al., 2017). Price increases in this setting affect drugs across the board, so this design does not allow us to study the effect of any one specific drug. However, it does let us investigate how interruptions in drug consumption more broadly affect health.

II.A Program Details and Identification Strategy

Since 2006, Medicare Part D has offered prescription drug coverage to seniors and disabled individuals in the US. Individuals can enroll in either stand-alone prescription drug plans (PDPs) that are offered alongside traditional medicare or Medicare Advantage (MA) plans where drug coverage is bundled with inpatient/outpatient care. The benefit's non-linear structure with respect to out-of-pocket costs is illustrated in Panel A of Figure II (Einav et al., 2015). Using the 2008 details to describe the plan, the calendar year begins with a deductible phase in which the beneficiary pays the entire cost of all drugs until she has spent $275. She then faces a 25% cost-sharing (coinsurance) rate that lasts until the total cost of all drugs purchased exceeds the budget cap of $2,510 (called the initial coverage limit). After this, the beneficiary falls into the coverage gap, or "donut hole" and again pays 100% of the cost of all drugs. Finally, after reaching $5,726 of total spending in the donut

hole, she enters the "catastrophic coverage" (which, to continue the analogy, represents the far side of the donut). Here, she pays only a 5% cost-sharing rate, or copays of $2.25 for generic or preferred drugs and $5.60 for other drugs.[7] The cutoff points for each coverage arm change slightly from year to year, as shown in Appendix Table A.1, but the basic structure remains the same. Importantly, starting in 2011, the donut hole began to close as a result of policy changes (fully closing in 2020), with cost-sharing rates for generic and branded drugs in the donut hole falling from 100 percent to 50% and 93%, respectively.

As first noted by Aron-Dine et al. (2015) and Kaplan and Zhang (2017), the spending limits that define cost-sharing are not pro-rated in the first year of enrollment. In other words, a person who enrolls in, say, February gets the exact same budget—e.g., $2510 of total spending before falling into the donut hole—as a person who enrolls in September. It is easy to see that a beneficiary's *monthly* budget will be far greater if they enroll in later calendar months: the same "pre-donut budget" of $2500 is spread over fewer months. This results in earlier enrollees being more likely to exceed their budget cap, landing them in the higher prices of the donut hole, while later enrollees are more likely to remain in the generous initial coverage phase with lower prices.

Individuals become eligible to enroll in Medicare Part D on the first day of their 65th birth month, meaning that enrollment month is primarily driven by birth month. If we believe that (i) birth month is as good-as-random, and (ii) people of different birth months select into enrollment months similarly, then the effect of any resulting variation in drug budgets on mortality can be identified. We test these three assumptions in detail below: first, via balance checks on range of health and utilization measures (Table II), to verify that enrollment months are observably similar. Second, we estimate a variety of "placebo estimates" of the effect of birth month and enrollment month on mortality, in a variety of settings where birth month and drug budgets are uncorrelated, to ensure there is no effect beyond the mechanical effect of being a month older or younger (Figure IV). Third, following Aron-Dine et al. (2015), we test whether beneficiaries select into enrollment

[7]The overall non-linear structure was largely the result of a political compromise balancing the desire to cover very sick beneficiaries (the catastrophic phase) with reducing the total cost of the program (the donut hole); a review is found in Oliver et al. (2004). Insurers may offer coverage that is "actuarially equivalent", or "enhanced" compared to the standard benefit. One common deviation from the standard design is to replace the deductible phase with uniform cost-sharing until the donut hole is reached. Additionally, most plans do not base cost-sharing on cost-sharing rates but rather copays based on drug tiers for each coverage arm. In practice, copays equate to roughly the same level of cost-sharing in each arm as the cost-sharing rates specified by the standard benefit (Einav et al., 2015).

months similarly across birth months (Appendix Figure A.2).[8] Like Aron-Dine et al. (2015), we prefer estimates based on enrollment month, as opposed to birth month, in our main specifications. Enrollment month more accurately measures the start of actual spending, as opposed to the start of eligibility, meaning it more accurately predicts drug prices. If we instead use birth month, we find similar, but less precise results (see Appendix Table A.2).

Since enrollment month affects prices, and prices might affect mortality via decreased drug quantities, it is tempting to estimate a two-stage least squares model. But this raises a thorny question: *which* price or quantity does enrollment month affect? Patients are well-known to change their consumption in response to both the current spot-price of a drug, and also the expected future price. In our setting, enrollment month changes both. Early-year enrollees will likely reduce utilization towards the end of the year (e.g. December) when their spot price is (quasi-randomly) high, but they may also reduce utilization earlier in the year (e.g. October) given their expected high end-of-year price. This in turn affects consumption earlier in the year, which also affects future price. Appendix Figure A.1 shows these messy causal relationships graphically. Therefore, using enrollment month to instrument for any one specific price or quantity will violate the exclusion restriction.

On the other hand, using an average or cumulative price or quantity risks smoothing over important temporal dynamics. As Aron-Dine et al. (2015) note, the end of the first calendar year of enrollment has several characteristics that make it different fron earlier months. As the year progresses, more and more beneficiaries enroll, yielding both larger samples and more identifying variation in enrollment month. Just as importantly, differences in price across enrollment months peak in December, as shown in Panel B of Figure II: early enrollees are most likely to land in the donut hole, while late enrollees are protected, before prices reset for all enrollees in January. Finally, our patient survey indicates that short drug interruptions are viewed as more innocuous than longer ones. This suggests that interruptions driven by high donut hole prices will increase as the time to January 1 decreases, and patients hold out for prices to reset (we will confirm this in Figure V).

[8] Unlike Medicare Part A, enrollment in Part D is not automatic. Beneficiaries can enroll during a 7-month long initial enrollment period that runs from three months before to three months after their 65th birth month. Coverage starts on the first day of the month after the individual enrolls, but not before the first day of the enrollee's birth month. If an individual chooses to enroll later, she faces a penalty of higher premiums for the remainder of her tenure on Part D. Of those enrolling within a year of turning 65, the majority enroll in their birth month (69%). Empirically, a small number of individuals enroll in the month before their birth month. This proportion is similar across birth months, and of similar magnitude in Aron-Dine et al. (2015).

As a result, we opt for a transparent, reduced-form estimate of the effect of enrollment month on December mortality shown in the following estimating equation:

$$\mathsf{Mortality}_i^{\mathsf{Dec}} = \theta_0 + \theta_1 \mathsf{EnrollMonth}_i + \mathsf{Year}_i\theta_2 + X_i\theta_3 + \epsilon_i \tag{1}$$

Enrollment month has the advantage of being causally upstream from all prices and quantities; it has the disadvantage of being somewhat difficult to interpret. Fortunately, it is easy to translate enrollment month into a highly policy-relevant quantity: a beneficiary's monthly "pre-donut budget", the average amount a beneficiary is allotted per month before entering the donut hole. For example, in 2008, the donut hole kicked in at $2,510 of total spending. So a February enrollee would have a monthly pre-donut budget of $228 (over the 11 months of coverage in her first year), while a September enrollee would have a monthly pre-donut budget of $628 (over 4 months). Concretely, this means we also estimate a version of Equation (1) where $\mathsf{EnrollMonth_i}$ is replaced with Budget_i. These upstream measures mean we do not need to specify one particular endogenous variable that influences mortality. Rather, we tie random variation in enrollment month directly to the overall cost-sharing policy it implies, and estimate its effect on mortality.[9] To summarize, our basic model relies on enrollment month to generate variation in the pre-donut budget and thus drug consumption. We assume (and test) that those with lower budgets will be more likely to reach the donut hole by the end of the year, and consequently, interrupt their drug consumption more, due to increased cost-sharing. We hypothesize that the cost-induced medication cutbacks could lead to adverse health events such as death.[10]

[9] We recognize that this budget too is not a perfectly sufficient statistic for plan design, as it ignores the re-introduction of coverage in the catastrophic phase, plan premiums, and plan-specific cost-sharing rates/copayments. However, we note that changes in price due to arm-specific prices or plan premiums are much smaller in magnitude than the initial coverage vs. donut hole difference. In addition, only 3% of beneficiaries ever even come close to the catastrophic phase, while the vast majority are affected by the donut hole, whether directly or via forward-looking behavior (Aron-Dine et al., 2015). So we believe that donut hole threshold, and by extension, the monthly pre-donut budget, is the key parameter to summarize coverage generosity.

[10] Of course, it's important to acknowledge that the donut hole might affect drug consumption differently from other cost-sharing policies (e.g., year-to-year changes in coinsurance rates) because it is such a sudden and temporary shock to drug expenditures. We can only study beneficiaries in their first year, so there could be later learning effects that modify the effect of the donut hole in later years. That said, Appendix Table D.3 shows that the demand response we measure is quite similar to other studies in the economics and medical literature, which gives us some reassurance that the effects we measure are generalizable outside of the specific setting we study.

II.B Focusing on Compliers

We depart from the prior literature with one important refinement to our analytic strategy. Variation in drug budget will affect drug consumption only for those whose spending is high enough to approach the donut hole budget cap. So our analysis hones in on a population whose drug consumption we expect to be most affected by enrollment month: those whose initial drug spending puts them on track to enter the donut hole if born earlier, but who are less likely to enter the donut hole if born later.

Definition 1. *Let W^* indicate a 65-year-old's beneficiary's (unobserved) spending on drugs at year-end, in the absence of cost-sharing, based on health and preferences. Let D be an indicator for whether she exceeds the budget cap and enters the donut hole by year-end. Let Z be enrollment month, restricted without loss of generality to a 2-month period where $Z = 0$ is earlier and $Z = 1$ later month. The population can be partitioned into 3 groups. For never-takers $\mathcal{N}$, W^* is so low that donut hole entry likelihood does not depend on Z: $Pr(D = 1|W^* \in \mathcal{N}, Z = 0) = Pr(D = 1|W^* \in \mathcal{N}, Z = 1)$. For compliers $\mathcal{C}$, monthly pre-donut budget increases and likelihood of entering the donut hole decreases in Z: $Pr(D = 1|W^* \in \mathcal{C}, Z = 0) > Pr(D = 1|W^* \in \mathcal{C}, Z = 1)$. For defiers $\mathcal{D}$, W^* is so high they predictably enter the donut hole, and likelihood of subsequently entering catastrophic coverage decreases in Z: $Pr(D = 1|W^* \in \mathcal{D}, Z = 0) < Pr(D = 1|W^* \in \mathcal{D}, Z = 1)$.*

To make Definition 1 concrete, we can consider three beneficiaries who will experience very different effects of enrollment month on drug consumption. First, consider a beneficiary whose preferred spending W^* is well below the donut hole budget cap (even if she enrolled in January). Such a never-taker's drug prices will not be affected by enrollment month, and her drug consumption will likewise be unchanged. Second, we have a beneficiary whose spending W^* approaches the donut hole threshold. If such a complier is born earlier in the year, she will likely end up reducing her drug consumption in response to the donut hole—either to avoid entering the donut hole (forward-looking behavior), or as a direct result of higher prices. If on the other hand she is born later in the year, her higher monthly pre-donut budget will shield her from these dynamics, and she will consume more drugs. Third, we turn to a beneficiary whose spending W^* is so high that she will certainly enter the donut hole, and then approach the catastrophic coverage phase, when prices return to near zero. This is rare—only the highest-spending 3% of the sample reach this point, as we will

show. But such a defier would show the exact opposite relationship between enrollment month and price as the compliers: very high-spending late enrollees now face higher, not lower, prices, because they do not have enough time to spend through the donut hole and reach the catastrophic phase. Thus early enrollees will consumer fewer drugs than later enrollees.

Failing to distinguish among these types is, at best, inefficient: a large fraction of the sample, the never-takers, will not have variation in drug consumption in response to the treatment. At worst, it generates unreliable estimates: the non-monotonicity of the defiers muddies the interpretation of the local average treatment effect of enrollment month on mortality. So we would like to find some way to partition the sample into types before estimating (1), to hone in on a homogeneous group of compliers with monotonic change in drug consumption in response to the instrument. But doing so is not straightforward because W^* is unmeasured, and realized spending W cannot be used to make the partitions: it is endogenous to enrollment month. For example, consider two beneficiaries who have spent \$1000 by year-end, one of whom enrolled in February and the other in September. Both end the year in the initial coverage phase, but grouping them together as never-takers would be an error: the former (who spent roughly \$100 per month to reach \$1,000) is likely a never-taker, but the latter (who spent \$300 per month, and would have spent \$3000 by year end had she been born in February) is likely a complier.

Proposition 1. *Since W^* is unmeasured, a proxy W' must be used to partition the sample into types $\{\mathcal{N}, \mathcal{C}, \mathcal{D}\}$. To be a good proxy, following Definition 1, $Pr(D = 1|W' \in \mathcal{N}, Z = 0) \approx Pr(D = 1|W' \in \mathcal{N}, Z = 1)$, $Pr(D = 1|W' \in \mathcal{C}, Z = 0) > Pr(D = 1|W' \in \mathcal{C}, Z = 1)$, and $Pr(D = 1|W' \in \mathcal{D}, Z = 0) < Pr(D = 1|W' \in \mathcal{D}, Z = 1)$. Further, to use the proxy for identification of the effect of Z, $W' \perp\!\!\!\perp Z$.*

One candidate proxy that might meet the criteria in Proposition 1 is a beneficiary's spending early in their Part D enrollment. The first criterion is that the proxy must partition the sample into groups with different impacts of enrollment month on donut hole entry. To test this, we take beneficiaries in a given enrollment month and bin them into percentiles of spending over their first 90 days. We then estimate the effect of enrollment rate on donut hole entry separately, for each (within-enrollment month) percentile bin of 90-day spending (shown in Appendix Figure A.3).[11]

[11] Using within-enrollment month percentile is attractive because is requires only a weaker assumption, that within-

We hypothesize that this will allow us to empirically partition the sample based on the likelihood of donut hole entry across enrollment months.

The second criterion in Proposition 1 is that the proxy W' must be uncorrelated with enrollment month. Early spending may fit this, if early enrollees do not display more forward-looking behavior than late enrollees, in anticipation of the donut hole budget cap. This is plausible, because all enrollees are learning about a highly complex program in their first few months of enrollment. On the other hand, given that forward-looking behavior is ubiquitous, it is also possible that even early spending is contaminated by responses to cost-sharing, and thus correlated with enrollment month. Using it as a proxy would thus induce differential sample selection across enrollment months and bias estimates of the enrollment month effect. While it is impossible to know exactly when beneficiaries begin to reduce their drug consumption in anticipation of donut hole limits, we can ask the empirical question of whether spending in the first few months of enrollment differs across enrollment months and verify it meets this criterion (we do so in Table II).

With this grouping strategy, we then have our primary estimating equation:

$$\mathsf{Mortality}_{i,j}^{\mathsf{Dec}} = \alpha_j + \mathsf{InitSpend}_i\beta_1 + (\mathsf{EnrollMonth}_i \times \mathsf{InitSpend}_i)\beta_2 + X_i\beta_3 + \epsilon_i \tag{2}$$

Here, $\mathsf{Mortality}_{i,j}^{\mathsf{Dec}}$ is December mortality, $\mathsf{InitSpend}_i$ are indicators for being in a specific range of within-enrollment month initial 90-day (from enrollment) spending percentile, and α_j is fixed effect capturing each individual Part D plan–year combination. Under the assumption that within each initial spending bin, pre-donut budget is quasi-randomly assigned (by enrollment month), estimates of β_2 in Equation 2 give unbiased estimates of the effect of enrollment month (and thus monthly pre-donut budget) on mortality.

II.C Data

Our main sample consists of a 20% random sample of first-time Medicare Part D enrollees in their initial enrollment period (birth month and three subsequent months) from 2007 to 2012. First,

enrollment month *ranking* of initial spending correlates with year-end prices. In an earlier version of this paper we instead used machine learning to predict one-year total spending using a range of early utilization measures (spending, number of fills, drug class indicators). We trained a model of one-year spending on early utilization out-of-sample (in the dual population that faces no cost-sharing) and applied it to our sample. However, we found that the simpler initial 90-day spending measure gave us enough precision to estimate our model.

we make sample restrictions common to the Part D literature. We subset to all beneficiaries who become eligible for Medicare because they turn 65, under the Old Age and Survivors Insurance (OASI).[12] This leaves us with 1,131,922 observations. We then remove all individuals dually-eligible for Medicaid or other low income subsidies, as they face low prices that do not change as a function of yearly spending (though we will later use this fact to construct falsification tests in this sample), which leaves us with 925,170 observations. We also remove all individuals that enroll in a deductible plan, as their initial claims vary with enrollment month due to the future price effects (Einav et al., 2015), bringing the sample to 605,502. A series of other minor subsets brings our sample to 557,999 beneficiaries.[13]

We make three additional exclusions with respect to the timing of enrollment month and death. First, in order to calculate mortality rates in December, we must exclude those who die before December 1. We will carefully check for and exclude mortality differences across enrollment months before December 1, which could indicate selection bias introduced by this exclusion. We also drop those who enroll in October and later: as Aron-Dine et al. (2015) note, these beneficiaries are still ramping up their drug consumption, whether because of new coverage or transitioning from a previous insurer. As a result, their December utilization is spuriously low compared to beneficiaries enrolling earlier in the year, who have reached steady state in terms of consumption by December. In addition, those born in October and later are legally allowed to enroll in January without penalty, because January is in their 4-month initial enrollment period (IEP). Empirically, January enrollment appears to be an outlier in terms of volume of patients enrolling, and January enrollees are observably different from all other enrollment months. So we follow Aron-Dine et al. (2015) here as well, and exclude them from our sample.[14] With these restrictions, our final analytic sample consists of 358,706 individuals.

Our analyses of specific drug filling decisions use the Medicare Part D drug claims made by beneficiaries in our sample, including fill date, total cost, out-of-pocket (OOP) cost, and 11-digit National Drug Code (NDC) identifiers. To classify drugs into clinically meaningful categories, we

[12] We exclude those who enroll in Medicare before age-64, for disability or end-stage renal disease.

[13] We include individuals in standalone PDPs and standard MA plans. The included MA plan types are HMO, HMO POS, Local PPO, Private FFS, and Regional PPO). We exclude individuals in special needs plans, those with non-standard ICL locations, and those not residing in the US 50 states or Washington, DC.

[14] We also find evidence that those born in January are less likely to delay enrollment, and that those born in other months (e.g. November) are more likely to delay enrollment to January as opposed to other months.

use the RxNorm and RxClass APIs to link NDCs to their corresponding Anatomical Therapeutic Chemical (ATC) codes, a hierarchical system for drug classification. This allows us to map, for example, a claim for Lipitor to the drug class of statins (HMG-CoA reductase inhibitors), within the 'lipid-modifying agents' category of cardiovascular drugs. We attempted to measure medical diagnoses, procedures, and other kinds of health care utilization besides drugs, using Medicare Parts A and B claims, including diagnoses, procedures, and admit/discharge dates. However, the subsample of individuals enrolled in standalone PDPs (non-MA), for whom we observe these data, is only about half our sample, so our analyses were largely under-powered.

III Results

III.A Sample Description

Summary statistics are shown in Table I. The sample is 90% white and mostly (60%) female. Roughly half of the sample is in a standalone PDP, with the other half in a Medicare Advantage plan. As our sample is relatively young, one-year mortality is low, 0.9 percentage points (p.p.), and one-year total spending is approximately $1,500. We also present the 10 most-used drug classes in our sample, which include statins, antihypertensives, diuretics, antidepressants, glucose-lowering drugs, and corticosteroids.

We check balance by regressing key 'pre-treatment' characteristics on enrollment month. Table II, Panel A, shows that estimates for race and sex are statistically and economically insignificant. Importantly, we also find that initial drug spending and consumption are balanced across enrollment months. As a more synthetic test, we also predict one-month mortality using all pre-treatment variables, and regress this on enrollment month.[15] This too is reassuring. Table II, Panel B shows a more direct balance check on 'pre-treatment' mortality: mortality in the first 3 months of enrollment, before cost-sharing begins to affect enrollment months differently (recall that this was already demonstrated in Panel A). We regress mortality in the 30, 60, and 90 days after enrollment on enrollment month. No estimates are significant, providing further evidence that baseline health

[15] Our model is estimated in a sample of 66+ year-olds, who we assume have the same relationship between mortality and covariates. The independent variables are measured over the first 90 days of coverage for the year (January-March), to mirror our main sample, and the dependent variable is mortality (April-December). We apply this model to generate predictions in our main sample, using covariates measured over the first 90 days of enrollment.

is similar between enrollment months. This also reassures us that conditioning our analytic sample on survival to December 1 does not introduce selection bias correlated with enrollment month.

Having established balance overall, we partition the sample into never-takers, compliers, and defiers (see Section II.B above). Appendix Figure A.3 shows the effect of enrollment month on the likelihood of entering the donut hole (y-axis), for each percentile of initial 90-day spending (x-axis).[16] In the first 70 percentiles of initial spending there is no relationship between enrollment month and donut hole entry, and only a very slight negative relationship in the 61-70th percentiles. Starting at the 71st percentile, we see significant and increasingly negative effects of enrollment month, as more and more earlier enrollees fall into the donut hole while later enrollees do not. Finally, at the 98th percentile, the relationship begins to reverse, and the last 2 percentile bins have a large positive effect, as earlier enrollees exit the donut hole and enter the catastrophic coverage.

Our primary goal in setting percentile cutoffs is to estimate Equation 2 in a homogeneous group of compliers, whose drug consumption monotonically increases in enrollment month. Because estimates for never-takers and defiers are not our primary focus, we are willing to tolerate some heterogeneity in these groups, to ensure homogeneity in the complier group. As a result, we assign the first 70 percentiles to the never-taker bin—even though, empirically, this is likely to include a small number of low-spending compliers from percentiles 61-70. We assign percentiles 98-100 to the defier bin. We emphasize that, on average, the 98th percentile experiences the impact of enrollment month more like a complier than a defier; but because our primary interest is in estimates from the complier bin, we set the cutoff very conservatively to minimize non-monotonicity of enrollment month in that bin. We performed a sensitivity analysis over a range of alternative cutoffs (e.g., starting the complier bin at the 61st percentile, or ending it at the 95th percentile) and found nearly identical results. Table I Column 2 and Table II Panel C show that the summary statistics and balance checks in compliers are very similar to those in the full population, apart from the fact that compliers spend and consume more drugs (as expected).

[16] There are not unique percentiles for the 60% of initial spending so we choose to estimate by decile for this group. However, we find precise null estimates for all 6 deciles.

III.B Mortality Effects of Drug Interruptions

Figure III summarizes our identification strategy and main result graphically. Panel A shows the proportion of beneficiaries that finish the year in each coverage arm, by initial spending bin. Like the finer percentile bins in Appendix Figure A.3, we observe a clear separation of (mostly) never-takers on the left, for whom enrollment month has no effect; compliers in the middle, where early enrollees end up in the donut hole more often (34.9% of February vs. 0.6% of September enrollees) but not the catastrophic coverage; and (mostly) defiers on the right, where early enrollees first enter the donut hole more often, then also enter the catastrophic coverage more often (the latter cannibalizes the share of early enrollees in the donut hole, hence the non-monotonic trend for the former). Panel B of Figure III shows that differences in donut hole exposure affect drug consumption, as measured by drug-days filled (i.e., the number days supplied, summed across all drugs). As expected, never-takers have no variation in days filled by enrollment month. Compliers fill substantially fewer days when they enroll early, a difference of 25.5 days between February and September. And defiers show the opposite pattern, with February enrollees filling 31.8 *more* days than September enrollees. (We take all estimates in this latter group with a grain of salt, due to the non-monotonicity of enrollment month effect on consumption, but emphasize that these beneficiaries are not the focus of our analysis.)

Panel C of Figure III shows the relationship between December mortality and enrollment month. For never-takers, there is no effect. For compliers, we find a significant and substantial negative relationship between mortality and enrollment month, that mirrors the increase in drug consumption shown in Panel B. And for defiers, we again see a large positive relationship, contrasting with compliers. Panel A of Table III summarizes these trends, via OLS estimates of β_2 from equation (2), both in terms of enrollment month (Column 2) and in terms of changes to monthly pre-donut budget (Column 3). Among compliers, a \$100 budget increase—24.4% relative to the average enrollment-month budget—leads to a mortality reduction of 0.0164 percentage points (p.p.), or 13.9% of the base mortality rate. The effect is negative, but small and insignificant, in never takers. The effect is positive and large in defiers; it is significant in the enrollment month specification, but not the pre-donut budget specification.

We might worry that variation in mortality across enrollment (or birth) months might be con-

founded by factors other than cost-sharing. Recall, though, that we have already seen two reassuring facts in this regard. First, no such variation in mortality exists early in enrollment, before cost-sharing and drug consumption trends start to diverge across enrollment months, as shown in Table II, Panel B. Second, any confounder would have to correlate not only to enrollment month, but also to exactly the *combination of enrollment month and spending patterns* we would expect, given Medicare policy: it would have to be present for medium-spenders, disappear for low-spenders, and then change sign for high-spenders. It is hard to imagine confounders that match these requirements.

Nevertheless, we directly address two known channels by which mortality variation linked to enrollment month could emerge, in the absence of the policy we study. The first and most obvious confounder is age: earlier enrollees are by construction older. But a simple calculation using US life tables (from the Social Security Administration) illustrates why this is unlikely to be a concern. Based on annual mortality rates for 65 vs. 66 years olds, we can roughly estimate the effect of being one month younger on monthly mortality: -0.001 p.p., or -0.76%.[17] As another point of comparison, the (imprecisely estimated) effect of enrollment month from the never-takers is -0.00036, or -0.68% (Table III). Just below (Section III.C), we describe how we construct a null distribution of effects of enrollment (or birth) months on mortality, across many observably similar samples lacking enrollment month effect on drug budgets. This yields a median estimate of -0.53%, which we interpret as the mean effect size of being one month younger in these populations. So these different methods of calculating age effects all give fairly consistent estimates of between -0.5-1%. They are all an order of magnitude smaller than the effect size from our main analysis, which captures the effect of enrollment month on cost-sharing plus the effect of being one month younger: -0.118 or -9.49%. (The equivalent estimate based on birth month is -10.8%, shown in Appendix Table B.1, Panel A.) This gives us confidence that aging is only a small part of the relationship we observe.

A second potential confounder is variation in health outcomes by birth season, which has been suggested to result from disease seasonality (Currie and Schwandt, 2013) or selection (Buckles and Hungerman, 2013).[18] Most of the literature focuses on peri- and post-natal outcomes, but two

[17] Using 2010 data, annual mortality was 1.59 p.p. for 65-year-olds and 1.74 p.p. for 66-year-olds, translating into monthly rates of roughly 0.133 p.p. and 0.146 p.p. The difference, -0.013 p.p., is approximately the effect of being one year younger on monthly mortality. We divide by 12 to get the effect of being one month younger, then divide by the mean base rate to get the percent decrease.

[18] While a full review of this literature is beyond the scope of our work, we refer the reader to Currie and Schwandt

large studies explore later-life outcomes in populations quite similar to our own. In a study of life expectancy at age 50, Doblhammer and Vaupel (2001) find that mortality peaks among May births and falls on either side. This cyclical pattern is stable across samples in the Northern Hemisphere (Austria and Denmark) and exactly reversed in Australia. In our sample, by contrast, mortality peaks among February births, while May births are solidly in the middle of the mortality distribution (Appendix Table A.2, Panel B). In addition, the birth month differences in our setting are orders of magnitude larger.[19] Another study by Zhang et al. (2019) considers cardiovascular disease and mortality over 40 years starting in middle age, among participants in the Nurses' Health Study. They find that while cardiovascular disease mortality peaks among April births, there was no difference in overall mortality across birth months. To the extent that both studies find some evidence of increased mortality among those born in April and May, this does not fit with our main finding of a large peak in early-year births, that varies widely as a function of drug spending.

III.C Falsification Tests

As we saw, the enrollment month effect varies widely across spending bins, following the idiosyncratic changes in drug budgets assigned by Medicare policy. We consider this our first falsification check, since few confounders could produce this very specific pattern. We develop this insight into a broader range of falsification checks: our analysis is closely tied to a specific Medicare policy, which is relevant only to a specific population at a specific time. So we replicate our analysis in a range of closely-related settings that lack these specific features linking enrollment month (or birth month) to drug budgets.

First, we follow the population of compliers over time, and show that the mortality effect of enrollment month is transient: it appears just as differences in prices and drug consumption peak at the end of the first calendar year, and disappears as soon as prices reset and drug consumption re-equalizes at the start of the second calendar year. Figure shows the last five months of the first calendar year of enrollment (August–December) and the first five months of their second calendar

(2013) for an excellent summary, and a very rigorous empirical exploration of mechanisms.

[19] Doblhammer and Vaupel (2001) find the maximum effect of one birth month change on life expectancy (at age 50, e_{50}) is 0.05-0.1 year, vs. an average of 27.5 years. Without access to the full life table, converting between e_{50} and mortality risk is not possible. But using the empirical rule of thumb from Pollard (2002), these e_{50} changes imply on the order of a 0.03% relative annual change in mortality risk (in each one-year age bin), compared to the effect of 10.8% we observe (0.0137 p.p. per birth month vs. base mortality of 0.127 p.p. for compliers.)

year (January–May) on the x-axis.[20] For each of these 10 months, we estimate the effect of enrollment month on three key outcomes. Panel A shows the effect of enrollment month on likelihood of entering the donut hole. This effect grows smoothly over time, as more beneficiaries enter the donut hole, then abruptly disappears in January, when all enrollment months re-enter the initial coverage phase. Panel B shows the effect of enrollment month on drug consumption. There is significant difference in drug consumption between earlier vs. later enrollees as early as September, and the difference steadily over the next 3 months. Then, in contrast to Panel A, there is an abrupt jump in December: while the enrollment month effect on donut hole entry is nearly identical from November to December, the effect on drug consumption more than doubles (1.75 drug-days for each later enrollment month in November, to 3.65 in December). Also in contrast to Panel A, drug consumption gradients reverse in January: as soon as prices reset, earlier enrollees—who have been waiting out the high prices of the donut hole, and filling fewer drug-days December as a result—now take advantage of lower prices to make up for their missed doses. Not only do earlier enrollees fill 2.86 *more* drug-days per enrollment month, they also fill *sooner* (0.30 days sooner for each earlier enrollment month, for those filling in January; estimate not shown). Finally, Panel C shows the effect of enrollment month on mortality. There is no significant mortality gradient across enrollment months from August to November—until December. Then, just as prices reset in January and earlier enrollees rush to fill their medications, the mortality effect attenuates: it is negative but insignificant in January, then disappears altogether from February onward. Overall, these results temporally tie mortality increases to the sudden increase in drug interruptions in December, likely driven by the facts shown in our survey (Figure I): patients view short interruptions in their drugs as largely innocuous, and many put off filling until prices reset.

Second, we leverage policy variation over time. In 2011, the donut hole began to close (described in Section II.B), reducing cost-sharing for earlier enrollees. This implies we should find larger mortality effects in 2010 and before. Panel B of Table III shows estimates of Equation (2) where instead of interacting enrollment month with bin of initial spending, we interact enrollment month with indicators for pre- vs. post-policy change. Over 2007-2010, we find that a \$100/month increase in the pre gap budget leads to a mortality reduction of 0.0104 p.p.; while there is no significant

[20] We could not produce stable estimates prior to August of year one, as sample size decreases with each month due to fewer enrollment months.

mortality effect in 2011-2012. Panel C repeats this analysis restricting to compliers. The effects are larger, with a mortality reduction of 0.0210 p.p. per \$100/month pre-gap budget increase, isolated to 2007-2010. So the effect of cost-sharing on morality is concentrated in the enrollees we expect, and over the time period when these enrollees are most affected by cost-sharing.

Third, we formalize the intuition of these placebo tests and construct an empirical "null distribution" of enrollment (or birth) month effects on mortality. We expect the estimates to be centered at zero (adjusting for the effect of age, which we discuss in detail below), and for our main estimate to be in the extreme tail. The placebo tests come from a large number of settings where Medicare beneficiaries look similar to the ones in our main sample, but lack the same idiosyncratic link between enrollment month and drug budgets. First, we extend the analysis in Figure V and follow our main sample further in time, estimating monthly effects from January of their second calendar year of enrollment until December of their fourth year (36 estimates: Appendix Figure B.1, Panel A). Second, we replicate the analysis in dual-eligibles, those on Medicaid or low-income subsidies, who do not face cost-sharing We pool all years together, then split into subsets defined by demographic factors or geography (46 estimates: Appendix Figure B.1, Panel B). Third, we broaden to a larger set of beneficiaries 66 years old and above—non-duals, duals, and disabled beneficiaries (ages 50-64), none of whom face the exact same cost-sharing as 65-year-old non-duals—whose initial spending makes them observably similar to our compliers (459 estimates: Appendix Figure B.1, Panel C).[21]

Figure IV pools results from all 541 falsification samples. We rank estimates of the enrollment month effect by magnitude on the x-axis.[22] The y-axis shows the cumulative fraction at least as large as x. The median estimate is -0.53%, likely reflecting the effect of age across these samples, as it is quite close to the age effect we estimated from other sources in Section III.B above (-0.76% from Social Security data, or -0.68% from the never-takers). The estimate from our main analysis, -9.49%, is shown in red. It is at the 97.4th percentile of mortality effects, larger in magnitude than 527 of 541 placebo estimates overall. In addition to the more tailored falsification tests above, this 'omnibus' test builds confidence that our observed mortality effect is in fact due to difference in the

[21] Because some of these populations lack an observable enrollment month, we use birth month as a proxy. More details are in the Appendix.

[22] Estimates are relative to the baseline mortality in each sample for comparability: falsification samples vary in their baseline mortality. Many samples are sicker than our 65 year-old non-dual primary sample, because of older age, lower income, and enrollment based on disability. As a result, mortality is higher, which if anything could make us better powered to detect a (spurious) effect.

pre-donut budget faced by enrollees, rather than a spurious correlation.

III.D Are These Effect Sizes Medically Plausible?

Table IV confirms in our setting a key finding of prior work (Einav et al., 2018; Brot-Goldberg et al., 2017): patients cut back many apparently high-value drugs, prescribed to prevent life-threatening adverse events: cardiovascular (e.g., statins, antihypertensives), glucose-lowering (e.g., insulin, oral hypoglycemics), and respiratory (e.g., steroids, inhalers). For each of these three classes, Column (1) shows the fraction of compliers who ever fill the drug. Column (2) shows estimates of how a $100 increase in monthly pre-donut budget affects drug-days consumed (in December). Earlier enrollees fill 5.2 more drug-days overall, with half of this total accounted for by cardiovascular, diabetes, and respiratory drugs (1.5, 0.6, and 0.5 more drug-days, respectively).[23]

Could short interruptions in these drugs cause mortality effects of the magnitude we find? Two sets of facts in the medical literature indicate that this is a possibility. First, it is a common misperception that drugs for chronic diseases begin to work slowly. Certainly, clinical trials for these drugs last many years—but particularly for rare outcomes like mortality, we should not conflate the time scale required to *measure* effects, with the time scale on which effects begin. Inspection of many published survival curves for chronic drugs shows that they start to diverge almost immediately, but only reach statistical significance after years. Statins provide an instructive example: the Kaplan-Meier curves in the landmark JUPITER trial begin to diverge at the origin (Ridker et al., 2008), and as Heeschen et al. (2002) note, appear to have protective effects in patients during hospitalizations for heart attack. This seems counter-intuitive, since the cholesterol-lowering effect of statins is to reduce atherosclerosis (heart disease) over long time periods. But statins also act via a range of other 'pleiotropic' mechanisms: they prevent blood clotting, and reduce inflammation and reactivity of blood vessels (Oesterle et al., 2017). These mechanisms yield large protective effects in the very short-term for patients with acute conditions like heart attack and stroke. Several other drugs have similar effects over multiple time-scales: the diabetes drug metformin lowers blood glucose in the short term, and also has a variety of longer-term anti-aging effects (Kulkarni et al., 2020); antibiotics

[23]We confirm in Appendix Table D.1 that deaths follow a similar pattern of drug consumption before they die: early (February-May) vs. late (June-September) enrollees have similar drug consumption over days 31-60 before death, then diverge sharply in the 1-30 days before death, with earlier enrollees filling 27.6 fewer drug-days than later enrollees, although estimates are imprecise.

used for COPD both treat acute infections and reduce long-term inflammation (Blasi et al., 2012). As a result, even short drug interruptions can mean large foregone treatment benefits, depending on the idiosyncratic time scale of the mechanisms mediating the drug's treatment benefit.

Second, simply thinking in terms of foregone benefit can be highly misleading. Drugs induce a complex set of physiological changes that put patients in a new physiological equilibrium—indeed, that is the point of taking drugs. Abruptly stopping a long-standing drug to which the body has adapted is well-known to precipitate a potentially dangerous set of withdrawal or rebound effects. This idea has entered the popular consciousness for opiates, but opiates are far from the only drug class with such effects: the list includes 7 of the 10 most commonly taken drugs in our sample (Table I)—statins (Heeschen et al., 2002), antihypertensives (Psaty et al., 1990), diuretics (Walma et al., 1997), antidepressants (Horowitz et al., 2021), corticosteroids (Jarad et al., 1999), and glucose-lowering drugs (Czosnowski et al., 2009). Because of the practical and ethical difficulties of studying drug interruptions directly, the best evidence for guidelines on drug tapering—which recommend slow transitions, under close medical supervision (Steinman and Reeve, 2023; Bain et al., 2008)—comes largely from older, small, idiosyncratic studies.[24] Our results add new weight to these recommendations.

If our results are unsurprising in light of the medical literature, they are far more surprising when viewed through the lens of economics: they run counter to the predictions of several economic models of behavior. If treatment benefit is as heterogeneous across individuals as most research indicates (Chandra and Skinner, 2012; Chandra and Staiger, 2020), under a variety of models, cutbacks should be concentrated among individuals in whom the drug is less clinically valuable. A moral hazard view would predict that the marginal drugs patients drop are disproportionately low-value. A Roy model of patient decision making with private information on heterogeneous treatment effects would likewise suggest that patients self-select into treatments that benefit them more. The corollary of this is that those with the highest potential benefit should be willing to pay more for a drug, and thus less likely to cut back when the price increases.

Our main results are only plausible if these models are incorrect, and drug interruptions also affect very high-value drugs. If by contrast interruptions are concentrated in low-treatment benefit

[24]Even for beta-blockers, perhaps the class of drugs for which there is the most evidence on rebound effects, a recent article can be summed up by its title: "Beta blocker rebound phenomenon is important, but we do not know its definition, incidence or optimal prevention strategies" (Koracevic et al., 2020).

patients, just as moral hazard or patient private information would predict, it would be hard to square with our results. If, on the other hand, high-benefit patients are interrupting their consumption, it could help make sense of the large treatment effects we observe. To distinguish between these views, it is critical to know *who* is interrupting their consumption, and estimate the health benefit an individual patient might get from a given treatment.

To develop such a measure, we use machine learning to form predictions on a patient's benefit from preventive medicines. The idea that drug benefit is proportional to the baseline risk of the outcomes the drug prevents is often simply assumed by doctors and in guidelines (e.g., the American College of Cardiology's 10-year risk calculator used to allocate treatments for cardiovascular disease). But there is strong and growing evidence to support this assumption, as well as specific clinical guidelines, for cardiovascular drugs. For example, the JUPITER, HOPE-3, CARDS, and ASCOT trials (reviewed by Bibbins-Domingo et al. (2016)) show 30-50% larger absolute risk reductions from statins in groups with higher predicted risk of heart disease, whether defined by age, diagnosed risk factors (e.g., diabetes), or biomarkers (e.g., LDL, CRP). Studies of polygenic risk scores show similar heterogeneity, with higher-risk participants getting nearly three times the absolute risk reduction (Natarajan et al., 2017). There is less strong evidence, but medical consensus and biological plausibility, for diabetic and respiratory drugs.[25]

So we identify three important drug classes G—cardiovascular, diabetes, respiratory—and compile a list of observable adverse outcomes, Y_G, the drugs are prescribed to prevent: heart attack and stroke for cardiovascular medicines, diabetic complications (e.g., foot amputation) for diabetes medicines, and respiratory failure. We form separate predictive models for each outcome, using a beneficiary's initial 90-day claims to predict the likelihood of adverse events over the next 270 days. These models are trained on an entirely separate sample of dual-eligible 66+ year olds, to ensure our predictions are out-of-sample. We restrict to those who are not taking class G (e.g., when predicting risk of heart attack or stroke, we exclude patients on statins), to obtain a prediction on the risk of complications if untreated.[26] Additional details are in Appendix C.

[25]Even if this model is far from optimal, in the sense that it captures 'true' treatment heterogeneity, if patients or doctors believe that high risk equates to high benefit, this measure will identify patients who *believe* they would benefit from a given treatment. We view this too as a useful fact to understand.

[26]In other words, we wish to predict $Y_{G=0}$, not $Y_{G=1}$ in potential outcomes notation, as our proxy for the benefit of drug G Naturally this choice of prediction target also induces selection bias, as documented in Mullainathan and Obermeyer (2022), who use machine learning to predict the yield of testing for heart attack in the tested then validate the model in the untested. Building on that work, Appendix Figure C.1 shows that true risk rises monotonically in

We then estimate the effect of changes in the pre-donut drug budget on drug class-specific consumption in the compliers, interacted with predicted risk. We specify risk as a simple indicator indexing the top one-third of the sample, based on where risk begins to increase rapidly (see Appendix Figure C.1). Table IV shows that the highest risk beneficiaries have elasticities at least as high, if not higher, for all three major drug classes. This trend is especially pronounced for cardiovascular drugs: for each \$100/month budget increase, low-risk patients fill 0.598 more cardiovascular drug-days, while high-risk patients fill 2.46 more. This finding is incompatible with a moral hazard model of behavior, or patient private information: either way, those at high risk of a cardiovascular event should have the most inelastic demand for treatment, proportional to their benefit from the drug. We find similar, although less pronounced, trends for diabetes and respiratory drugs. Importantly, Appendix Table D.2 demonstrates that these effects are of similar magnitude in high- and low-income zip codes alike. So while socioeconomically disadvantaged patients may have both worse health and less ability to pay for treatments, our results are unlikely to be driven by socioeconomic factors alone.[27]

IV Conclusion: Errors and Misinformation

Medicine and economics share a fundamental respect for individuals' preferences and decision-making. Our results, alongside a nascent literature in both fields, raise a dilemma for this perspective, if patients are making decisions that are very far from optimal. A rough calculation highlights the problem. Starting with the effect of enrollment month on cumulative drug spending (from September-December, when drug-day differences begin to emerge across enrollment months), we divide by the effect of enrollment month on mortality to obtain an estimate of the implied life-year valuation: \$11,321 (95% confidence interval: \$6,195-73,858).[28] Another way to look at the patient calculus is that, at the usual life-year valuation of \$100,000 per year, a 65-year-old complier in our

predicted risk for the treated just as the untreated, establishing face validity.

[27] Of course, there is variation in income within zip codes, often quite a bit, so this does not by any means rule out income effects or liquidity constraints. However, to the extent that we see similar behaviors in rich and poor areas alike, it forms some upper bound on how important these effects can be on average.

[28] This is based on the full cost of the drug, which is an upper bound on how much patients actually pay. We estimate a two-stage-least squares (2SLS) regression of December mortality on (instrumented) spending from October-December of year 1, in compliers. The inverse of this estimate is then "dollars per life." We then divide the dollars per life estimate by average life expectancy at 65 from Social Security data, weighted by the proportion of males/females in our sample, to estimate life-year valuation. Finally we calculate the implied life-year valuation at the bounds of the 95% confidence interval from the 2SLS coefficient.

sample would have to believe that she had at most 2.19 years left to live. In our sample, at a median follow-up period of 5 years, 93.2% of compliers are still alive; average life expectancy in the general population at age 65 is 19.2 years. This is hard to square with the idea that patients are equalizing marginal benefit with marginal cost of drugs, and suggests that the price elasticity of demand is an insufficient statistic for welfare, as has been noted by both Baicker et al. (2015) and Einav and Finkelstein (2018). This idea, termed 'behavioral hazard,' has far-reaching implications for the design of health insurance, particularly as insurers place more emphasis on cost-sharing.

Behavioral economics has identified several distortions that could contribute to our results. Costs, for example, might be over-weighted relative to their true value. If a patient arrives at the pharmacy counter to find that her drug basket has shot up in cost relative to her expectations, it could cause costs to be salient and overweighted (Bordalo et al., 2013, 2020). If costs deviate from previously set reference points, they may be viewed as losses and thus overweighted (Kahneman and Tversky, 1979). Present bias (Laibson, 1997; O'Donoghue and Rabin, 1999) could likewise cause patients to overweight present costs over future benefits. Alternatively, patients could be relying on heuristics—like filling *the* most important drug, dropping the most expensive drug—effectively substituting simpler problems for the more difficult full calculation of marginal costs and benefits (Tversky and Kahneman, 1974). Or patients could disengage from the cost-benefit calculus altogether, because of inattention (Handel and Schwartzstein, 2018; Gabaix, 2019), choice fatigue (Augenblick and Nicholson, 2016; Iyengar and Kamenica, 2010), or judging the problem as unsolvable and simply giving up (Ackerman and Thompson, 2017), could work similarly. Exploring whether these factors might be at play is a fruitful direction for future work in behavioral science, with potentially large real-world impact.

It is also worth considering a simpler explanation: patients could simply be misinformed. The survey of Medicare patients taking medications we fielded, while small, is perhaps the first evidence of its kind on how patients view short drug interruptions.[29] The 200 patients we surveyed were quite similar to the compliers in our sample: the median respondent was on at least 5 medications, and 75% indicated they had hypertension or high cholesterol. In general, these patients view short

[29]We partnered with the firm Survey Healthcare Global (SHG) to recruit patients ages 61-70 who report taking at least one medication, from a panel of 600,000+ patients and caregivers maintained by the firm. The survey, which took place in July 2022, asked patients about the consequences of drug interruption: how long before the risk of a serious health problem increased, and what kind of health problem they would worry about. The survey took less than 5 minutes to complete and had a 100% completion rate.

interruptions as innocuous: two-thirds doubt any acute events (hospitalization, deaths) would result from even a month-long interruption. Most cannot imagine *any* issues with missing their drugs for a week. These results grounds our empirical result, that drug interruptions peak sharply in December, in patient beliefs, and shows why the phenomenon is so temporary: patients are willing to hold out for low prices on the horizon in January. Our results argue strongly that this view is mistaken.

Ultimately, the decision to ingest a drug lies with the patient. However our results suggest that both physicians and policy-makers are missing opportunities to improve the architecture of these decisions. Policy-makers should remember that drug cost-sharing policies have major implications for patient health, as well as health are costs. And physicians should remind their patients that, for a variety of chronic medications, even short interruptions can be deadly.

Author Affiliations

Chandra: Harvard Kennedy School, Harvard Business School, and NBER.

Flack: Stanford University.

Obermeyer: University of California, Berkeley School of Public Health and NBER.

References

Abaluck, J., Caceres Bravo, M., Hull, P., and Starc, A. (2021). Mortality Effects and Choice Across Private Health Insurance Plans. *The Quarterly Journal of Economics*, 136(3):1557–1610.

Ackerman, R. and Thompson, V. A. (2017). Meta-Reasoning: Monitoring and Control of Thinking and Reasoning. *Trends in Cognitive Sciences*, 21(8):607–617.

Aron-Dine, A., Einav, L., Finkelstein, A., and Cullen, M. (2015). Moral Hazard in Health Insurance: Do Dynamic Incentives Matter? *The Review of Economics and Statistics*, 97(4):725–741.

Augenblick, N. and Nicholson, S. (2016). Ballot Position, Choice Fatigue, and Voter Behaviour. *The Review of Economic Studies*, 83(2):460–480.

Baicker, K., Mullainathan, S., and Schwartzstein, J. (2015). Behavioral Hazard in Health Insurance. *The Quarterly Journal of Economics*, 130(4):1623–1667.

Bain, K. T., Holmes, H. M., Beers, M. H., Maio, V., Handler, S. M., and Pauker, S. G. (2008). Discontinuing Medications: A Novel Approach for Revising the Prescribing Stage of the Medication-Use Process: DISCONTINUING MEDICATIONS. *Journal of the American Geriatrics Society*, 56(10):1946–1952.

Bibbins-Domingo, K., Grossman, D. C., Curry, S. J., Davidson, K. W., Epling, J. W., García, F. A. R., Gillman, M. W., Kemper, A. R., Krist, A. H., Kurth, A. E., Landefeld, C. S., LeFevre, M. L., Mangione, C. M., Phillips, W. R., Owens, D. K., Phipps, M. G., and Pignone, M. P. (2016). Statin Use for the Primary Prevention of Cardiovascular Disease in Adults: US Preventive Services Task Force Recommendation Statement. *JAMA*, 316(19):1997–2007.

Blasi, F., Mantero, M., and Aliberti, S. (2012). Antibiotics as immunomodulant agents in COPD. *Current Opinion in Pharmacology*, 12(3):293–299.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2013). Salience and Consumer Choice. *Journal of Political Economy*, 121(5):803–843.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2020). Memory, Attention, and Choice. *The Quarterly Journal of Economics*, 135(3):1399–1442.

Broderick, J. P., Bonomo, J. B., Kissela, B. M., Khoury, J. C., Moomaw, C. J., Alwell, K., Woo, D., Flaherty, M. L., Khatri, P., Adeoye, O., Ferioli, S., and Kleindorfer, D. O. (2011). Withdrawal of Antithrombotic Agents and Its Impact on Ischemic Stroke Occurrence. *Stroke*, 42(9):2509–2514.

Brot-Goldberg, Z. C., Chandra, A., Handel, B. R., and Kolstad, J. T. (2017). What does a Deductible Do? The Impact of Cost-Sharing on Health Care Prices, Quantities, and Spending Dynamics. *The Quarterly Journal of Economics*, 132(3):1261–1318.

Buckles, K. S. and Hungerman, D. M. (2013). Season of Birth and Later Outcomes: Old Questions, New Answers. *The review of economics and statistics*, 95(3):711–724.

Cannon, C. and Freeman, M. (2023). Low density lipoprotein-cholesterol (LDL-C) lowering after an acute coronary syndrome. In *UpToDate*.

Chandra, A., Gruber, J., and McKnight, R. (2010). Patient Cost-Sharing and Hospitalization Offsets in the Elderly. *American Economic Review*, 100(1):193–213.

Chandra, A. and Skinner, J. (2012). Technology Growth and Expenditure Growth in Health Care. *Journal of Economic Literature*, 50(3):645–680.

Chandra, A. and Staiger, D. O. (2020). Identifying Sources of Inefficiency in Healthcare. *The Quarterly Journal of Economics*, 135(2):785–843.

Choudhry, N. K., Avorn, J., Glynn, R. J., Antman, E. M., Schneeweiss, S., Toscano, M., Reisman, L., Fernandes, J., Spettell, C., Lee, J. L., Levin, R., Brennan, T., and Shrank, W. H. (2011). Full Coverage for Preventive Medications after Myocardial Infarction. *New England Journal of Medicine*, 365(22):2088–2097.

Currie, J. and Schwandt, H. (2013). Within-mother analysis of seasonal patterns in health at birth. *Proceedings of the National Academy of Sciences*, 110(30):12265–12270.

Czosnowski, Q. A., Swanson, J. M., Lobo, B. L., Broyles, J. E., Deaton, P. R., and Finch, C. K. (2009). Evaluation of glycemic control following discontinuation of an intensive insulin protocol. *Journal of Hospital Medicine*, 4(1):28–34.

Doblhammer, G. and Vaupel, J. W. (2001). Lifespan depends on month of birth. *Proceedings of the National Academy of Sciences of the United States of America*, 98(5):2934–2939.

Einav, L. and Finkelstein, A. (2018). Moral Hazard in Health Insurance: What We Know and How We Know It. *Journal of the European Economic Association*, 16(4):957–982.

Einav, L., Finkelstein, A., and Polyakova, M. (2018). Private Provision of Social Insurance: Drug-Specific Price Elasticities and Cost Sharing in Medicare Part D. *American Economic Journal: Economic Policy*, 10(3):122–153.

Einav, L., Finkelstein, A., and Schrimpf, P. (2015). The Response of Drug Expenditure to Non-linear Contract Design: Evidence from Medicare Part D. *The Quarterly Journal of Economics*, 130(2):841–899.

Finkelstein, A., Taubman, S., Wright, B., Bernstein, M., Gruber, J., Newhouse, J. P., Allen, H., Baicker, K., and Group, O. H. S. (2012). The Oregon Health Insurance Experiment: Evidence from the First Year*. *The Quarterly Journal of Economics*, 127(3):1057–1106.

Gabaix, X. (2019). Behavioral inattention. *Handbook of Behavioral Economics: Applications and Foundations 1*, 2:261–343.

Goldin, J., Lurie, I. Z., and McCubbin, J. (2020). Health Insurance and Mortality: Experimental Evidence from Taxpayer Outreach. *The Quarterly Journal of Economics*, 136(1):1–49.

Hancox, R. J. (2006). Concluding remarks. *Clinical Reviews in Allergy & Immunology*, 31(2):279–287.

Handel, B. and Schwartzstein, J. (2018). Frictions or Mental Gaps: What's Behind the Information We (Don't) Use and When Do We Care? *Journal of Economic Perspectives*, 32(1):155–178.

Heeschen, C., Hamm, C. W., Laufs, U., Snapinn, S., Bohm, M., and White, H. D. (2002). Withdrawal of Statins Increases Event Rates in Patients With Acute Coronary Syndromes. *Circulation*, 105(12):1446–1452.

Horowitz, M. A., Jauhar, S., Natesan, S., Murray, R. M., and Taylor, D. (2021). A Method for Tapering Antipsychotic Treatment That May Minimize the Risk of Relapse. *Schizophrenia Bulletin*, 47(4):1116–1129.

Iyengar, S. S. and Kamenica, E. (2010). Choice proliferation, simplicity seeking, and asset allocation. *Journal of Public Economics*, 94(7):530–539.

Jarad, N. A., Wedzicha, J. A., Burge, P. S., and Calverley, P. M. A. (1999). An observational study of inhaled corticosteroid withdrawal in stable chronic obstructive pulmonary disease. *Respiratory Medicine*, 93(3):161–166.

Kahneman, D. and Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2):263–291.

Kaplan, C. M. and Zhang, Y. (2017). Anticipatory Behavior in Response to Medicare Part D's Coverage Gap. *Health Economics*, 26(3):338–351.

Koracevic, G., Micic, S., Stojanovic, M., Tomasevic, M., Kostic, T., Velickovic Radovanovic, R., Lovic, D., Djordjevic, D., Randjelovic, M., Koracevic, M., et al. (2020). Beta blocker rebound phenomenon is important, but we do not know its definition, incidence or optimal prevention strategies. *Hypertension Research*, 43(7):591–596.

Kulkarni, A. S., Gubbi, S., and Barzilai, N. (2020). Benefits of Metformin in Attenuating the Hallmarks of Aging. *Cell Metabolism*, 32(1):15–30.

Laibson, D. (1997). Golden Eggs and Hyperbolic Discounting. *The Quarterly Journal of Economics*, 112(2):443–478.

Miller, R. R., Olson, H. G., Amsterdam, E. A., and Mason, D. T. (1975). Propranolol-Withdrawal Rebound Phenomenon: Exacerbation of Coronary Events after Abrupt Cessation of Antianginal Therapy. *New England Journal of Medicine*, 293(9):416–418.

Miller, S., Johnson, N., and Wherry, L. R. (2019). Medicaid and Mortality: New Evidence from Linked Survey and Administrative Data. Technical Report w26081, National Bureau of Economic Research.

Mullainathan, S. and Obermeyer, Z. (2022). Diagnosing physician error: A machine learning approach to low-value health care. *The Quarterly Journal of Economics*, 137(2):679–727.

Natarajan, P., Young, R., Stitziel, N. O., Padmanabhan, S., Baber, U., Mehran, R., Sartori, S., Fuster, V., Reilly, D. F., Butterworth, A., Rader, D. J., Ford, I., Sattar, N., and Kathiresan, S. (2017). Polygenic risk score identifies subgroup with higher burden of atherosclerosis and greater relative benefit from statin therapy in the primary prevention setting. *Circulation*, 135(22):2091–2101.

Neumann, P. J., Cohen, J. T., and Weinstein, M. C. (2014). Updating Cost-Effectiveness — The Curious Resilience of the $50,000-per-QALY Threshold. *New England Journal of Medicine*, 371(9):796–797.

Newhouse, J. P. and Group, I. E. (1993). Free for All?: Lessons from the RAND Health Insurance Experiment.

O'Donoghue, T. and Rabin, M. (1999). Doing It Now or Later. *American Economic Review*, 89(1):103–124.

Oesterle, A., Laufs, U., and Liao, J. K. (2017). Pleiotropic Effects of Statins on the Cardiovascular System. *Circulation Research*, 120(1):229–243.

Oliver, T. R., Lee, P. R., and Lipton, H. L. (2004). A Political History of Medicare and Prescription Drug Coverage. *The Milbank Quarterly*, 82(2):283–354.

Osterberg, L. and Blaschke, T. (2005). Adherence to Medication. *New England Journal of Medicine*, 353(5):487–497.

Pollard, J. (2002). Improving Mortality: A Rule of Thumb and regulatory Tool. *Journal of Actuarial Practice 1993-2006*.

Psaty, B. M., Koepsell, T. D., Wagner, E. H., LoGerfo, J. P., and Inui, T. S. (1990). The Relative Risk of Incident Coronary Heart Disease Associated With Recently Stopping the Use of -Blockers. *JAMA*, 263(12):1653–1657.

Raffoul, M., Moore, M., Kamerow, D., and Bazemore, A. (2016). A Primary Care Panel Size of 2500 Is neither Accurate nor Reasonable. *The Journal of the American Board of Family Medicine*, 29(4):496–499.

Ridker, P. M., Danielson, E., Fonseca, F. A., Genest, J., Gotto, A. M., Kastelein, J. J., Koenig, W., Libby, P., Lorenzatti, A. J., MacFadyen, J. G., Nordestgaard, B. G., Shepherd, J., Willerson, J. T., and Glynn, R. J. (2008). Rosuvastatin to Prevent Vascular Events in Men and Women with Elevated C-Reactive Protein. *New England Journal of Medicine*, 359(21):2195–2207.

Steinman, M. and Reeve, E. (2023). Deprescribing. In *UpToDate*.

Tversky, A. and Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, 185(4157):1124–1131.

Waddell, T. K., Rajkumar, C., Cameron, J. D., Jennings, G. L., Dart, A. M., and Kingwell, B. A. (1999). Withdrawal of hormonal therapy for 4 weeks decreases arterial compliance in postmenopausal women. *Journal of Hypertension*, 17(3):413.

Walma, E., Dooren, C. v., Prins, A., Does, E. v. d., and Hoes, A. (1997). Withdrawal of long term diuretic medication in elderly patients: a double blind randomised trial. *BMJ*, 315(7106):464–468.

World Health Organization (2003). Adherence to long-term therapies : evidence for action. Technical report, World Health Organization. number-of-pages: 196.

Young, B. A., Lin, E., Von Korff, M., Simon, G., Ciechanowski, P., Ludman, E. J., Everson-Stewart, S., Kinder, L., Oliver, M., Boyko, E. J., and Katon, W. J. (2008). Diabetes Complications Severity Index and Risk of Mortality, Hospitalization, and Healthcare Utilization. *The American journal of managed care*, 14(1):15–23.

Zhang, Y., Devore, E. E., Strohmaier, S., Grodstein, F., and Schernhammer, E. S. (2019). Birth month, birth season, and overall and cardiovascular disease mortality in US women: prospective cohort study. *BMJ*, 367:l6058.

Tables

TABLE I
SAMPLE DESCRIPTIVE STATISTICS

	(1)	(2)
	Mean (Standard Deviation)	
	Entire Sample	Compliers
Panel A: Demographics, Spending, Health		
White (%)	89.6 (30.5)	92 (27.1)
Female (%)	59.2 (49.1)	58.8 (49.2)
One-year total spending ($)	1,478 (2,789)	2,828 (1,910)
One-year mortality (p.p.)	0.873 (9.3)	1.16 (10.7)
Panel B: Top 10 Drug Classes (% on class)		
Lipid-lowering drugs (statins)	34 (47.4)	59.2 (49.1)
ACE inhibitors	20.5 (40.4)	30 (45.8)
Beta blockers	18.9 (39.2)	31.4 (46.4)
Thiazide diuretics	18 (38.4)	26.7 (44.2)
Antidepressants	14.1 (34.8)	24.8 (43.2)
Corticosteroids	13 (33.6)	22.4 (41.7)
Acid blockers (GERD)	12.4 (33)	24.3 (42.9)
Antiiinfectives	11.6 (32)	17.4 (37.9)
Glucose-lowering drugs (oral)	11.2 (31.5)	20.9 (40.7)
Decongestants	11 (31.3)	20 (40)
Observations	358,706	96,849

Notes: Column 1 shows descriptive statistics for the entire sample. Column 2 shows the sub-sample of compliers, whose donut hole exposure will be most affected by enrollment month (based on high initial spending: 71-97th percentiles, measured in the first 90 days of enrollment). One-year spending is measured from the first day of enrollment. One-year mortality is measured from December 1 of the first calendar year of enrollment, to parallel our analysis. The percent on a drug class is measured by the presence of any claim in a given class in the first 90 days of enrollment.

TABLE II
BALANCE OF KEY VARIABLES ACROSS ENROLLMENT MONTHS

	(1)	(2)
	Mean	Enrollment Month Effect (Std. Error)
Panel A: Key Characteristics		
White (%)	89.6	0.004 (0.022)
Female (%)	59.2	0.058 (0.036)
Initial 90-day spending ($)	357	0.18 (0.49)
Initial 90-day consumption (fills)	5.35	0.0036 (0.0042)
Predicted mortality (p.p.)	0.411	-0.00048 (0.00047)
Panel B: Initial Mortality (p.p., cumulative)		
1-30 days from enrollment	0.058	-0.000053 (0.0017)
1-60 days from enrollment	0.128	0.0013 (0.0026)
1-90 days from enrollment	0.201	0.0037 (0.0032)
Panel C: Initial Mortality, Complier Subset (p.p., cumulative)		
31-60 days from enrollment	0.082	-0.0037 (0.0041)
31-90 days from enrollment	0.181	-0.0011 (0.006)

$^{*}p < .1, ^{**}p < .05, ^{***}p < .01$

Notes: Panel A: Sample mean (Column 1) and enrollment month coefficients from a regression of key 'pre-treatment' variables on enrollment month (Column 2), for for the entire sample ($n = 358{,}706$). Predicted mortality is estimated by fitting a model with demographics and initial (3-month) drug data to predict subsequent (9-month) mortality, in a sample of 66-year-olds. Panel B: Sample mean (Column 1) and enrollment month coefficients from a regression of initial mortality, measured over different time windows after enrollment, on enrollment month (Column 2). This Panel includes the entire sample ($n = 363{,}306$), though rows 2-3 of Panel B exclude those that die in the first 30 days of enrollment ($n = 358{,}496$). Panel C: The same as Panel B, but restricted to compliers, whose donut hole exposure will be most affected by enrollment month (based on high initial spending; $n = 96{,}849$). Compliers here are identified using the first 30 days of spending (unlike our main specification, which uses 90 days), so we are unable to report the estimate for 1-30 days.

TABLE III
MORTALITY EFFECTS OF CHANGES IN DRUG BUDGETS

	(1)	(2)	(3)
	December Mortality (p.p.)	Enrollment Month Effect (p.p./mo)	Pre-Donut Budget Effect (p.p./$100)
Panel A: By Initial Spending			
Never-takers (1-70th pct.)	0.053	-0.000362	-0.000771
		(0.00199)	(0.00317)
Compliers (71-97th pct.)	0.118	-0.0112**	-0.0164**
		(0.00498)	(0.00786)
Defiers (98-100th pct.)	0.288	0.0469**	0.0549
		(0.0219)	(0.0402)
Panel B: Pre- vs. Post-Donut Hole Change			
2007-2010	0.083	-0.0055**	-0.0104**
		(0.00266)	(0.00427)
2011-2012	0.067	0.00405	0.00611
		(0.00321)	(0.00511)
Panel C: Pre- vs. Post-Donut Hole Change, Complier Subset			
2007-2010	0.120	-0.0137**	-0.021**
		(0.00626)	(0.00993)
2011-2012	0.113	-0.0063	-0.00993
		(0.00814)	(0.0129)

$^{*}p < .1, ^{**}p < .05, ^{***}p < .01$

Notes: Panel A: December mortality rate (Column 1), and regression estimates of the effect of enrollment month on mortality (Column 2), for three groups of beneficiaries (n = 358,706): Never-takers (lowest 70% of initial 90-day spending), who are unlikely to enter the donut hole irrespective of enrollment month; compliers (71-97th percentile) who are likely to enter the donut hole if enrolling earlier but not later; and defiers (98-100th percentile) who are likely to enter both the donut hole then the catastrophic coverage. Column (3) translates enrollment month into a beneficiary's 'pre-donut budget' (in $100/month), the amount she can spend before entering the donut hole, which is a mechanical function of enrollment month. Panel B: effect of enrollment month (or pre-donut budget) on mortality for the full sample (n = 358,706), estimated before vs. after a policy change that began to close the donut hole in 2011. Panel C: Same as Panel B, but restricted to the complier bin (n = 96,849; the four coefficients for never-takers and defiers are omitted for clarity). Robust standard errors are in parentheses.

TABLE IV
DRUG BUDGET EFFECTS ON DRUG CONSUMPTION, BY PREDICTED RISK

	(1)	(2)	(3)	(4)	(5)	(6)
	All		*Bottom 2/3 Risk*		*Top 1/3 Risk*	
	Mean	Est.	Mean	Est.	Mean	Est.
All Classes	126.40	4.93*** (0.289)	111.40	3.34*** (0.32)	155.50	7.13*** (0.552)
Cardiovascular	50.20	1.42*** (0.157)	40.50	0.598*** (0.172)	68.90	2.46*** (0.302)
Diabetes	10.20	0.618*** (0.0701)	9.80	0.515*** (0.0848)	10.90	0.813*** (0.123)
Respiratory	5.30	0.459*** (0.0453)	4.80	0.407*** (0.0544)	6.00	0.549*** (0.0805)

$^{*}p < .1, ^{**}p < .05, ^{***}p < .01$

Notes: N= 358,706. Effect of cost-sharing on consumption of key drug classes. Cardiovascular classes include statins, beta-blockers, ACE inhibitors, calcium channel blockers, angiotensin receptor blockers, and thiazide diuretics. Diabetes drugs include both insulin and oral hypoglycemic agents. Respiratory drugs are inhaled and oral treatments for chronic pulmonary disease. Column (1) shows the mean number of drug-days compliers filled in December: days supply, summed across all prescriptions filled, and grouped by drug class. Column (2) presents regression estimates (and robust standard errors) of drug-days on the pre-gap budget (in $100s). Columns (3) and (4) show mean drug days and regression estimates, restricted to the 2/3 we predict to be at the lowest risk of the adverse events each drug class prevents: heart attack and stroke for cardiovascular drugs, diabetic complications for diabetes, and respiratory failure for pulmonary drugs. Columns (5) and (6) do the same for the highest-risk 1/3. (For "all classes", we use predicted cardiovascular event risk, since both cardiovascular drugs and cardiovascular mortality are common.)

Figures

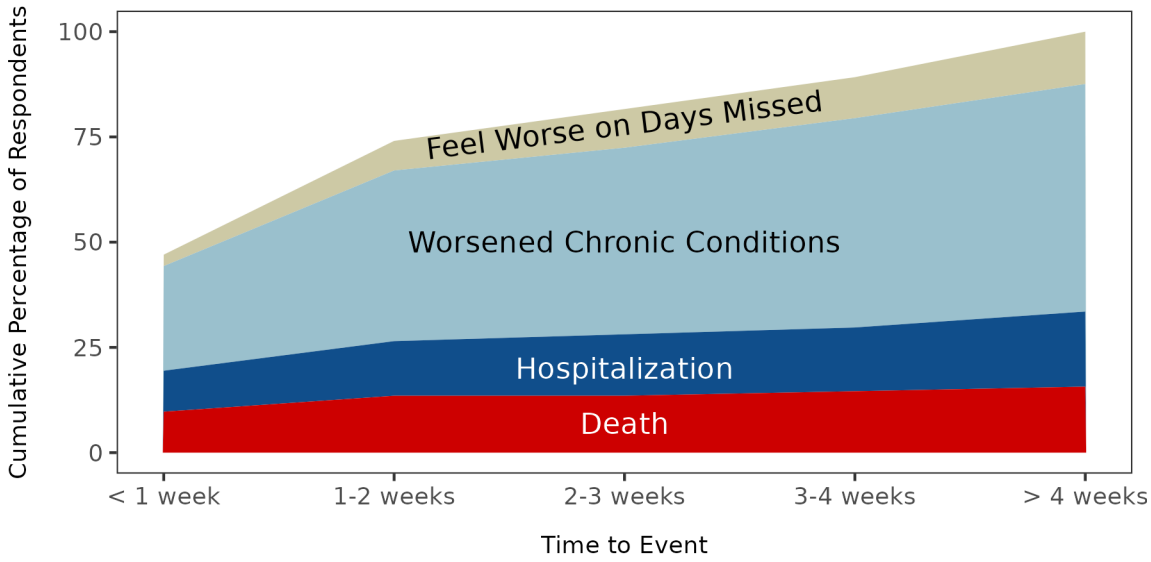

FIGURE I
Patient Beliefs on the Health Effects of Drug Interruptions

Notes: Results of a survey of 200 Medicare-age patients taking medications. The y-axis shows the cumulative percentage of patients who believed a given type of health problem could result from interrupting consumption of their medications. (Respondents were able to choose multiple problems, so each individual is assigned the most severe problem, ordered from feeling worse to death.) The x-axis shows the minimum number of weeks until such a health problem occurred. The exact wording of the questions was: "Patients often miss doses of their medications (research has found that up to 57% of doses are missed). Imagine a situation where you missed doses of your own medications. How long would it take before your risk of a serious health problem increased?" and "Think about the kinds of health problems that could arise from missing your medications. Which of the following could happen?" Possible responses were: "no change to your health", "you feel worse on days you miss the medications", "your chronic conditions get worse, in a way that eventually harms your health", "you need to be hospitalized", "death". 15 respondents selected "no change to your health," and thus do not contribute to the totals in the Figure.

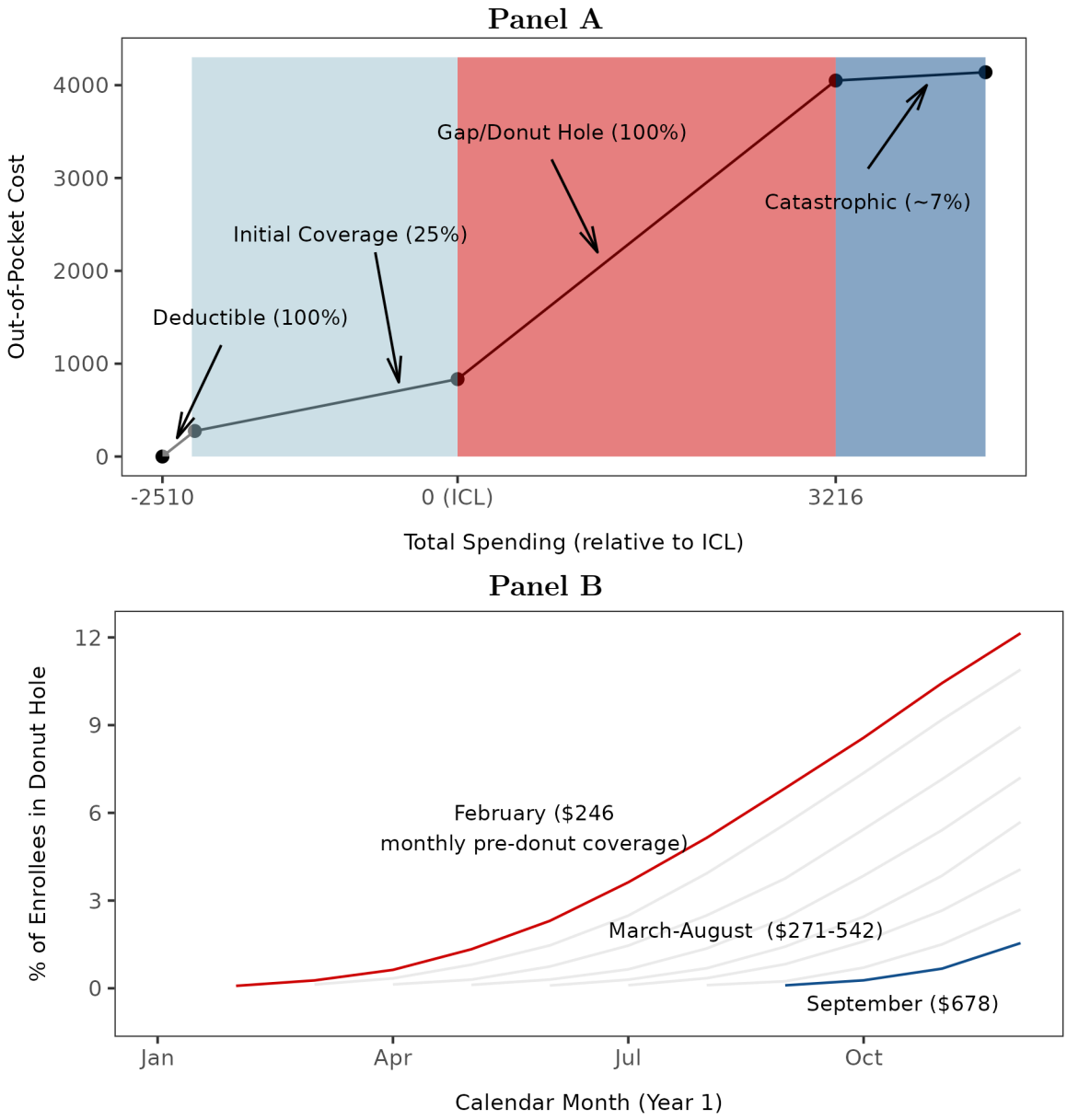

FIGURE II
Medicare Drug Benefit Design and Donut Hole Consequences

Notes: Panel A: Part D standard benefit design, adapted from Einav et al. (2015), using 2008 program details. The initial coverage limit (ICL) is the budget cap where beneficiaries transition from initial coverage to the donut hole. Panel B: Percentage of beneficiaries who enter the donut hole by the end of their first calendar year of enrollment, by enrollment month. February enrollees appear on top (red), March-August enrollees in the middle (gray), and September enrollees at the bottom (blue). The monthly pre-donut budget, the amount each beneficiary can spend before entering the donut hole, is shown in parentheses beside the enrollment month.

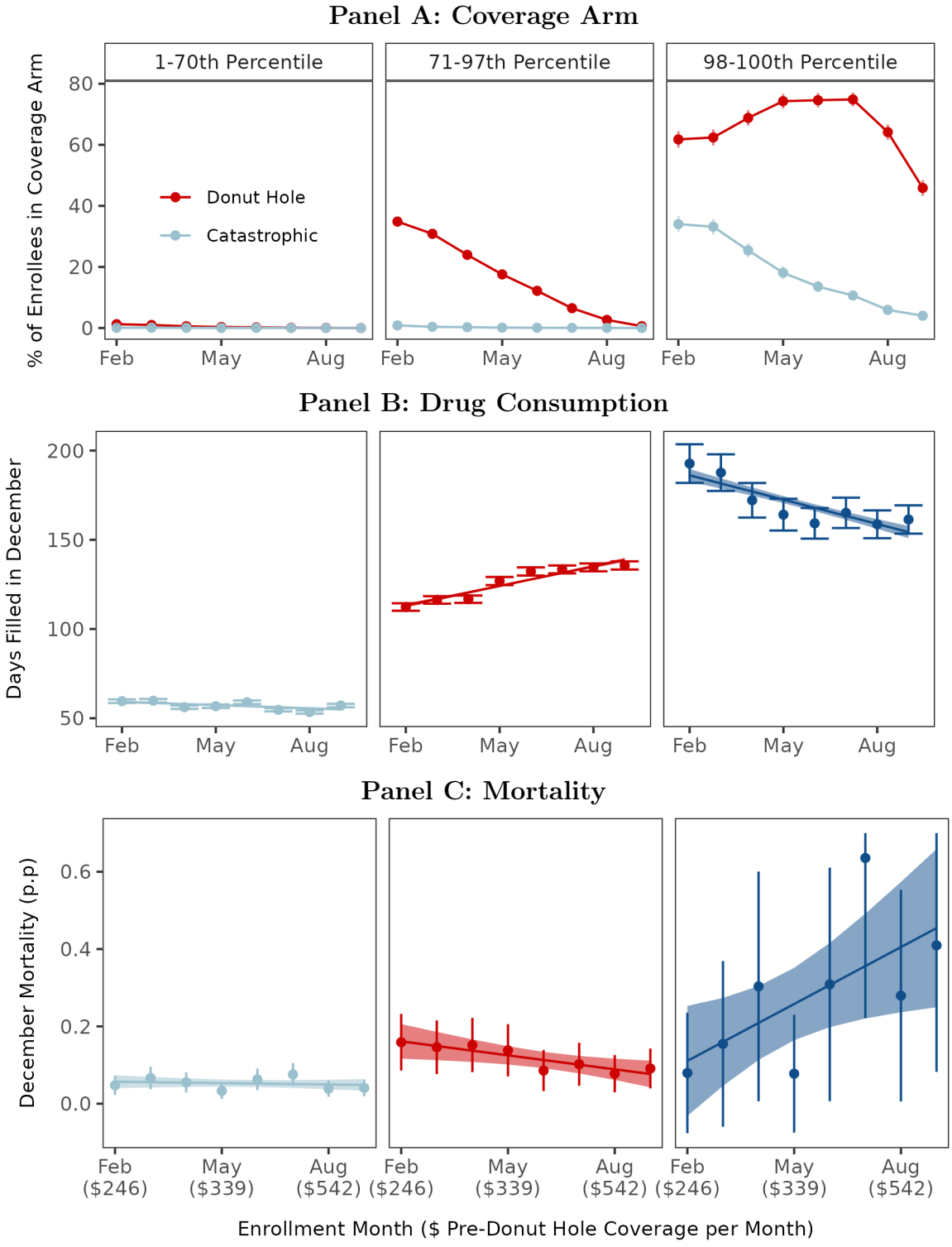

FIGURE III
Effect of Enrollment Month on End-of-Year Coverage Arm, Utilization, and Mortality

Notes: Panel A: Proportion of beneficiaries in a given coverage arm at year-end (y-axis), by enrollment month (x-axis) and initial 90-day spending bin (sub-panels, left to right: Never-takers, Compliers, Defiers). Initial coverage is omitted for clarity. Panel B: Drug-days filled in December (y-axis) by enrollment month and spending bin. Panel C: December mortality (y-axis) by enrollment month and spending bin; some upper confidence intervals in the right-most panel are truncated for visibility. The x-axis also shows monthly pre-donut budget for each month in parentheses.

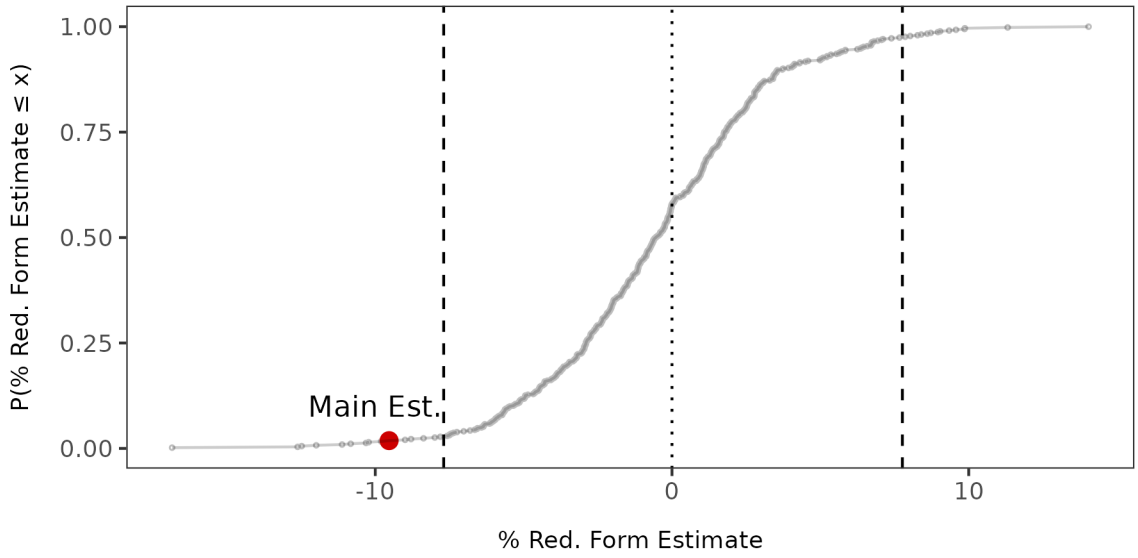

FIGURE IV
Distribution of 'Placebo' Estimates of Enrollment or Birth Month on Mortality

Notes: Regression estimates of the effect of enrollment month or birth month on mortality, for compliers (initial spending in the 71-97th percentiles, measured in the first 90 days of enrollment), in a variety of settings: Non-dual enrollees from age 66-85, dual enrollees from age 66-85, and disabled enrollees from age 50-64. Estimates are divided by mean mortality in each sample to get a percentage change. This Figure pools all falsification tests together; Appendix Figure B.1 provides further detail on the separate types of tests that contribute. Vertical lines show the 2.5 and 97.5 percentiles. The main (non-placebo) estimate is shown as a red dot.

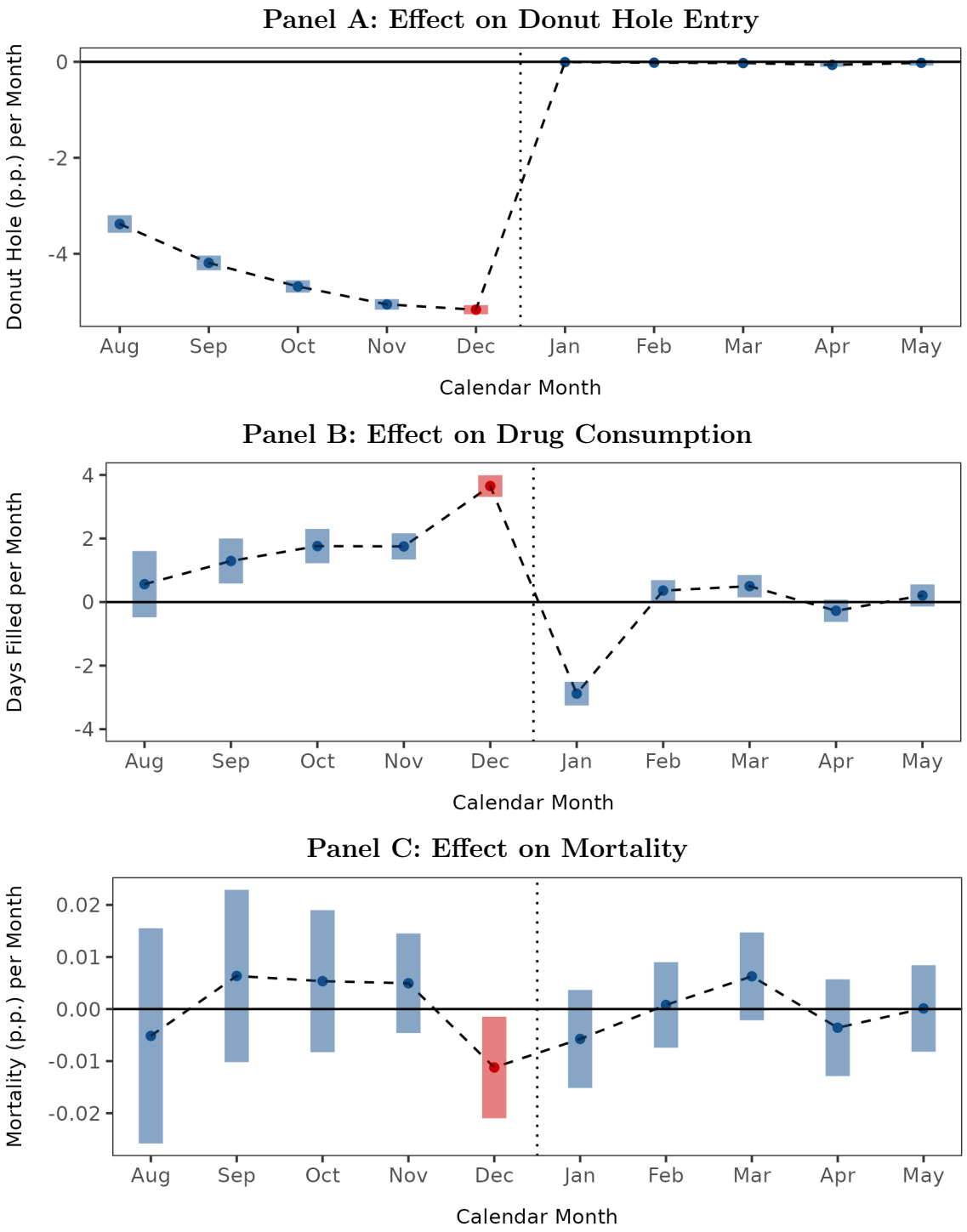

FIGURE V
Effects of Enrollment Month on Donut Hole Entry, Drug Consumption, and Mortality Over Time

Notes: Effects (and 95% confidence intervals) of enrollment month on key variables over time, for compliers (71-97th percentiles of initial spending). Each point shows the result of a regression, one for each month from August of year 1 to May of year 2. December is shown in red. For months prior to December, we exclude those who have not been enrolled for three months, in order to identify a similar group of compliers over time (e.g. September regression includes February-June enrollees). Panel A: Effects on donut hole entry. Panel B: Effects on the total number of drug-days supplied, summed over all drugs. Panel C: Effects on mortality.

Appendix A Prices, Part D Design, and Enrollment

FIGURE A.1
Channels for Price Effects

Notes: This graph shows how enrollment month can affect December mortality through multiple price and forward-looking behavior channels. Enrolling earlier leads to a higher price both in November and December. The higher November price affects November consumption which in turn affects the December price due to the non-linear contract. The higher December price affects consumption in December but may also affect November consumption if patients respond to the future price. Additionally, prices in future months (usually higher, but lower in January of year 2 after spending levels reset) may also affect the current month's consumption, if patients are forward-looking. These issues compound as we consider additional months (e.g. October). We are therefore unable to attribute health changes in December to any single price or quantity.

TABLE A.1
PART D COVERAGE ARM SPENDING THRESHOLDS AND COINSURANCE RATE BY YEAR

	(1)	(2)	(3)	(4)
	Spending Threshold		Gap Coinsurance Rate	
Year	Coverage Gap	Catastrophic	Generic	Branded
2007	$2,400	$5,451	100%	100%
2008	$2,510	$5,726	100%	100%
2009	$2,600	$6,154	100%	100%
2010	$2,830	$6,440	100%	100%
2011	$2,840	$6,648	93%	50%
2012	$2,930	$6,658	86%	50%

Source: q1medicare.com

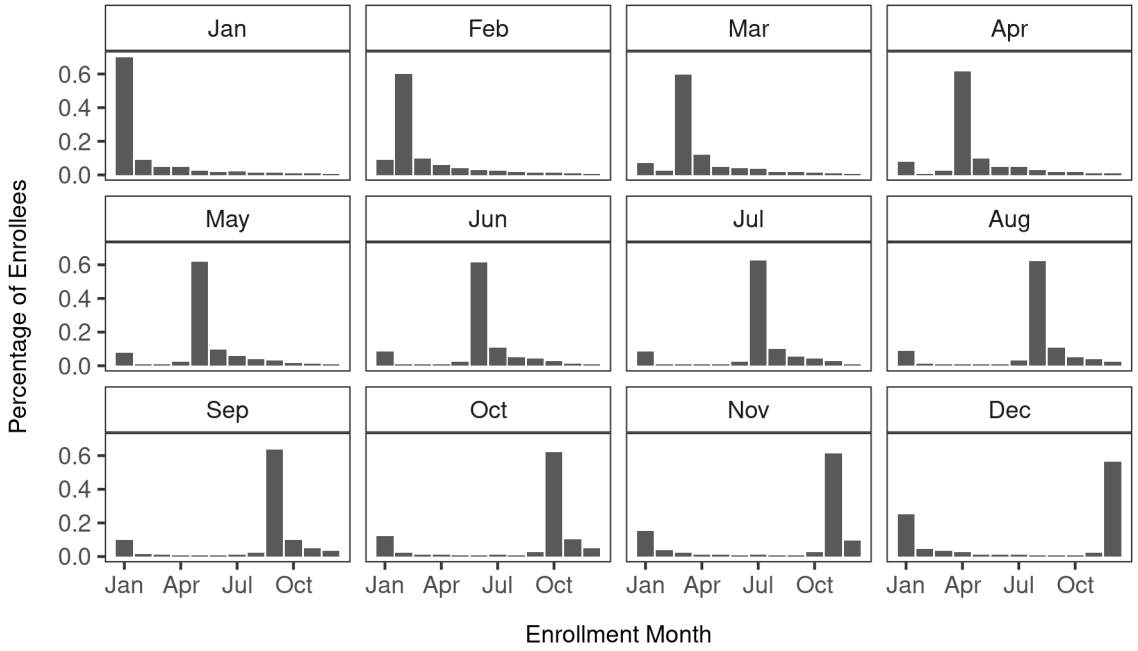

FIGURE A.2
Enrollment Timing by Birth Month

Notes: Each panel corresponds to a different birth month, and plots the percentage of beneficiaries (from that birth month) enrolling in each of the 12 calendar months during the first year of Part D eligibility.

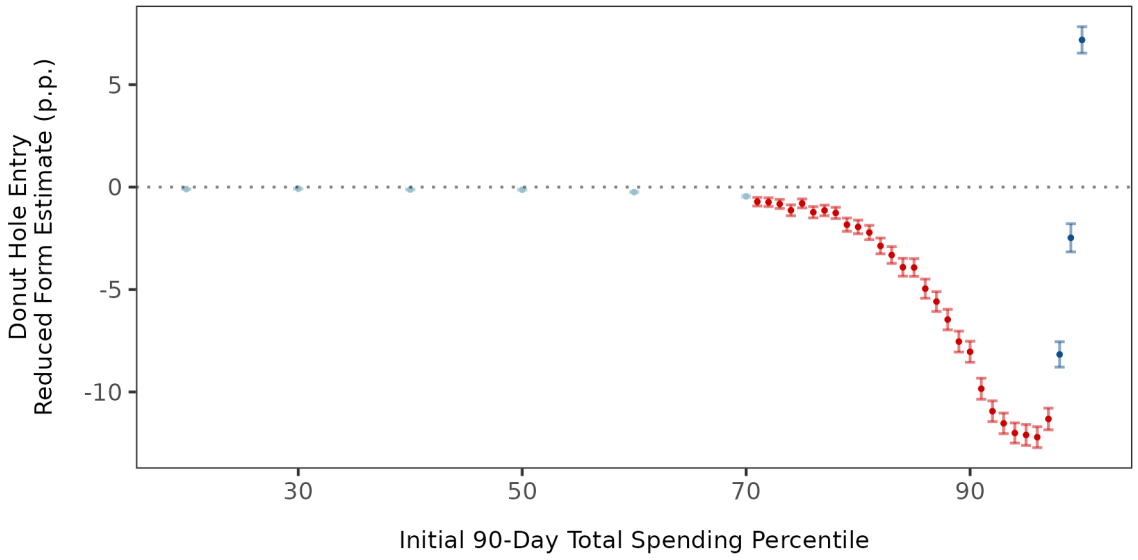

FIGURE A.3
Effect of Enrollment Month on Donut Hole Entry, by Fine Bin of Initial Spending

Notes: Each point represents the linear coefficient from a regression performed within a percentile of initial 90-day spending. The dependent variable is an indicator for ending year 1 in the donut hole, and the independent variable is enrollment month (and the usual controls, which are not shown). Initial spending levels below the 71st percentile are grouped into deciles because there is not enough variation in these levels of spending to define unique percentiles. The first and second deciles are grouped together for the same reason. Coefficients on the horizontal line (at zero) indicate no effect of enrollment month on donut hole incidence. Negative coefficients correspond to bins of initial spending where later enrollment months are more likely to end up in the donut hole, and positive coefficients correspond to bins where later enrollees are more likely to end up in the donut hole. We use these fine-binned estimates to create the larger bins displayed in Figure III and subsequent tables. The colors in this figure match those in Figure III.

TABLE A.2
BIRTH MONTH ESTIMATES

	(1)	(2)
	Dec. Mortality Rate (p.p.)	Birth Month Estimate (p.p./mo)
Panel A: By Initial Spending		
1-70th percentile	0.046	0.000258
		(0.00224)
71-97th Percentile	0.127	-0.0137**
		(0.00594)
98-100th percentile	0.285	0.055**
		(0.0224)

Panel B: Mortality by birth month (71-97th %)

Feb.	Mar.	Apr.	May.	Jun	Jul	Aug	Sep
0.185	0.176	0.147	0.144	0.088	0.11	0.069	0.108
(0.046)	(0.044)	(0.041)	(0.04)	(0.031)	(0.033)	(0.026)	(0.033)

$^{*}p < .1, ^{**}p < .05, ^{***}p < .01$

Notes: N= 358,706. In Panel A, present estimates of equation (2) where enrollment month is replace with birth month. In Panel B, we report the mortality rates by birth month for those in the 71-97th percentiles of initial spending. For both panels, we restrict to the subset of beneficiaries that enroll in their birth month (so enrollment and birth month are the same).

Appendix B Falsification Tests

We group our falsification checks into three categories, based on the population they draw on. First, we extend the intuition of Figure V and follow our main sample for four years after the first enrollment year. Specifically, we estimate equation 2 replacing December year 1 mortality with monthly mortality from January year 2 to December to December of year 4 (36 estimates). Regressions exclude those who died in a previous month. We also exclude those who are right-censored (e.g. those enrolling in 2012 are excluded from year 3 regression). Note that this particular exercise suffers from the fact that cost-sharing differences at the end of year 1 may lead to health effects in subsequent months. However, we expect this would bias estimates towards our main estimate (Table III). These estimates are shown in Appendix Figure B.1, Panel A.

Second, we focus exclusively on dual enrollees, who face little-to-no cost sharing for prescription drugs so there are no differences in cost-sharing by birth month. As a result, they provide a very useful comparison sample. Here, we pool all ages (66-85) together, but split the sample into subsets based on observable characteristics. Concretely, we construct 5 samples based on spending ventiles, 2 samples based on gender (male/female), 2 based on race (white/non-white), and 37 based on states. (Of note, we restrict to states with at least 10,000 observations in the dual sample.) For each of the 46 sub-samples we regress December mortality on birth month. These estimates are shown in Appendix Figure B.1, Panel B.

Finally, we turn to a larger sample of dual-eligibles ages 66-85, non-duals ages 66-85, and disabled enrollees ages 50-64. In both the disabled and older non-dual populations, almost all individuals are enrolled for the entire year and non-January enrollment is not driven by birth month, meaning there are are also no birth-month driven cost-sharing variation. To mirror our design, which focuses on a subset of high-spending individuals, we use claims from January-March to place enrollees into the same three initial spending spending bins: 1-70th, 71-97th, and 98-100th percentiles. We then estimate a total of 459 sample/month specific regressions (e.g. 66 year-old non-duals in December) of monthly mortality on birth month interacted with spending bin. Because these populations lack an observable enrollment month, we use birth month as a proxy. These estimates are shown in Appendix Figure B.1, Panel C.

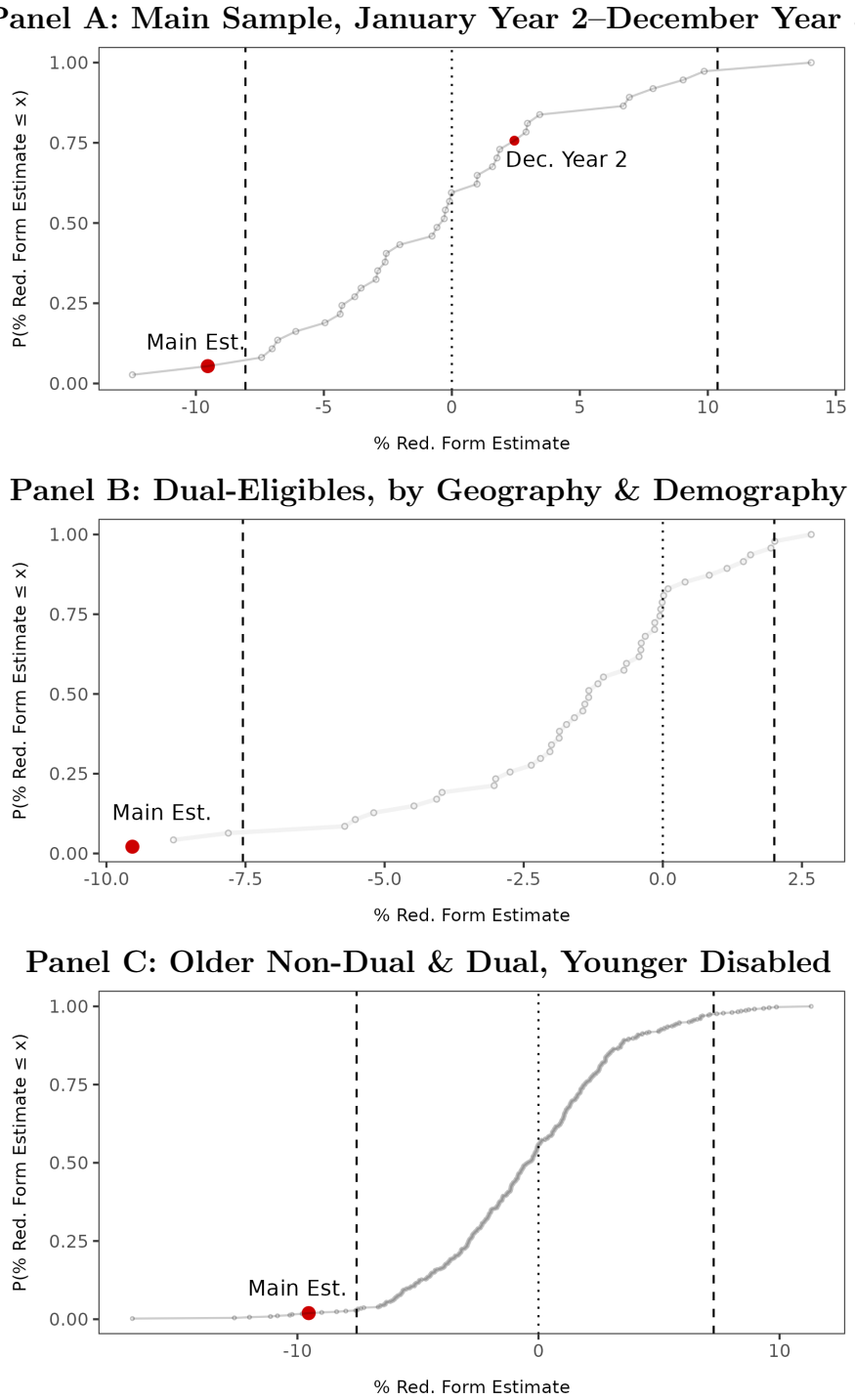

FIGURE B.1
Detail: Distribution of 'Placebo' Estimates of Enrollment or Birth Month on Mortality

Notes: Regression estimates of the effect of enrollment month or birth month on mortality, for compliers (based on spending in the 71-97th percentiles, January-March), in a variety of settings. Panel A: 36 enrollment month effects on monthly mortality in our main sample, restricted to compliers (initial spending in the 71-9th percentiles) from January of year 2, to December of year 4. Panel B: 46 birth month effects on monthly mortality in dual eligibles, ages 66-85, split into subsets based on: male/female (2), white/non-white (2), spending quintile (5), and state (restricted to states with at least 10,000 beneficiaries: 37). Panel C: 459 birth month effects from non dual enrollees from age 66-85, dual enrollees from age 66-85, and disables enrollees from age 50-64. Vertical lines show the 2.5 and 97.5 percentiles. The main (non-placebo) estimate is shown as a red dot.

Appendix C Risk Prediction

We define 3 sets of drug–event pairs G: cardiovascular, diabetes, and pulmonary. For each of these classes, we compile a list of observable adverse outcomes, Y_G: heart attack and stroke for cardiovascular medicines (ICD9 codes 410-411 and 433-435), diabetic complications (using the Diabetes Complications Severity Index (Young et al., 2008), e.g., foot amputation) for diabetes medicines, and pulmonary collapse requiring mechanical assistance (ICD9 codes 5188, 7991, 9604, and 9607) for pulmonary medicines. We group medications into these classes using ATC3 codes.

We construct a sample of dual-eligible enrollees, where each observation is a beneficiary-year. We define the outcome as an indicator for having a class of acute event (e.g., heart attack or stroke) over April to December. We form a set of predictors including race, sex, state and drug filling behavior (number of claims, total spending, ATC4 indicators) in the first 90 days of the year (roughly January-March). In total this 1,770 features. We then define separate training samples for each drug–event pair, restricting to those *not treated* with the particular medications in question (e.g., when predicting risk of heart attack or stroke, we exclude patients on statins, antihypertensives, etc.). We do this to form a prediction on the risk of complications if untreated (i.e., in potential outcomes notation, we wish to estimate $Pr(Y_{G0} = 1|X)$, not $Pr(Y_{G1} = 1|X)$. Naturally this choice of prediction target also induces selection bias: we form predictions on Y_{G0} in patients selected into treatment status $T = 0$, but then wish to generate predictions on in patients with arbitrary treatment status. In our setting, our predictions are likely to underestimate risk on average, because doctors select patients into treatment $T = 1$ on the basis of higher risk. This is similar in spirit to Mullainathan and Obermeyer (2022), who predict the yield of testing for heart attack in the tested, and apply the model to generate predictions in the untested. We verify that, as a check of face validity, risk increases in predicted risk for both groups (Figure C.1).

Our machine learning model consists of an ensemble of two predictors, LASSO (ℓ1-regularized regression) and gradient boosted trees (a combination of multiple tree-based models, each fit to the residual of the last). 70% of the sample is used to train the LASSO/gradient boosted models and 10% is used to ensemble them via no-intercept OLS of the outcome on the individual model predictions. We then validate the model using the final 20% of the sample. The model follows Mullainathan and Obermeyer (2022) closely. We then apply this model to generate predictions in

the main sample of 65 year-old beneficiaries (using the same predictors, measured in the first 90 days of enrollment). We make separate predictions for each (but use the same predictors for each).

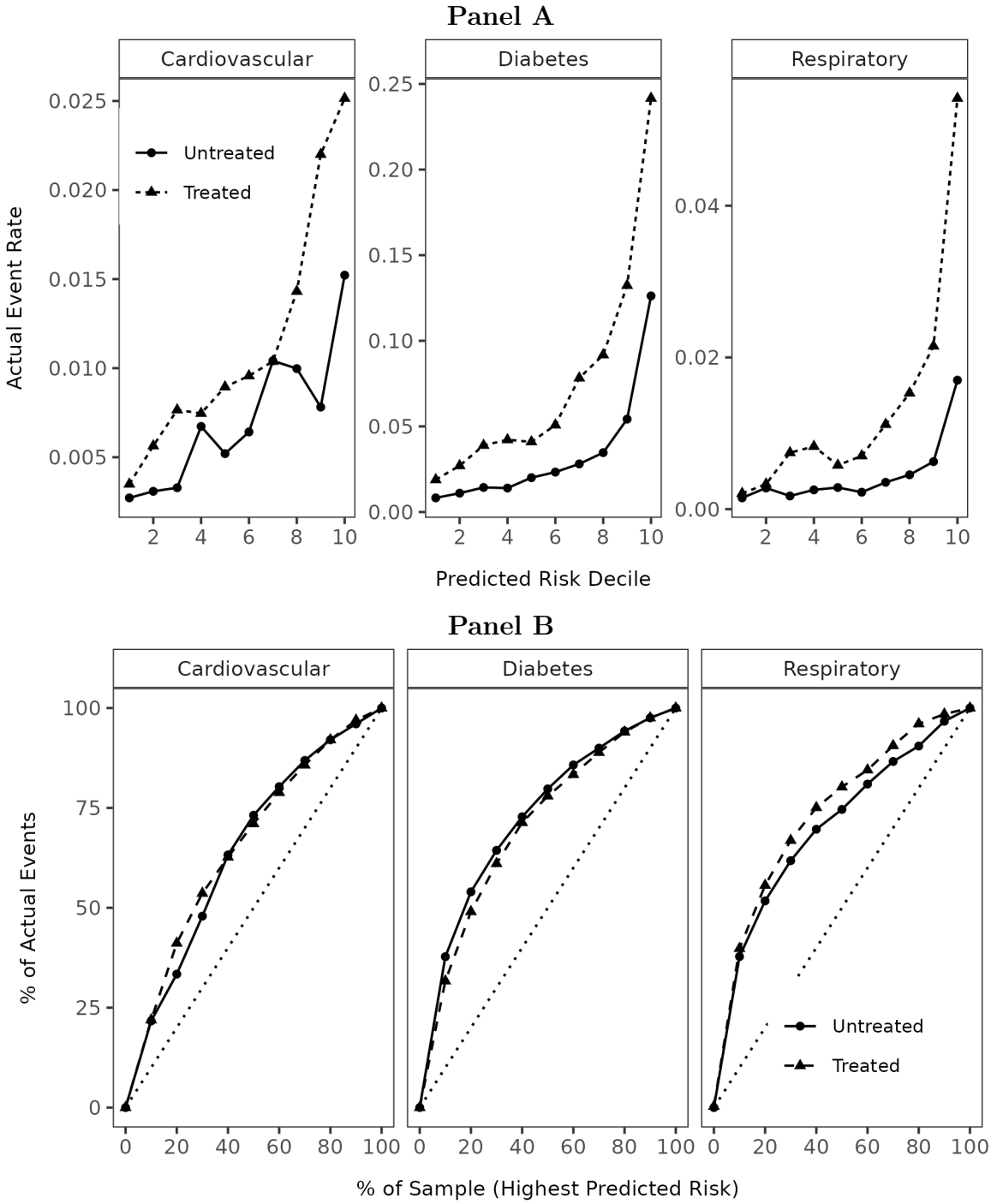

FIGURE C.1
Acute Event Risk Prediction Calibration and Capture

Notes: Panel A plots the actual event rate (for acute health events in each category, from days 90-360 of enrollment) by decile of predicted risk (using data from the first 90-days) and treatment status. An individual in the sample is considered treated if she fills a claim in the category in the first 90 days of enrollment. Panel B plots the cumulative percent of actual events captured by each successive decile (e.g. 0.1 is top 10 percent of sample in terms of risk) and treatment status. This figure uses only the stand alone PDP (non-MA) subsample for whom we observe Parts A claims.

Appendix D Effects of Cost-Sharing on Drug Consumption

Here, we provide evidence that the individuals that die in December exhibit similar utilization cut-backs as the entire sub-sample of compliers. Estimates in table IV are based on both individuals that survive and die in December. While we have no reason to expect the patterns to not hold in the mortality sub-sample, we are able to address this potential concern directly. To do so, we compare the utilization of early enrollees (February-May) to that of late enrollees (June-September) in the 1-30 days before death in December, to utilization in the 31-90 days before death. We hypothesize that early enrollees and later enrollees have (more) similar utilization 2-3 months before death, but that early enrollees have substantially lower utilization in the 1-30 days before death. Days filled by enrollment/pre-death timing along with a difference-in-difference estimate are presented in Table D.1 below. In the 31-90 days before death, early and late enrollees fill an almost identical amount of days per-month (148.0 vs 148.4). In the 30 days before death however, early enrollees fill only 101.9 days compared to the 130.0 days filled by late-year enrollees. This gives us a difference in differences estimates of 27.6 days filled. Scaling this by the average difference in the pre-donut budget between early and late year enrollees ($234), gives us an estimate of 11.8 days per $100/month of pre-donut budget. In conclusion, we observed the same, if not larger, utilization cut-backs in the death sub-sample as in the broader sample of compliers.

TABLE D.1
UTILIZATION CUTBACKS IN DECEMBER DEATHS

	(1)	(2)	(3)
	Days Filled		
	Late Enrollees (June-September)	Early Enrollees (February-May)	Diff-in-Diff Estimate
31-90 days before death (÷2)	148.4	148	
	(12.3)	(10.7)	
1-30 days before death	130	101.9	
	(16.2)	(13.5)	
Difference	46.06	18.44	27.61
	(17.23)	(20.41)	(26.71)

Notes: Column (1) presents mean days filled for those in the 71-97th percentiles of initial spending, that enrolled in June-September, and died in December of year 1, for the 31-90, and 1-30 days before death. Column (2) presents the same but for those that enrolled between February-May. Column (3) presents a difference in differences estimate which captures how much more early enrollees cut-back prior to death in December than late enrollees.

TABLE D.2
UTILIZATION RESPONSE BY RISK AND INCOME

	(1)	(2)	(3)	(4)	(5)	(6)
	Pre-Donut Budget Effect (Days/100$)					
	All		Bottom 2/3 Income		Top 1/3 Income	
	Mean	Est. (S.E.)	Mean	Est. (S.E.)	Mean	Est. (S.E.)
All Classes	126.40	4.93*** (0.289)	129.20	4.8*** (0.36)	121.30	5.2*** (0.493)
Cardiovascular	50.20	1.42*** (0.157)	51.00	1.4*** (0.193)	48.80	1.43*** (0.275)
Diabetes	10.20	0.618*** (0.0701)	11.00	0.623*** (0.0875)	8.60	0.567*** (0.118)
Respiratory	5.30	0.459*** (0.0453)	5.40	0.482*** (0.0564)	5.00	0.433*** (0.0776)

$^{*}p < .1, ^{**}p < .05, ^{***}p < .01$

Notes: N = 352,454. In column (1) we present the mean number of December days filled for all drugs and by broad class, for those in the 71-97th percentiles of initial spending. In column (2) present regression estimates (and robust standard errors) of days filled in December on the pre-donut budget (in $100s of dollars). Cardiovascular classes include statins, beta-blockers, ACE inhibitors, calcium channel blockers, angiotensin receptor blockers, and thiazide diuretics. Diabetes drugs include both insulin along with oral hypoglycemic agents. Respiratory drugs are drugs for obstructive airway diseases. In columns (3) through (6) we present means and regression estimates separately by quantile of five-digit zip code median income (from the American Community Survey). We exclude individuals for whom the American Community Survey has missing income for their zip code.

TABLE D.3
DEMAND RESPONSE: COMPARISON TO LITERATURE

	(1)	(2)	(3)
		Chandra et al.	
	Our Estimate (95% CI) (2SLS, Claims per $)	PPO	HMO
Overall			
Number of claims	-0.0377 (-0.0408, -0.0346)	-0.0143	-0.0387
	Our Estimate (95% CI) (2SLS, % with claims per $)	Choudhry et al.	Einav et al.
By Class			
ACE or ARB	-0.0051 (-0.0062, -0.004)	-0.0043	
Beta Blockers	-0.0043 (-0.0062, -0.0023)	-0.0032	-0.0034
Statins	-0.003 (-0.0036, -0.0024)	-0.0022	-0.0028

Notes: Here, we attempt to present demand response estimates on the same scale as three key previous studies. To generate comparable estimates, we set up a two-stage least squares—with all the caveats that enrollment month can affect utilization via many mechanisms, of which this is just one—that uses the December price in dollars (either overall or class specific) per fill as the endogenous variable. In Chandra et al. (2010), the authors present demand response estimates from two policy changes for different types of plans (HMO and PPO), hence the two columns here. Einav et al. (2018) report elasticities instead of a derivative; we multiply the elasticities reported in their study by $\frac{P}{Q}$ (estimated in our sample) to obtain a comparable derivative to the one we estimate. For Choudhry et al. (2011), we simply divide the quantity change in utilization for a class by the average copayment amount (prior to intervention, which erased copayments).