NBER WORKING PAPER SERIES

THE HEALTH COSTS OF COST-SHARING

Amitabh Chandra
Evan Flack
Ziad Obermeyer

Working Paper 28439
http://www.nber.org/papers/w28439

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2021, Revised February 2024

We acknowledge funding from the NSF Graduate Research Fellowship (Flack) and National Institute on Aging grant P01AG005842 (Obermeyer). We thank, without implicating, Ned Augenblick, Stefano DellaVigna, Amy Finkelstein, Ben Handel, Peter Hull, Jon Kolstad, Jon Gruber, Robin McKnight, Sendhil Mullainathan, Mohan Ramanujan, Julian Reif, Jon Skinner, Josh Schwartzstein, Amar Venugopal, Anna Zink, and numerous seminar participants for their thoughtful comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed additional relationships of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w28439

The Health Costs of Cost-Sharing
Amitabh Chandra, Evan Flack, and Ziad Obermeyer
NBER Working Paper No. 28439
February 2021, Revised February 2024
JEL No. I12,I13

ABSTRACT

What happens when patients suddenly stop taking their medications? We study the health consequences of drug interruptions caused by large, abrupt, and arbitrary changes in price. Medicare's prescription drug benefit as-if-randomly assigns 65-year-olds a drug budget as a function of their birth month, beyond which out-of-pocket costs suddenly increase. Those facing smaller budgets consume fewer drugs and die more: 0.0164 percentage points per month (13.9%) for each $100 per month budget decrease (24.4%). This estimate is robust to a range of falsification checks, and lies in the 97.4th percentile of 541 placebo estimates, formed in similar populations that lack the same idiosyncratic budget policy. Several facts help make sense of this large effect. First, patients stop taking drugs that are both 'high-value,' and suspected to cause life-threatening withdrawal syndromes when stopped. Second, using machine learning, we identify patients at the highest risk of drug-preventable adverse events. Contrary to the predictions of standard economic models, high-risk patients (e.g., those most likely to have a heart attack) cut back more than low-risk patients on exactly those drugs that would benefit them the most (e.g., statins). Finally, patients appear unaware of these risks. In a survey of 65-year-olds, only one-third believe that missing their drugs for up to a month could have any serious consequences. We conclude that, far from curbing waste, cost-sharing policies cause patients to miss opportunities to buy health at very low cost ($11,321 per life-year).

Amitabh Chandra
John F. Kennedy School of Government
Harvard University
79 JFK Street
Cambridge, MA 02138
and NBER
amitabh_chandra@harvard.edu

Ziad Obermeyer
School of Public Health
University of California, Berkeley
2121 Berkeley Way
Berkeley, CA 94704
and NBER
zobermeyer@berkeley.edu

Evan Flack
Department of Economics
Stanford University
Stanford, CA 94305
flack@stanford.edu

A data appendix is available at http://www.nber.org/data-appendix/w28439

I Introduction

Randomized trials quantify the benefits of starting a drug, but far less is known about the costs of stopping a drug. A researcher interested in this topic would struggle to obtain ethical approval: experimentally withdrawing prescribed medications is infeasible in most settings. But the non-linear contract structure of many health insurance plans creates sharp variation in out-of-pocket drug prices over time, causing millions of patients to stop prescriptions abruptly every year (Einav et al., 2015; Einav and Finkelstein, 2018; Einav et al., 2018).

In general, these price-driven drug interruptions have not been a cause for concern, at least to 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 this harms their health.[1] Patients may be deciding, rationally, that financial savings outweigh health gains, particularly if they have private information about the benefits (or side effects) of treatment. Indeed, patients themselves seem unbothered. 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 (e.g., hospitalization, death). A majority (53%) predict *no* negative health consequences, even simply feeling worse, from missing their drugs for up to a week.

The medical literature, by contrast, gives more cause for concern. Beyond the obvious reason—stopping a drug means forgoing any benefits shown in clinical trials—interrupting drugs can cause withdrawal or 'rebound' effects. This phenomenon is difficult to study experimentally, but there are some rare, striking signals from the medical literature that even short interruptions can have life-and-death consequences. Consider a classic study that randomly alternated a drug for high blood pressure (the beta-blocker propranolol) with placebo, to identify an optimal dose (Miller et al., 1975). Of 20 patients participating in what should have been a fairly mundane dose-finding trial, 10 had adverse cardiac events, 6 of which were heart attack or death; *all* adverse events occurred in the multiple short placebo periods—4 of 44 total weeks—when the drug was abruptly withdrawn. In another study, researchers noted that patients taking statins sometimes failed to get them for idiosyncratic reasons (Heeschen et al., 2002). In the setting of hospitalization for heart attack, those

[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.

who failed to receive their statin in the hospital had three-fold higher risk of death or repeat heart attack than similar patients who got their statin, and a 69% higher risk than patients who *never* took statins.[2] Rebound effects have been suggested for a large number of drugs in wide use by older patients, and guidelines suggest tapering under medical supervision rather than simply stopping them (Steinman and Reeve, 2023; Bain et al., 2008). But the empirical evidence backing these guidelines is weak, relying mostly on animal models, or highly confounded real-world studies that simply compare patients who do vs. do not interrupt their medications. Making matters worse, physicians are unlikely to notice even large effects: detecting a 30% increase in mortality from a baseline of 1% would require perfect awareness of drug interruptions in a sample size of 40,000 (vs. a typical panel: 2000 patients (Raffoul et al., 2016)). So in practice, even if doctors advise against interruptions, their recommendations are lost on patients—as evidenced by our survey results.

We study the mortality effects of drug interruptions at scale, exploiting an abrupt and as-good-as random price shock affecting millions of Medicare Part D beneficiaries. We build on an identification strategy pioneered by Aron-Dine et al. (2015) and Kaplan and Zhang (2017): a non-linear contract structure causes drug prices to vary by birth month for enrollees in their first year of coverage. Historically, 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 out-of-pocket costs jump to 100%. Whether or not a beneficiary enters the donut hole is of course not random: it depends on prior consumption. Critically, however, the cap is not pro-rated, giving all enrollees get the same 'pre-donut budget' whether they enroll early or late in the year. So by the time a later enrollee spends her first dollar, an earlier enrollee has been spending down her budget for months, approaching the donut hole. Because Part D eligibility begins in the month someone turns 65, birth month generates exogenous budget variation in the first year of enrollment.[3]

An important nuance is that the donut hole affects beneficiaries differently, depending on their drug spending. For example, most people end the year with far less than $2500 of total spending, meaning even early enrollees do not risk entering the donut hole. We focus on a set of 'middle-

[2]These observations illustrate a key principle of rebound effects. The body develops equilibrating reactions that compensate the drug's effect (e.g., opioids desensitize opioid receptors, requiring escalating doses to get the same effect (Dumas and Pollack, 2008)). When the drug is suddenly stopped, these endogenous mechanisms are no longer balanced, producing an opposite effect that will be problematic if the drug was doing something helpful.

[3]Reassuringly, baseline characteristics are similar for early vs. late enrollees, and balanced on demographics, drug consumption, and mortality in the first 90 days of enrollment.

spenders,' whose initial drug spending puts them on track to enter the donut hole if they enroll early in the year, but not if they enroll late.[4] We focus our analysis on the month of December, when our power to detect an effect is highest: sample size increases over time as 45,000 more beneficiaries enroll every month, and drug consumption differences across enrollment months grow as more early enrollees enter the donut hole.

The complex interaction of budget caps, initial drug spending, and enrollment month produces sharp variation in mortality. Enrolling one month earlier increases December mortality by 0.0112 p.p., or 9.5%. We translate this into policy-relevant terms by noting that enrollment month implies a monthly drug budget before full cost-sharing sets in[5]: each \$100/month drug budget decrease (a 24.4% change vs. the average budget) increases mortality by 13.9%. We verify that changes in consumption mirror changes in assigned drug budgets and mortality, with earlier enrollees consuming significantly fewer drugs.[6] This large mortality effect in middle-spenders contrasts with the absence of effect in low-spenders, just as we expect: low-spenders remain far from the donut hole, even if they enroll early, and show no effet of enrollment month on consumption. At the same time, the highest-spending 2-3% show a large, significant, and opposite-sign effect of enrollment month vs. middle-spenders. Again, this is just as we expect: in this group, earlier enrollees spend *through* the donut hole and enter the "catastrophic coverage", where cost-sharing drops from 100% to near zero, resulting in more drug consumption; later enrollees remain stuck in the donut hole. To summarize, enrollment month has effects on both drug consumption and mortality that vary widely across sub-populations, in a way that exactly matches the interaction of Part D budget caps and initial spending. It is hard to imagine a confounder with such a complex and idiosyncratic structure.

To build confidence in our estimates, we conduct a broad set of falsification tests and checks, beginning with two obvious potential confounders. First, while earlier enrollees are by construction

[4]We verify that initial spending, like other 'pre-treatment' characteristics, is uncorrelated with enrollment month. This means specifically that forward-looking behavior has not yet induced differential selection into initial spending bins, allowing us to identify a similar group of middle-spenders across enrollment months. We consider high-spenders, who not only enter the donut hole but also approach or enter the catastrophic coverage, in detail in Section II.C.

[5]Monthly budget is a mechanical function of enrollment month: e.g., a \$2500 budget gives a February enrollee \$227/month and a September enrollee \$625/month.

[6]We cannot easily link mortality to drug consumption using two-stage least squares in our setting, as the inter-temporal dependence of drug consumption and mortality would introduce bias. Intuitively, the enrollment month effect on mortality is mediated via a feedback loop of prices and quantities over time. Using any one of these quantities as the endogenous variable would violate the exclusion restriction. In particular, Appendix B shows that the combination of (i) prior drug consumption effects on current mortality (ii) inter-temporal substitution of drugs across periods produces a large upward bias in the estimated effect of consumption on mortality in any one period.

slightly older, mortality differences are 10 times larger than those implied by age differences (based on Social Security data). Second, well-known health differences across birth seasons have a different temporal pattern, and are far too small to explain our results (Doblhammer and Vaupel, 2001). Next, we replicate our analysis in a wide range of closely-related settings, to ensure mortality effects are only present when enrollment month affects drug budgets. For example, we replicate our analysis before vs. after reforms to Medicare that began to close the donut hole in 2011, and find that the effect size mirrors the extent of cost-sharing. We also track the evolution of enrollment month effects on drug consumption and mortality in our cohort of middle-spenders, before and after December of the first year. We find a significant difference in mortality only in December, exactly when differences in consumption across enrollment months peak and our power to detect an effect is greatest. The effect fades, then disappears after prices reset for all enrollees in January. Finally, we formally generate a large number of "placebo estimates": enrollment month effects on mortality, in Medicare populations for whom enrollment month does not affect drug budgets. This includes, for example, 66 year-olds, who are no longer affected by the enrollment month quirk we exploit at age 65; 65-year old dual-eligibles who face no cost-sharing; and 64-year old 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, our main estimate lies in the 97.4th percentile of mortality effects, larger in absolute value than 527 of 541 placebo estimates.

We make sense of this large mortality effect in two ways. First, like Einav et al. (2018) and Brot-Goldberg et al. (2017), we find that patients interrupt drugs that are 'high-value.' Many are also suspected to produce dangerous rebound syndromes when interrupted, with potentially life-threatening consequences: statins, antihypertensives, glucose-lowering agents, inhalers and steroids for pulmonary disease. Second, among those taking a given drug, we document a *positive* correlation between treatment benefit and likelihood of interrupting the drug. For example, beneficiaries at the highest risk of heart attack and stroke cut back four times as much on cardiovascular drugs (e.g., statins, antihypertensives) vs. lower-risk patients (2.46 vs. 0.60 drug-days per $100 change in monthly pre-donut coverage, a 3.6% vs. 1.5% reduction).[7] Similar patterns exist for diabetes and respiratory drugs. These differences are unlikely to be explained by income alone, as we see similar

[7]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. Again, we caution against comparing the magnitudes of changes in drug-days and mortality, given the biases detailed in Appendix B.

cutbacks in both high- and low-income zip codes. One potential driver of this effect is that higher-risk patients are more likely to interrupt *all* of their medications when prices increase, causing large absolute reductions in those on more drugs at baseline.

Our findings may be surprising from the point of view of standard economic models of behavior that emphasize moral hazard or private information: patients should not interrupt drugs with large benefits. Yet they fit with a growing literature in economics linking insurance coverage to lower mortality. Most closely related, Abaluck et al. (2021) find that switching into a plan with donut hole coverage reduces mortality by 9.8%.[8] More broadly, Miller et al. (2021) show Medicaid expansion reduced mortality by 9.4%, and Goldin et al. (2020) find that Obamacare tax incentives reduced mortality by 6.3%. Together, these studies show that health insurance can have large health impacts, by affecting how patients use—and especially under-use—high-value care (Baicker et al., 2015).

Cost-sharing has been a cornerstone of health insurance design for decades, driven by worries about wasteful spending and moral hazard. Our results indicate that, far from reining in low-value care, cost-sharing is itself highly wasteful. 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 of $100-200,000: Neumann et al. (2014)).[9] An optimistic view of these results is that policy makers have a unique opportunity to purchase health at negligible cost, by improving the design of prescription drug insurance.

II Empirical Strategy

Studying the health effects of drug interruptions is difficult. Experimentally stopping prescribed medications is impractical and unethical, and observational comparisons are highly confounded: 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 who take a lower fraction of assigned doses have worse outcomes—whether they are in the treatment or the placebo group (Osterberg and Blaschke, 2005). This fact elegantly demonstrates how difficult

[8] This estimate is from an two-stage least squares with donut hole coverage as the endogenous variable; of course, many other aspects of plans also vary so these comparisons are necessarily approximate. Further discussion of how our estimates relate to this literature is in Appendix Table E.2.

[9] Given that drug consumption can also offset inpatient spending (Chandra et al., 2010), this number is likely to be an over-estimate of the cost. Unfortunately, because half of our sample is in Medicare Advantage and we do not observe in- or out-patient spending, we are unable to study this directly.

it is to study this phenomenon: the myriad unmeasured links between drug-taking behavior and health can complicate comparisons of patients who do vs. do not interrupt their medications, even in the setting of a randomized trial. The idiosyncratic structure of the Medicare drug benefit, which creates exogenous and abrupt changes in drug prices, allows us to study the effects of drug consumption interruptions on mortality.

II.A Medicare policy context

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 plan details from 2008 to describe the plan, the calendar year begins with a deductible phase in which the beneficiary pays the entire cost of all drugs until spending reaches $275. She then faces a 25% cost-sharing (coinsurance) rate that lasts until spending exceeds the budget cap of $2,510 (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 (based on list-prices, which are significantly higher than net prices). Finally, after reaching $5,726 of total spending, she enters the 'catastrophic coverage' (to continue the analogy, this phase represents the far side of the donut). Here, she pays only 5% of the drug cost, or a copay of between $2.25 and $5.60 for each drug (depending on whether it is generic or preferred or not, respectively).[10] The cutoff points for each coverage arm change slightly from year to year, as shown in Appendix Table A.1, but the basic structure remained the same until 2011, when the donut hole began to close as a result of policy changes: cost-sharing rates for generic and branded drugs in the donut hole fell from 100% to 50% and 93%, respectively. The donut hole was officially eliminated by the Affordable Care Act in 2020, but many aspects of its

[10]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 use coinsurance, but rather use 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 coinsurance rates specified by the standard benefit (Einav et al., 2015).

structure persist in both Medicare (CMS, 2024) and and private insurance.[11]

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 'pre-donut hole' budget, $2510 of total spending, 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 $2510 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.[12]

Individuals become eligible to enroll in 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 with respect to health, and (ii) different birth months select into enrollment months similarly,[13] then the effect of any resulting variation in drug budgets on mortality can be identified. Like Aron-Dine et al. (2015), we prefer estimates based on enrollment month as opposed to birth month, and use these 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.[14] In the empirical work, we test the underlying assumptions in detail: first, by checking that those from different birth or enrollment months are similar, on range of observable health and utilization measures (Table II). We also compare our estimates to those from the literature on birth-seasonal variation in health to see if these known effects can account for our results.

[11] For a helpful overview, see KFF (2023).

[12] Very high-spending earlier enrollees are also more likely to enter the catastrophic coverage after the donut hole, for the same reason.

[13] 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. We explore the selection of enrollment timing from birth months in Appendix Figure A.1. 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).

[14] Using birth month instead of enrollment months for our main analysis yield similar, if less precise, results, just as we would expect: see Appendix Table C.2.

II.B Model

We would ideally like to estimate the effect of drug consumption Q, a vector of consumption measures for individual drugs (specifically, number of drug-days consumed in some period), on mortality, denoted by indicator Y. In addition to Q, Y is determined by patient factors, which can be either observed (vector X) or unobserved (vector W). Unobserved factors affect both mortality and consumption, so OLS regression of Y on Q will be confounded, hence the need for an instrument Z that exogenously shifts drug consumption: enrollment month. For ease of exposition, we begin with a simplified scenario where Z takes two values, indicating whether patients are assigned to enroll early in their first year, receiving the usual Medicare monthly drug budget ($Z = 0$), or late in the year, receiving a higher monthly drug budget ($Z = 1$).[15]

We begin by considering the effect of Z on drug prices via the donut hole.[16] Let function $g(S, Z, t; \theta)$ represent the Medicare policy that calculates cumulative spending from beneficiary i's enrollment month $Z_i = z$ to period t, $S_{i,z:t} = \sum_{j=z}^{t} S_{ij}$, and compares it to policy limits θ, in particular the donut hole limit $\theta_D 9$.[17] The function determines whether the beneficiary is in the donut hole in t, measured by indicator $D_{it} = \mathbb{1}\{S_{i,z:t} \geq \theta_D\}$. The instrument mechanically affects the calculation of cumulative spending, making earlier enrollees more likely to enter the donut hole: in potential outcomes notation, $E[D_{it}^0] \geq E[D_{it}^1]$, where the superscript corresponds to the two potential enrollment months.[18] This sets the vector of individual drug prices: $P_{it} = P_0 + \delta D_{it}$, the baseline price P_0 plus the donut hole premium δ. Earlier enrollees thus face higher prices, $E[P_{i,t}^0] > E[P_{it}^1]$, which likely leads to lower drug consumption $E[Q_{it}^0] < E[Q_{it}^1]$. We hypothesize that these differences in consumption drive mortality differences such that $E[Y_{it}^0] > E[Y_{it}^1]$.

Given our interest in the effect of budget-induced drug consumption changes on mortality, a natural estimation approach would be two-stage least squares, instrumenting for drug consumption using enrollment month. Unfortunately, any choice of endogenous variable for the first stage will

[15] This simplified binary case sets up our potential outcomes framework; it will generalize to our actual empirical setting, where Z runs from 2 for February enrollees to 9 for September enrollees, with larger values always implying larger pre-donut hole budgets.

[16] The donut hole is the policy feature that affects the largest fraction of beneficiaries. We consider the catastrophic coverage policy, which only the highest-spending 2-3% of beneficiaries enter, in Section II.C below.

[17] To simplify notation, we write θ_D as a constant, but in our analysis we account for the changing values of all policy thresholds across years.

[18] While D_{it} is determined by spending, whether or not someone actually enters the donut hole depends on their consumption choices, so $E[D_{i,t}^z], z \in \{0, 1\}$ is taken over the distribution of individual behaviors, contingent on being assigned to enroll early or late.

violate the exclusion restriction, if the instrument affects mortality via a complex feedback loop connecting quantities to prices over multiple time periods, because any of these may affect mortality. A pernicious consequence of this problem is that estimates of the effect of consumption on same-period mortality will be biased upwards, as we show more formally in Appendix B. The intuition for this is two-fold. First, because prior and current drug consumption are correlated, and both affect mortality, a regression of mortality on same-period consumption will be biased: the effect of prior consumption will be mis-attributed to the current period.[19] Second, because of the well-documented phenomenon of inter-temporal substitution across time periods (Aron-Dine et al., 2015), past consumption is greater than current-period consumption, inflating the bias: small changes in current consumption will appear to have large effects on same-period mortality, reflecting in part the large changes in prior consumption.

As a result, we report for a transparent, reduced-form estimate of the effect of enrollment month on mortality in period t, using the following estimating equation:

$$Y_{it} = \gamma_0 + \gamma_1 Z_i + X_i \gamma_2 + \gamma_{year} + \gamma_{plan} + \epsilon_{it} \tag{1}$$

where γ_{year} is a set of calendar year fixed effects, γ_{plan} is a set of fixed effects for Part D plans, X_i is a vector of sex and race indicators, and the instrument Z_i is a scalar that takes integer values from 2 (February enrollment) to 9 (September enrollment).[20] Our primary interest is γ_1, the effect of enrolling one month later. The reduced form estimate of enrollment month is not generalizable to other (non-Medicare) settings, where enrollment month is not a meaningful quantity. However, it does have a natural policy-relevant interpretation: enrollment month mechanically implies a monthly dollar budget—how much a beneficiary can spend before reaching the full cost-sharing of the donut hole: $B_i = \frac{\theta_D}{(T - Z_i + 1)}$, where $T = 12$ (the total number of months). This has direct relevance to other health insurance settings that use similar budget structures.

While we do not implement formal two-stage least squares to link mortality to drug consumption, we do estimate the instrument's effect on drug consumption, both for individual drugs and total

[19] We cannot easily get around this by using cumulative consumption, because the measurement of consumption (and all other variables) in prior periods is correlated with the instrument—enrollment month: we will not observe consumption for later enrollees as we go back in time.

[20] We exclude January, October, November, and December enrollees for reasons described in Section II.D.

consumption, in a single period t, via

$$Q_{it} = \alpha_0 + \alpha_1 Z_i + X_i \alpha_2 + \alpha_{year} + \alpha_{plan} + \eta_{it} \tag{2}$$

This is useful to confirm that budget limits produce the expected changes in drug consumption, and to study where and when beneficiaries cut back. However, we emphasize that it is misleading to correlate the magnitude of changes in estimated consumption to changes in estimated mortality: any such effort suffers from the same bias as described above and in Appendix B, where small changes in consumption will appear to have outsized effects on same-period mortality.

Given the inter-temporal dynamics described above, the choice of t is important. We favor a time period near the end of the first calendar year—specifically December, like Aron-Dine et al. (2015)—for two reasons. First, more and more beneficiaries enroll and reach their steady-state drug consumption, meaning we have larger samples as the year progresses. Second, differences in drug consumption across enrollment months accelerate over the course of the year, as shown in Figure II, Panel B, before prices reset for all enrollees in January. The combination of these two factors means our power to detect an effect increases dramatically and non-linearly over the course of the first calendar year, peaking in December, as shown in Appendix Figure C.3. For example, our power to detect an effect of the same magnitude as December would be 72 percent lower in November, and 76 percent lower in October.

II.C Heterogeneity in the Effect of Enrollment Month on Prices

We have so far focused on the dominant effect of enrollment month Z_i: earlier enrollment increases the likelihood of entering the donut hole by year-end ($D_{i,12}$), resulting in higher prices and lower consumption. But the effect of Z_i is heterogeneous. For example, after the first calendar year of enrollment (where all beneficiaries look like our early enrollees, since cumulative spending is calculated starting in January), the majority of Part D beneficiaries have cumulative year-end spending well below the donut hole budget limit $E[S_{i,1:12}] << \theta_D \approx \$2500$.[21] For such low-spenders $i \in \mathcal{L}$, the effect of enrollment month on donut hole entry is likely negligible, because even the earliest enrollees will end the year far from the donut hole limit. By contrast, a handful of the

[21] This is based on data from after the first enrollment year, i.e., absent enrollment month effects.

highest-spending beneficiaries $\mathcal{H}$ are *less* likely to end the year in the donut hole: they do enter the donut hole, but then exit it by exceeding the second policy limit $E[S_{i,1:12}|i \in \mathcal{H}] \geq \theta_C \approx \$6000$, landing in the catastrophic coverage phase where cost-sharing returns to very low levels.

We would ideally like to focus attention on a group of middle-spenders $\mathcal{M}$ whose spending is 'just right': they spend enough to be affected by the donut hole budget limit by year-end, but not enough to approach the catastrophic coverage limit. In this group, we expect large effects of enrollment month on donut hole entry by year-end, and thus consumption and mortality, in a way that is relatively easy to interpret. An analogy can help see why: the middle spenders are like compliers, for whom donut hole entry is monotonically lower for later enrollees vs. earlier enrollees: $E[D^0_{i,12}|i \in \mathcal{M}] > E[D^1_{i,12}|i \in \mathcal{M}]$. The lowest-spenders are never-takers, for whom $E[D^0_{i,12}|i \in \mathcal{L}] = E[D^1_{i,12}|i \in \mathcal{L}] = 0$. The highest-spenders are defiers, $E[D^0_{i,12}|i \in \mathcal{H}] < E[D^1_{i,12}|i \in \mathcal{H}]$, who violate the monotonicity assumption: early enrollees are more likely to enter the donut hole (resulting in higher prices), but also more likely enter the catastrophic coverage (resulting in lower prices)—i.e., the effect of enrollment month Z_i on prices and consumption is ultimately reversed. Focusing on the middle-spenders, who experience *only* the (prospect of) budget decreases, would produce the most straightforward estimate of the effect of Z_i.

In order to assign beneficiaries to groups $\mathcal{G}_i \in \{\mathcal{L}, \mathcal{M}, \mathcal{H}\}$, we cannot simply use realized year-end spending—our identification strategy relies on the fact that cumulative year-end spending $S_{i,z:12}$ is endogenous to Z_i. Just after enrollment, though, the donut hole limit may be sufficiently far away that spending is temporarily unaffected by Z_i: enrollees are all just starting to learn about a highly complex program. While this is plausible, if even initial spending is affected by forward-looking behavior and thus correlated with enrollment month, using initial spending to assign groups would induce differential sample selection across enrollment months: higher-spending earlier enrollees, who have already started to cut back, would be inappropriately grouped with lower-spending later enrollees, who have not. Fortunately, this is an empirical question: we can directly inspect beneficiaries' spending trajectories in their first months of Part D coverage, to determine whether initial spending varies by enrollment month. To do so, we first test for differences in spending in the first month of enrollment $E[S_{i,z}|Z_i = 1] - E[S_{i,z}|Z_i = 0]$ by regressing $S_{i,z}$ on scalar Z_i. If we find no significant effect (at $p < 0.05$), it implies that forward-looking behavior has not materialized by the end of the first month. We proceed to test for differences in month $(z+1)$, both individually and cu-

mulative since enrollment, and so on until we find a significant coefficient on Z_i. Appendix Table A.2 shows that such a difference emerges only in month $z+3$. So we define 'initial spending' as spending in the first three months of enrollment $S_{i,z:z+2}$. We calculate percentile bins of initial spending for each enrollment month, and use these to define a vector of indicators $\sigma_i = [\sigma_{i1}, \sigma_{i2}, ...\sigma_{ij}, ...\sigma_{i100}]$, which are set to 1 if beneficiary i falls into the j^{th} percentile of initial spending within her enrollment month. Using within-enrollment month percentile is attractive because is requires only a weaker assumption, that the *ranking* of initial spending within-enrollment month is correlated with the likelihood of entering the donut hole at the end of the year.

This lets us estimate the empirical analog of $E[D^1_{it} - D^0_{it}|\mathcal{G}_i]$ using our instrument and spending bins: $E[D_{it}|Z_i = 1, \sigma_i] - E[D_{it}|Z_i = 0, \sigma_i]$.[22] Focusing on December ($t = 12$), we regress

$$D_{i,12} = \phi_0 + (Z_i \times \sigma_i)\phi_1 + X_i\phi_2 + \phi_{year} + \phi_{plan} + \omega_{i,12} \tag{3}$$

and inspect vector ϕ_1, coefficients for the vector of spending-percentile indicators σ_i interacted with scalar Z_i. This measures the effect of enrollment month on donut hole entry at year-end for each spending percentile bin j. If $\phi_{1j} = 0$, it implies $E[D_{it}|Z_i = 1, \sigma_i] = E[D_{it}|Z_i = 0, \sigma_i]$, and all beneficiaries in spending percentile j are assigned to $\mathcal{L}$. Likewise, if $\phi_{1j} < 0$ then $i \in \mathcal{M}$, and if $\phi_{1j} > 0$ then $i \in \mathcal{H}$. With spending groups defined, we run our primary estimating Equation 1 separately by group:

$$Y_{i,12} = \gamma_0 + (Z_i \times \mathcal{G}_i)\gamma_1 + X_i\gamma_2 + \gamma_{year} + \gamma_{plan} + \epsilon_{i,12} \tag{4}$$

The vector γ_1 will measure the effect of enrollment month on mortality, separately by initial spending group $\mathcal{G}$. Our hypothesis is that these mortality effects will mirror the effects on donut hole likelihood from Equation 3, with $\gamma_{1\mathcal{L}} = 0$, $\gamma_{1\mathcal{M}} < 0$, and $\gamma_{1\mathcal{H}} > 0$.

II.D Data

Our main sample for estimation 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 2007-12.

[22] We take this empirical approach, rather than trying to forecast whether or not they will exceed policy budget limits (θ_D, θ_C) based on initial spending, because we know cutbacks start well in advance of these limits.

First, we make sample restrictions common to the Part D literature—though we will later use some of these excluded populations for falsification tests. We subset to those who become eligible for Medicare because they turn 65, under the Old Age and Survivors Insurance, which means excluding those who enroll in Medicare before age 65, for disability or end-stage renal disease. 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, 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.[23]

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 exclude those who die before December 1. We carefully check mortality differences across enrollment months before December 1, which could indicate selection bias introduced by this exclusion, and find none: see Table II and Figure IV. We also drop those who enroll in October and later, because as Aron-Dine et al. (2015) note, these beneficiaries are still ramping up their drug consumption. 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. Finally, we follow Aron-Dine et al. (2015) and exclude January enrollees from our sample, for several reasons. 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.[24] With these restrictions, our final analytic sample consists of 358,706 individuals.

Our analyses of drug consumption use the Medicare Part D drug claims made by beneficiaries in our sample, including fill date, total cost, out-of-pocket cost, and 11-digit National Drug Code (NDC) identifiers. To classify drugs into clinically meaningful categories, we use the RxNorm and RxClass APIs to link NDCs to their corresponding Anatomical Therapeutic Chemical (ATC) codes,

[23]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. We include individuals in standalone PDPs and standard MA plans that are HMO, HMO POS, Local PPO, Private FFS, and Regional PPO.

[24]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.

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 the cardiovascular category.[25]

III Results

III.A Sample Description

Summary statistics are shown in of Table I. The overall sample is 90% white and 60% female. Roughly half is in fee-for-service Medicare (standalone PDP, with the rest in Medicare Advantage. As our sample is relatively young and not dual-eligible (excluding the poorest beneficiaries, for whom enrollment month does not affect cost-sharing), mortality is low, 0.9 percentage points (p.p./year), and drug spending around $1,500/year. The 10 most-used drug classes in our sample include statins, antihypertensives, diuretics, antidepressants, glucose-lowering drugs, and corticosteroids.

We check balance by regressing key 'pre-treatment' characteristics on enrollment month. Columns 1-2 of Table II present means and estimated coefficients on enrollment month for the entire sample, respectively. Panel A shows that estimates for race, sex, and number of drug prescriptions filled in the first 90 days are statistically and economically insignificant. Balance on spending is discussed above (Section II.C and shown in Appendix Table A.2. As a more synthetic test, we predict one-month mortality using all pre-treatment variables, and regress this on enrollment month.[26] This too is reassuring. Panel B shows a more direct balance check on mortality in the first 3 months of enrollment, by regressing mortality in the 30, 60, and 90 days after enrollment on enrollment month. Because cost-sharing begins to affect enrollment months differently after this period (Appendix Table A.2), and the latest enrollment months has already entered December by the fourth month after enrollment, this is the last time we can compare mortality across enrollment months for

[25] We also attempted to measure medical diagnoses, procedures, and health care utilization besides drugs, using Medicare Parts A and B claims, including diagnoses, procedures, and admit/discharge dates. However, we are underpowered to detect effects for two reasons. First, the subsample of individuals enrolled in standalone PDPs (non-MA), for whom we observe these data, is half our sample. Second, 69% of all deaths happened outside of the hospital, implying a sudden event that did not result in prior utilization.

[26] This predictive model is estimated in a sample of 66+ year-old dual enrollees, who we assume have the same relationship between mortality and covariates. The independent variables are demographics (race, sex) along with consumption metric 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.

the entire sample. No estimates are significant, providing further evidence that baseline health is similar between enrollment months. This provides some reassurance that conditioning our analytic sample on survival (i.e., to run our model of December mortality beneficiaries must be alive on December 1) does not introduce selection bias correlated with enrollment month.

Having established balance, we now turn to defining initial spending bins as described above in Section II.C. Appendix Figure A.3 shows the effect of enrollment month on the likelihood of entering the donut hole by year end, for each percentile of initial spending.[27] In the first 60 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 percentile bin has a large positive effect, as earlier enrollees exit the donut hole and enter the catastrophic coverage.

Our primary goal in setting spending group cutoffs is to estimate Equation 4 in a homogeneous group of middle-spenders, whose drug consumption monotonically increases in enrollment month. Estimates for low-spenders and high-spenders are not our primary focus, so we are willing to tolerate some heterogeneity in these groups, to ensure homogeneity in the middle-spender group. We thus assign the first 70 percentiles to the low-spender bin—even though, empirically, this is likely to include a small number of lower-middle-spenders from percentiles 61-70 who are slightly affected by the donut hole—and percentiles 98-100 to the high-spender bin. On average, the 98-99th percentiles likely experience the impact of enrollment month more like a middle-spender than a high-spender (i.e., the effect of enrollment month on donut hole entry is negative); but because our primary interest is in estimates from the middle-spender 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 middle-spender bin at the 61st percentile, or ending it at the 95th percentile) and found nearly identical results, shown in Appendix C.1.

Table I, Columns 2-4, present summary statistics by spending group. Middle-spenders, our

[27] As discussed in Section II.C and Appendix Table A.2, we use 90-day initial spending because it is balanced across enrollment months, while spending past 90 days is likely affected by forward-looking behavior. There are not unique percentiles for the 70% of initial spending so we estimate by decile for this group. We find precise null estimates for all 6 deciles.

main population of interest, are more likely to be fee-for-service, and have higher drug spending (by construction) and mortality than the overall sample mean. Table II, Columns 3-4, present balance checks within the middle-spending group. While we do find a statistically significant association between enrollment month and sex in the middle-spenders sub-sample, it is quite small in both percentage terms (0.29%) and orders of magnitude too small to explain the mortality differences we find. Drug consumption, measured by number of prescriptions filled in the first 90 days, is balanced across enrollment months. Our most important balance tests— predicted mortality based on observables, and actual initial mortality in the first months of enrollment—are likewise reassuring.

III.B Mortality Effects of Drug Interruptions

Figure III shows our identification strategy and main result graphically, using a set of indicators for enrollment month. In Figure III, Panel A, we see the proportion of beneficiaries in each coverage arm in December, by spending group. This aggregates the finer percentile bins in Appendix Figure A.3, and shows a similar pattern of clear separation of (mostly) low-spenders on the left, for whom enrollment month has no effect on donut hole entry; middle-spenders, 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) high-spenders 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 these differences in donut hole exposure affect drug consumption, as measured by December drug-days filled (i.e., the number days supplied, summed across all drugs). As expected, low-spenders have no variation in days filled by enrollment month. Middle-spenders fill substantially fewer days when they enroll early: February enrollees, for example, fill 25.5 fewer drug-days in December than September enrollees. And high-spenders show the opposite pattern, with February enrollees filling 31.8 *more* drug-days in December than September enrollees.[28] We take all estimates in this latter group with a grain of salt, due to the non-monotonicity of the enrollment month effect on consumption; the complexity of interpreting these coefficients means these beneficiaries are not the focus of our analysis. Appendix Table D.1 shows the specific drug

[28]Figure IV will show that these decreases are indeed interruptions, particularly at year-end: differences across enrollment months are small earlier in the year, then accelerate dramatically, peaking in December.

classes most affected by price increases, which include statins, beta-blockers, ACE inhibitors, and antidepressants; notably, opioids are unaffected.

Panel C of Figure III shows the relationship between December mortality and enrollment month. For low-spenders, there is no effect. For middle-spenders, we find a significant and large negative relationship between mortality and enrollment month, that mirrors the increase in drug consumption shown in Panel B. And for high-spenders, we see a large positive relationship, again contrasting with middle-spenders. Panel A of Table III summarizes these figures, and reports OLS estimates of β_2 from equation (4), both in terms of enrollment month (Column 2) and monthly pre-donut budget (Column 3) as scalars. Concretely, among middle-spenders, a \$100 budget increase—24.4% relative to the average enrollment-month budget in our sample—leads to a mortality reduction of 0.0164 percentage points (p.p.), or 13.9% of the base mortality rate. In low-spenders, the effect is small and insignificant. In high-spenders, the effect is positive, large, and significant.[29]

A central question is whether variation in mortality across enrollment (or birth) months is confounded by factors other than cost-sharing. Recall that we have already seen two reassuring facts in this regard. First, no such variation in mortality exists in the first few months after enrollment, before cost-sharing and drug consumption trends start to diverge across enrollment months (see Table II, Panel B). Second, any confounder would have to correlate not only to enrollment month, but also to the exact *combination of enrollment month and initial spending patterns* we would expect, given Medicare policy: large for medium-spenders, absent for low-spenders, and opposite-sign for high-spenders.

Nevertheless, we directly address two known potential confounders. The first and most obvious is age: those who enroll earlier in the year are older than those who enroll later, so our results indicate older enrollees die more—not a novel finding. We emphasize, however, that if aging alone were responsible, this trend would not be affected by initial spending differences. In addition, a simple calculation using US life tables (from the Social Security Administration) illustrates why this is unlikely to be a concern. Comparing annual mortality rates for 65- vs. 66-year-olds, we estimate the effect of being one month younger is to decrease monthly mortality by roughly -0.001

[29]The significant coefficient is in the enrollment month specification, but not the pre-donut budget specification, which is not surprising. We would not expect the latter to be a good way to scale enrollment month for those who have greatly exceeded the donut hole budget and entered the catastrophic phase.

p.p., or -0.76%.[30] This is quite similar to the effect of enrollment month from the low-spenders, albeit imprecisely estimated, which in the absence of enrollment month effects on budget are likely to be solely attributable to age: -0.00036, or -0.68% (Table III). Another data point comes from a null distribution of effects of enrollment (or birth) months on mortality, across many observably similar samples lacking enrollment month effect on drug budgets (which we describe in Section III.C, below). This yields a median estimate of -0.53%, which again likely reflects the effect of being one month younger in these populations. In other words, several different methods of calculating the aging effect all give fairly consistent estimates, between -0.5 and 1%—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 p.p. or -9.49%. (The equivalent estimate based on birth month is -0.009 p.p., or -7.7%, shown in Appendix Table C.2, Panel A.) This gives us confidence that aging is only a small part of the relationship we observe.

A second potential confounder is health differences by birth season, which has been suggested to result from disease seasonality (Currie and Schwandt, 2013) or selection (Buckles and Hungerman, 2013).[31] Most of the literature focuses on peri- and post-natal outcomes, but two large studies explore later-life outcomes in populations similar to our own. In a study of life expectancy at age 50, Doblhammer and Vaupel (2001) find that mortality peaks among May births, a pattern that is consistent across multiple cohorts in the Northern Hemisphere (Austria and Denmark; it is exactly reversed in Australia). This echoes a Nurses' Health Study by Zhang et al. (2019) that finds cardiovascular disease mortality peaks among April births (although that study found no overall mortality differences). In our sample, by contrast, mortality peaks among February births, and May births are in the middle of our distribution of mortality effects (Appendix Table C.2, Panel B shows these effects by birth month, rather than enrollment month, for comparability to this literature). In addition, our mortality effects are again orders of magnitude larger.[32] To summarize, the birth

[30] 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. We interpret the difference, -0.013 p.p., as the effect of being one year younger on monthly mortality. We translate this annual effect into a monthly effect, then divide by mean monthly mortality to get the percent decrease.

[31] While a full review of this literature is beyond the scope of our work, we refer the reader to Currie and Schwandt (2013) for an excellent summary, and a very rigorous empirical exploration of mechanisms.

[32] 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 remaining e_{50} of 27.5 years. Without access to the full life table, converting between e_{50} and annual mortality risk is not possible, but using the 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 middle-spenders.)

seasonality literature demonstrates small, cyclical increases in mortality among those born in April and May, while we find large mortality increases earlier in the year, that increase linearly with enrollment month, and that are larger and correlate with Medicare drug budgets.

III.C Falsification Tests

As shown in Figure III, the effect of enrollment month on mortality varies across spending bins, following the pattern of idiosyncratic changes in drug budgets created by Medicare policy. We consider this our first falsification check because it would be so hard for potential confounders to match the very specific set of idiosyncratic budget limits, initial spending, and calendar-month effects we exploit. We build on this insight to develop a more comprehensive set of falsification checks that replicate our analysis in a range of closely-related settings lacking an enrollment month (or birth month)–drug budget link. This lets us further verify that our findings are only seen in the presence of Medicare budget limits on drug consumption.

First, Figure IV isolates the mortality effect to the end of the first calendar year: it appears just as differences in drug consumption peak in December, and disappears as soon as prices reset and drug consumption re-equalizes in January. The Figure is constructed by following our main population of middle-spenders over time, before and after December. In each calendar month, we summarize the effect of enrollment month on being in the donut hole (Panel A), drug consumption (Panel B), and mortality (Panel C), via a linear coefficient. For example, the red point in Panel A is the coefficient on enrollment month, from a regression of donut hole entry in December on enrollment month. This is analogous to ϕ_1 from Equation 3, but estimated in all middle-spenders. In other words, this red point summarizes the middle panel of Figure III, Panel A as a single linear coefficient. Each point on the graph similarly represents an estimate of the enrollment month effect on an outcome, one for each calendar month, from August of the first calendar year to May of the second year.[33] Panel A shows that the effect of enrollment month on donut hole entry grows smoothly over time in year one, as more beneficiaries enter the donut hole. It then disappears in January of year two, when all enrollment months re-enter the initial coverage phase.

[33]For months prior to December, we exclude those who have not been enrolled for three months, in order to identify a similar group of middle-spenders over time (e.g. September regression includes February-June enrollees). We could not produce stable estimates prior to August of year one, as the sample size decreases with each month due to fewer enrollment months.

Similarly, each point in Panel B is the estimated linear coefficient of monthly drug-days filled on enrollment month (analogous to α_1 from Equation 2, but estimated in all middle-spenders). Differences in consumption between early and late enrollees appear as early as September, and this effect grows steadily over the next 3 months. But in contrast to Panel A, the magnitude abruptly jumps up in December: the effect on drug consumption more than doubles (1.75 drug-days difference per enrollment month in November, vs. 3.65 in December). These consumption gradients completely 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. Indeed, 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, conditional on filling in January; estimate not shown). These patterns fit with the beliefs documented in our survey (Figure I): patients view short interruptions in their drugs as largely innocuous, and put off filling until prices reset.

Finally, Panel C presents estimated linear coefficients of monthly mortality on enrollment month ($\gamma_{1\mathcal{M}}$ from Equation 4). Recall that our balance checks showed no differences in mortality in the first three months of enrollment. This Panel shows a similar analysis, by calendar month rather than enrollment month. We find no significant mortality gradient across enrollment months from August to November—then a large a significant effect in December. In January, just as prices reset 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 show that mortality increases are a transient phenomenon tied to abrupt increases in drug interruptions in December, and December alone. This is close to what we would expect based on our effect size and power calculations (Appendix Figure C.3).

Our second falsification test leverages policy variation over time, to show that the mortality effect is proportional to the degree of cost-sharing mandated by evolving Medicare policy. In 2011, the donut hole began to close, allowing us to compare mortality effects of enrollment month in our main sample of middle-spenders, before vs. after a policy change that reduced cost-sharing faced by earlier enrollees. Panel B of Table III shows estimates based on Equation (4), but with enrollment month interacted with indicators for pre- vs. post-policy change, and restricted to middle-spenders. The effects are larger before the attenuation of the donut hole, with a mortality reduction of 0.0210

p.p. per $100/month pre-gap budget increase, isolated to 2007-2010. In other words, 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.

Our most comprehensive set of falsification tests situates our main estimate in an empirical "null distribution" of enrollment (or birth) month effects on mortality, from a range of related populations and time periods. The placebo effects are estimated in samples of Medicare beneficiaries who are similar to those in our estimation sample, but lack the idiosyncratic link between enrollment month and drug budgets. As a result, we expect this distribution 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 left tail. Our first set of estimates extends the analysis in Figure IV to 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: shown separately in Appendix Figure C.1, Panel A). Our second set replicates the analysis in older dual-eligibles, who do not face cost-sharing. We pool all years together, then split into subsets defined by demographic factors and geography (46 estimates: Appendix Figure C.1, Panel B). Third, we broaden to a larger set of beneficiaries 66 years old and above—non-duals, older duals, and disabled dual 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 middle-spenders (459 estimates: Appendix Figure C.1, Panel C).[34]

Figure V reports the distribution of results across these 541 falsification samples. We rank estimates of the enrollment month effect by magnitude on the x-axis.[35] 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: 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 low-spenders). The estimate from our main analysis, -9.49%, is shown in red. It is at the 97.4th percentile of mortality effects, larger in absolute magnitude than 527 of 541 placebo estimates overall (and more negative than 532 of 541 estimates, or 98.3%). Appendix Figure C.2 shows similar results in a plot of t-

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

[35]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.

statistics from these regressions. This 'omnibus' test complements the more tailored falsification tests above, and builds confidence that our observed mortality effect is in fact due to difference in the pre-donut budget faced by enrollees, rather than any spurious correlation.

III.D Are These Effect Sizes Medically Plausible?

Table IV shows that a consistent finding of prior work is true in our setting as well (Einav et al., 2018; Brot-Goldberg et al., 2017): patients cut back on many drugs that have been shown to prevent life-threatening adverse events in randomized trials: cardiovascular (e.g., statins, antihypertensives), glucose-lowering (e.g., insulin), and respiratory (e.g., steroids, inhalers). For each of these three classes, Column (1) shows the fraction of middle-spenders 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).[36]

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 plausible. First, it is a common misperception that drugs for chronic diseases work slowly. Certainly, clinical trials for these drugs last many years—but we should not conflate the time scale required to *measure* effects for rare, noisy outcomes like mortality, 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 short-term 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, which

[36] We again caution against any effort to tie the magnitude of changes in estimated consumption to changes in estimated mortality in the same time period, given the bias discussed in Section III.B: lagging mortality effects of drugs, combined with inter-temporal substitution across periods, means that scaling the mortality effect by consumption will bias estimation of the effect of consumption (and thus consumption changes due to cutbacks).

is why clinical guidelines mandate initiating a statin immediately on diagnosis of heart attack as a result (Cannon and Freeman, 2023). Several other chronic medications 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 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, inferring the effect of drug interruptions from the effect of drug initiation may be misleading in the presence of rebound effects. 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, can precipitate a potentially dangerous set of effects in several settings. This idea has entered the popular consciousness in the setting of opiate withdrawal, but opiates are far from the only drug class with such effects. Indeed, rebound effects have been noted for 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 *in vitro* experiments, or older, idiosyncratic studies.[37] Our results add new weight to these recommendations, suggesting that drug interruptions can precipitate serious adverse events. Interestingly, we find that 69% of December deaths in our sample occur *outside* the hospital, suggesting a catastrophic, sudden event.[38]

If our results fit with medical intuition, they are more discordant with some parts of economics, as they run counter to the predictions of several standard models of behavior. Consider a population of patients prescribed a drug by a doctor, for whom treatment benefit is heterogeneous across indi-

[37]Even for beta-blockers, perhaps the class of drugs for which there is the most evidence, 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).

[38]This estimate comes from enrollees in standalone PDPs, 52% of our sample, in whom we also observe hospitalizations. This combination of high Medicare Advantage prevalence and high out-of-hospital death rate means we are under-powered to detect changes in hospitalizations in this study.

viduals, as research indicates (Chandra and Skinner, 2012; Chandra and Staiger, 2020). If the price of that drug suddenly increases, traditional models predict that cutbacks should be concentrated among individuals for whom the drug yields fewer benefits. Under moral hazard, the marginal drugs patients drop are disproportionately low-benefit. Likewise, a Roy model of patient decision making with private information on heterogeneous treatment effects predicts 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 should thus cut back less when the price increases. So a key question here is: *who* is interrupting their consumption? If interruptions occur largely in patients with low treatment benefit, 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 interrupt their consumption due to relatively small price increases suggests, it would help make sense of the large mortality effects we see; it would also raise a fascinating new set of questions about why this might happen.

To develop a measure of an individual's health benefit from a drug, we use machine learning to form predictions on patient risk. Focusing on a set of drugs used to prevent key adverse events, we assume that the benefit of a given drug is proportional to the baseline risk of those outcomes the drug prevents. For example, we assume the benefit of a statin is proportional to the risk of heart attack and stroke. This assumption is supported by both a substantial body of evidence, as well as specific clinical guidelines, particularly for cardiovascular drugs. Major randomized trials (e.g., JUPITER, HOPE-3, CARDS, and ASCOT, 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). Clinical guidelines also reflect this assumption, for example use of the American College of Cardiology 10-year risk calculator to guide treatment for cardiovascular disease. There is similar medical consensus and biological plausibility, if less strong empirical evidence, for diabetes and respiratory drugs.[39]

Concretely, we identify three important drug classes M—cardiovascular, diabetes, respiratory—

[39]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.

and compile a list of observable adverse outcomes that the drugs are prescribed to prevent: heart attack and stroke for cardiovascular medicines, diabetic complications (e.g., foot amputation) for hypoglycemic medicines, and respiratory failure for inhalers and steroids. This allows us to define indicators Y_M, one for each of the three drug classes, that indexes whether a beneficiary experienced an adverse event preventable by drug M. We form separate predictive models for each outcome, and restrict to those who are not taking class M (e.g., when predicting risk of heart attack or stroke, we exclude patients on statins and other medications for coronary artery disease), to obtain a prediction on the risk of complications if untreated.[40] These models use a beneficiary's demographics and initial 90-day claims to predict the likelihood of adverse events over the next 270 days, and are trained on an entirely separate sample of dual-eligible 66+ year olds, to ensure our predictions are out-of-sample. We turn the model's continuous risk predictions into simple indicators, $\hat{Y}_{Mi}$, that index the highest-risk one-third of the sample, based on where risk begins to increase rapidly (see Appendix Figure D.1). Additional details are in Appendix D.

Strikingly, the highest-risk beneficiaries cut back at least as much, if not more, on those medications that benefit them the most. We run Equation 2 with enrollment month interacted with $\hat{Y}_{Mi}$, separately for each drug class M and restricting to middle-spenders as usual. Table IV shows the results, which are especially pronounced for cardiovascular drugs: for each \$100/month budget decrease, low-risk patients fill 0.598 fewer cardiovascular drug-days (a 1.5% reduction), while high-risk patients fill 2.46 fewer (3.6%).[41] This finding is incompatible with standard economic models of behavior, or private information: those at high risk of a cardiovascular event should have the most inelastic demand for treatment, proportional to their benefit from the drug.[42] 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

[40] In potential outcomes notation, we wish to predict Y_M^0, not Y_M^1, as our proxy for the benefit of drug M. Naturally this choice of prediction target also induces selection bias, as noted in Mullainathan and Obermeyer (2019), 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 D.1 shows that true risk rises monotonically in predicted risk for the treated just as the untreated, establishing face validity of the predictor irrespective of treatment status.

[41] We emphasize the relative, not the absolute, magnitude of these changes, which reflect only December differences in consumption and not cumulative differences over the year; see Appendix B.

[42] As noted above, our population of middle-spending patients who have been prescribed a basket of drugs and then must decide whether or not to continue them after a sudden price increase. Among this population, which is already filling 126 drug-days per month, lower- vs. higher- risk patients should be more elastic. These results do not apply to low-spenders, who are at such low risk that they are not prescribed any medications, who would also have zero demand for treatment regardless of price.

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.[43] We do identify one potential contributor to this effect: a subgroup of beneficiaries chooses to fill *no* drugs when prices increase—no matter how many drugs they were on prior to the price shock, or their individual health risks. Indeed, those with higher initial consumption are in fact *more* likely to interrupt their entire regimen of prescribed drugs.[44] Table D.3 shows that enrollment month makes it more likely for all middle-spenders to fill not medications (Panel A), but also shows that those in the higher two terciles of initial drug consumption (in the first 90 days of enrollment) are far more affected in both absolute and relative terms. Mechanically, this behavior results in large absolute reductions in drug use for higher-risk patients, who are on more drugs to begin with.

IV Conclusion: Errors and Misinformation

Clinical medicine and economics share a fundamental respect for individuals' preferences and decision-making. Our results, alongside a substantial literature in both fields, present a dilemma for this perspective, if individuals make decisions that are inconsistent with the general preference to stay alive. A rough calculation highlights the incongruous life-year valuations implicit in the price-driven cutbacks we study. We divide the effect of enrollment month on cumulative drug spending (from September-December, when drug-day differences begin to emerge across enrollment months), by the effect of enrollment month on mortality to infer the life-year valuations underlying the decision to interrupt drugs in our sample: \$11,321 (95% confidence interval: \$6,195-73,858).[45] Put another way, at a widely-used life-year valuation of \$100,000 per year (Neumann et al., 2014), a 65-year-old middle-spender in our sample would have to believe that she had at most 2.19 years left to live. This

[43] 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.

[44] This is reassuring that this phenomenon is not simply a floor effect, i.e. due to left-censoring: it is more common, not less, in those with more drug fills to begin with.

[45] 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 middle-spenders (we emphasize that this is a very approximate exercise and all the caveats from Appendix B apply). The inverse of this estimate is 'dollars per life-year,' which we divide 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.

contrasts sharply with average life expectancy in the general population at age 65—19.2 years—and observed outcomes in our sample: at a median follow-up period of 5 years, 93.2% of middle-spenders are still alive. This is hard to square with the idea that patients are equalizing marginal benefit with marginal cost of drugs, particularly in the absence of obvious zip code differences that might indicate income or liquidity constraints. It supports the idea 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 'behavioral hazard' has far-reaching implications for the design of health insurance, particularly as insurers place more emphasis on cost-sharing.

Behavioral economics provides several potential explanations for our results, in the form of predictable distortions in the cost-benefit calculus. Costs, for example, might be over-weighted relative to their true value for several reasons. If a patient arrives at the pharmacy counter to find that her drug basket has shot up in cost relative to her expectations, costs may be highly salient (Bordalo et al., 2013, 2020); if costs deviate from previously set reference points, they may be viewed as losses (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 or frictions (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). This last set of mechanisms in particular could explain the phenomenon of patients choosing to fill none of their medications in response to price increases. Time-specific factors in December may contribute: while the absence of zip-level effects suggest that household budgets around the holidays are less likely to explain our results, the administrative burdens of getting a doctor's visit or responses to questions may be larger than usual at this time. Exploring which of 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: the seriousness of interrupting drugs is simply not known to patients—and potentially also their doctors. Our survey of patients taking medications, while small, is some of the first evidence of its kind on how patients view short drug inter-

ruptions. We partnered with the firm Survey Healthcare Global (SHG) to recruit 200 patients ages 61-70 who reported taking at least one prescription medication, in July 2022. Survey respondents were selected to be similar to the middle-spenders in our sample: the median respondent was on at least 5 medications, and 75% indicated they had hypertension or high cholesterol. The survey took less than 5 minutes to complete and had a 100% completion rate.[46] As shown in Figure I, patients view short 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 beliefs are one explanation for a central driver of our results—that drug interruptions peak sharply in December—and may also explain why the phenomenon is so temporary: patients are willing to hold out for low prices just over the horizon in January, because they view short-term interruptions as innocuous. 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.

[46]The sample was drawn from a panel of over 600,000 patients and caregivers maintained by the firm; more information is available on SHG's website. The exact wording of the questions was: (1) "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?" Possible responses were: "<1 week", "1-2 weeks", "2-3 weeks", "3-4 weeks", ">4 weeks". (2) "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". For this question, 15 respondents selected "no change to your health," and thus do not contribute to the totals in the Figure.

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.

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.

CMS (2024). Costs in the coverage gap | Medicare.

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.

Dumas, E. O. and Pollack, G. M. (2008). Opioid Tolerance Development: A Pharmacokinetic/Pharmacodynamic Perspective. *The AAPS Journal*, 10(4):537.

Dupont, W. D. and Plummer, W. D. (1998). Power and sample size calculations for studies involving linear regression. *Controlled Clinical Trials*, 19(6):589–601.

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 Nonlinear Contract Design: Evidence from Medicare Part D. *The Quarterly Journal of Economics*, 130(2):841–899.

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.

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.

KFF (2023). An Overview of the Medicare Part D Prescription Drug Benefit.

Koracevic, G., Micic, S., Stojanovic, M., Tomasevic, M., Kostic, T., Velickovic Radovanovic, R., Lovic, D., Djordjevic, D., Randjelovic, M., Koracevic, M., and Ristic, Z. (2020). Beta blocker rebound phenomenon is important, but we do not know its definition, incidence or optimal prevention strategies. *Hypertension Research: Official Journal of the Japanese Society of Hypertension*, 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. (2021). Medicaid and Mortality: New Evidence From Linked Survey and Administrative Data. *The Quarterly Journal of Economics*, 136(3):1783–1829.

Mullainathan, S. and Obermeyer, Z. (2019). A Machine Learning Approach to Low-Value Health Care: Wasted Tests, Missed Heart Attacks and Mis-Predictions. Technical Report w26168, National Bureau of Economic Research.

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.

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.

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.

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)	(3)	(4)
	All	*Spending Group*		
		Low	Middle	High
Panel A: Demographics, Spending, Health				
White (%)	89.6 (30.5)	88.5 (31.9)	92 (27.1)	92.5 (26.4)
Female %	59.2 (49.1)	59.5 (49.1)	58.8 (49.2)	56.6 (49.6)
Standalone PDP (%)	51.7 (50)	46 (49.8)	64.2 (47.9)	71.2 (45.3)
Initial 90-day fills	5.35 (5.75)	3.16 (3.62)	9.75 (5.71)	16.9 (9.18)
One-year total spending ($)	1,478 (2,789)	626 (1,158)	2,828 (1,910)	9,185 (10,119)
One-year mortality (p.p.)	0.873 (9.3)	0.683 (8.24)	1.16 (10.7)	2.78 (16.4)
Panel B: Top 10 Drug Classes (%)				
Lipid modifiers	34 (47.4)	22.9 (42)	59.2 (49.1)	64.7 (47.8)
ACE inhibitors	20.5 (40.4)	16.4 (37.1)	30 (45.8)	30.8 (46.2)
Beta blockers	18.9 (39.2)	13.3 (34)	31.4 (46.4)	37.7 (48.5)
Thiazide diuretics	18 (38.4)	14.4 (35.1)	26.7 (44.2)	25.7 (43.7)
Antidepressants	14.1 (34.8)	8.8 (28.4)	24.8 (43.2)	39.7 (48.9)
Corticosteroids	13 (33.6)	8.5 (27.9)	22.4 (41.7)	33.5 (47.2)
Acid blockers (GERD)	12.4 (33)	6.8 (25.1)	24.3 (42.9)	37 (48.3)
Anti-infectives	11.6 (32)	8.9 (28.4)	17.4 (37.9)	23.7 (42.5)
Hypoglycemics (oral)	11.2 (31.5)	6.4 (24.5)	20.9 (40.7)	33.2 (47.1)
Decongestants	11 (31.3)	6.7 (25)	20 (40)	31.6 (46.5)
Observations	358,706	251,093	96,849	10,764

Notes: Column 1 shows mean (standard deviation) for the entire sample. Columns 2-4 show the same by initial 90-day spending group: low (1-70th within-enrollment month percentile), middle (71-97th), high (98-100th). 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. All participants are exactly 65 years old.

TABLE II
BALANCE OF KEY VARIABLES ACROSS ENROLLMENT MONTHS

	(1)	(2)	(3)	(4)
	Entire Sample		*Middle-spenders*	
	Mean	Enrollment Month Effect (Std. Error)	Mean	Enrollment Month Effect (Std.Error)
Panel A: Demographics and Key Characteristics				
White (%)	89.6	0.00403 (0.0222)	92.0	0.0219 (0.0378)
Female (%)	59.2	0.0582 (0.0356)	58.8	0.165** (0.0687)
Initial 90-day fills (count)	5.35	0.00364 (0.00419)	9.75	0.00374 (0.008)
Predicted mortality (p.p)	0.411	-0.000477 (0.000467)	0.458	0.00175 (0.00116)
Panel B: Initial Mortality (cumulative p.p., from enrollment)				
30 days	0.058	-0.000053 (0.0017)	-	-
60 days	0.128	0.0013 (0.0026)	0.082†	-0.0037 (0.0041)
90 days	0.201	0.0037 (0.0032)	0.181†	-0.0011 (0.006)

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

Notes: Panel A: Sample mean (Column 1) and coefficient on enrollment month (scalar), from regression of key 'pre-treatment' variables on enrollment month (Column 2), for the entire sample (n = 358,706). Predicted mortality is estimated by fitting a model with demographics and initial (3-month) drug claims to predict subsequent (9-month) mortality, in an independent sample of 66+ year-old dual enrollees. Columns (3-4) restrict to middle-spenders (n = 96,849), based on initial spending in the first 90 days of enrollment. Panel B: Sample mean (Column 1) and coefficient on enrollment month (scalar), from regression of mortality in the first 30, 60, and 90 days after enrollment (Column 2). Columns (3-4) restrict to middle-spenders, based on initial spending in the first 30 days of enrollment (unlike our main specification, which uses 90 days). Because we use the first 30 days to assign spending bins, we are unable to report the estimate for 0-30 days, and mortality rates in middle-spenders reflect a one month shorter period than for the entire sample (i.e., 31-60 and 31-90 days, denoted by the †). Balance checks on initial spending are in Appendix Table A.2 and discussed in Section III.A.

TABLE III
MORTALITY EFFECTS OF CHANGES IN DRUG BUDGETS

	(1) December Mortality (p.p.)	(2) Enrollment Month Effect (p.p./mo)	(3) Pre-Donut Budget Effect (p.p./$100)
Panel A: Enrollment Month Effect, By Initial Spending			
Low-spenders	0.053	-0.000362	-0.000771
		(0.00199)	(0.00317)
Middle-spenders	0.118	-0.0112**	-0.0164**
		(0.00498)	(0.00786)
High-spenders	0.288	0.0469**	0.0549
		(0.0219)	(0.0402)
Difference: Middle vs. Low	-	0.0109**	0.0156*
		(0.00535)	(0.00846)
Difference: Middle vs. High	-	-0.0581***	-0.0713*
		(0.0225)	(0.0411)
Panel B: Pre- vs. Post-Donut Hole Closing (Middle-Spenders)			
Full Donut Hole (2007-10)	0.120	-0.0137**	-0.021**
		(0.00626)	(0.00993)
Closing Donut Hole (2011-12)	0.113	-0.0063	-0.00993
		(0.00814)	(0.0129)
Difference: Full vs. Closing	-	0.00745	0.0111
		(0.0103)	(0.0163)
Panel C: Key Falsification Estimates (Middle-Spenders)			
Main Sample: December, age 66	0.12	0.003	0.0076
		(0.0051)	(0.0083)
Duals: December, age 65	0.38	-0.0013	-0.0144
		(0.013)	(0.021)
Disabled: December, age 64	0.593	-0.0017	-0.0066
		(0.021)	(0.032)

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

Notes: Panel A: December mortality rate (Column 1) and coefficient γ_2 on enrollment month from equation (4) (Column 2), by initial spending: low-spenders (lowest 70% of initial 90-day spending: unlikely to enter the donut hole irrespective of enrollment month); middle-spenders (71-97th percentile: more likely to enter the donut hole if enrolling earlier; and high-spenders (98-100th percentile: likely to enter both the donut hole then the catastrophic coverage if enrolling earlier). Column (3) translates enrollment month into 'pre-donut budget' (in $100/month) before full cost-sharing. Panel B: Coefficient on enrollment month and pre-donut budget for middle-spenders ($n = 96,849$), before vs. after the donut hole began to close in 2011. 'Difference' rows are pairwise tests for equality of coefficients. Panel C: Selected falsification tests, estimating enrollment month effect on mortality in settings with no policy link between enrollment month and drug prices. Robust standard errors are in parentheses.

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

	(1)	(2)	(3)	(4)	(5)	(6)
	All		*Lowest 2/3 Risk*		*Top 1/3 Risk*	
	Mean	Est. (SE)	Mean	Est. (SE)	Mean	Est. (SE)
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: Effect of cost-sharing on December drug consumption, for key drug classes. Cardiovascular drugs includes statins, beta-blockers, ACE inhibitors, calcium channel blockers, angiotensin receptor blockers, and thiazide diuretics. Diabetes includes both insulin and oral hypoglycemic agents. Respiratory includes inhaled and oral treatments for chronic pulmonary disease. Column (1) shows mean number of drug-days filled by middle-spenders in December (days supply, summed across all prescriptions filled, and grouped by drug class). Column (2) presents estimates (and robust standard errors) of the effect of pre-donut budget (in $100s) on drug-days. Columns (3)-(6) show mean drug days and similar regression estimates, by 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. (For "all classes", we use predicted cardiovascular event risk, since both cardiovascular drugs and cardiovascular mortality are the most common.)

Figures

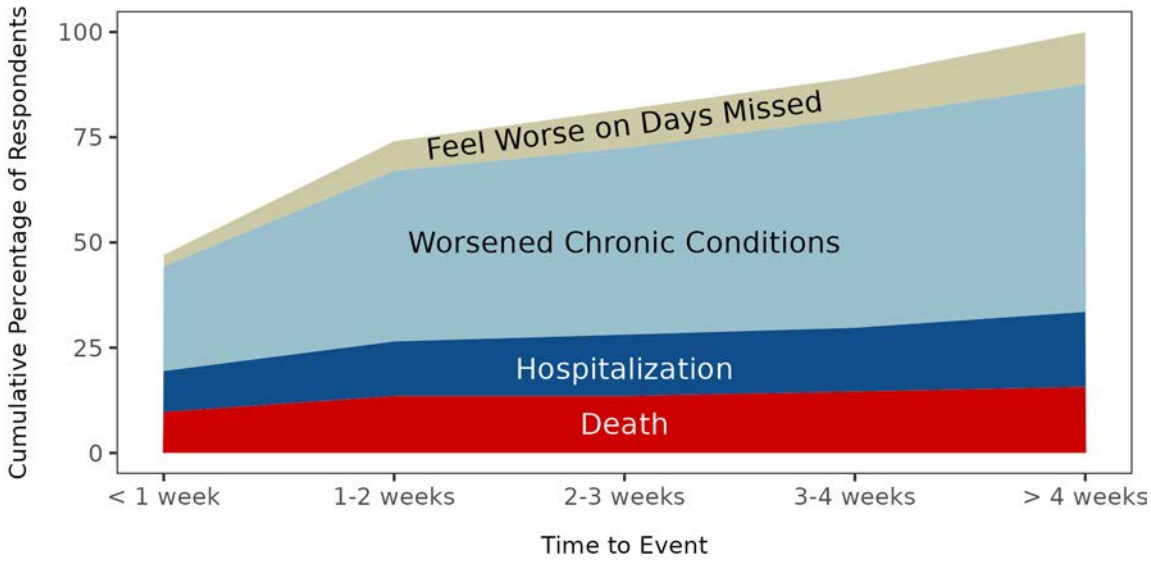


FIGURE I
Patient Beliefs on the Health Effects of Drug Interruptions

Notes: Results of a survey of Medicare-age patients taking at least one prescribed medication. The y-axis shows the cumulative percent 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 respondents thought it would take for such a health problem to occur.

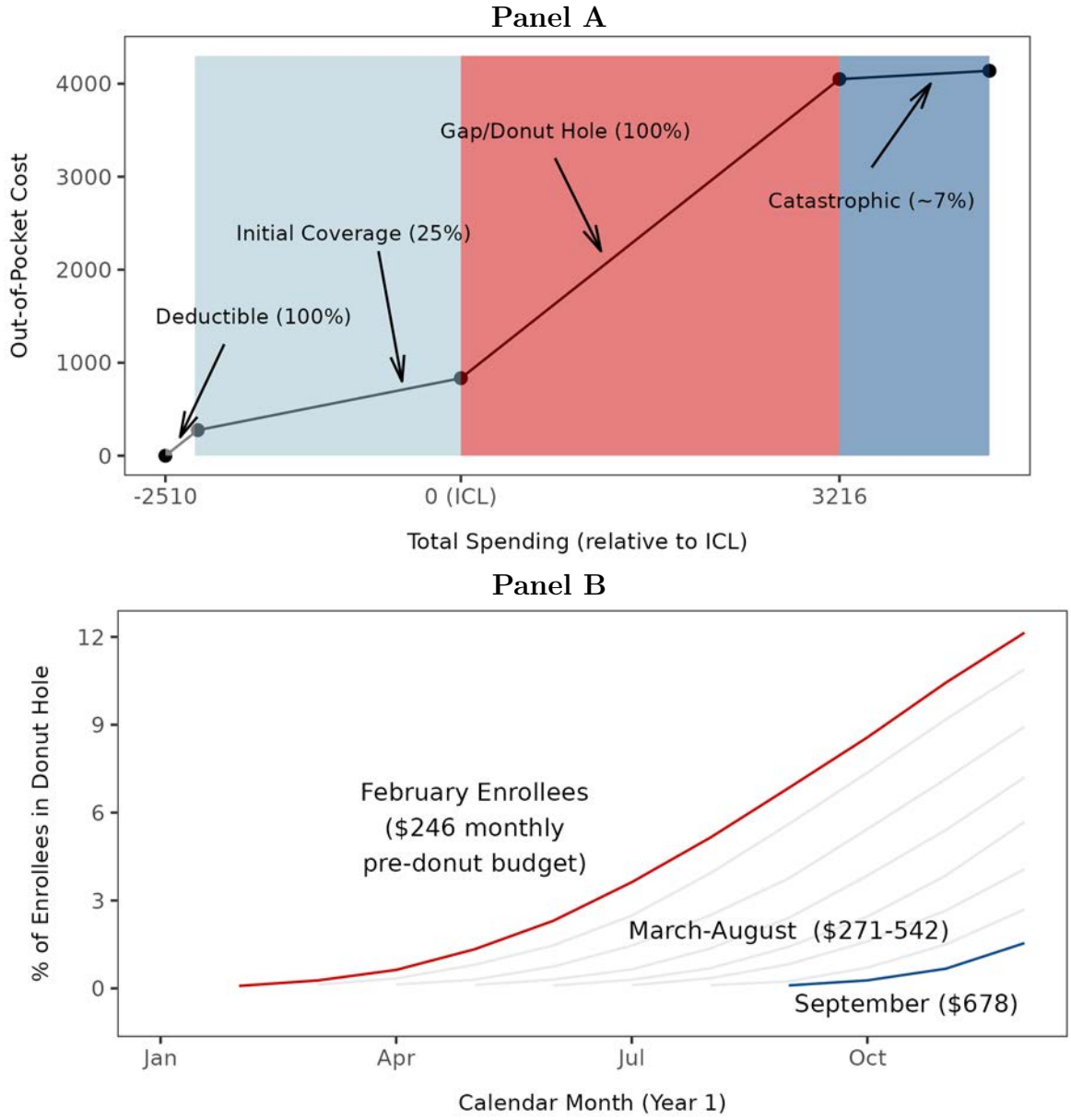


FIGURE II
Medicare Drug Benefit Design and Donut Hole Consequences

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.

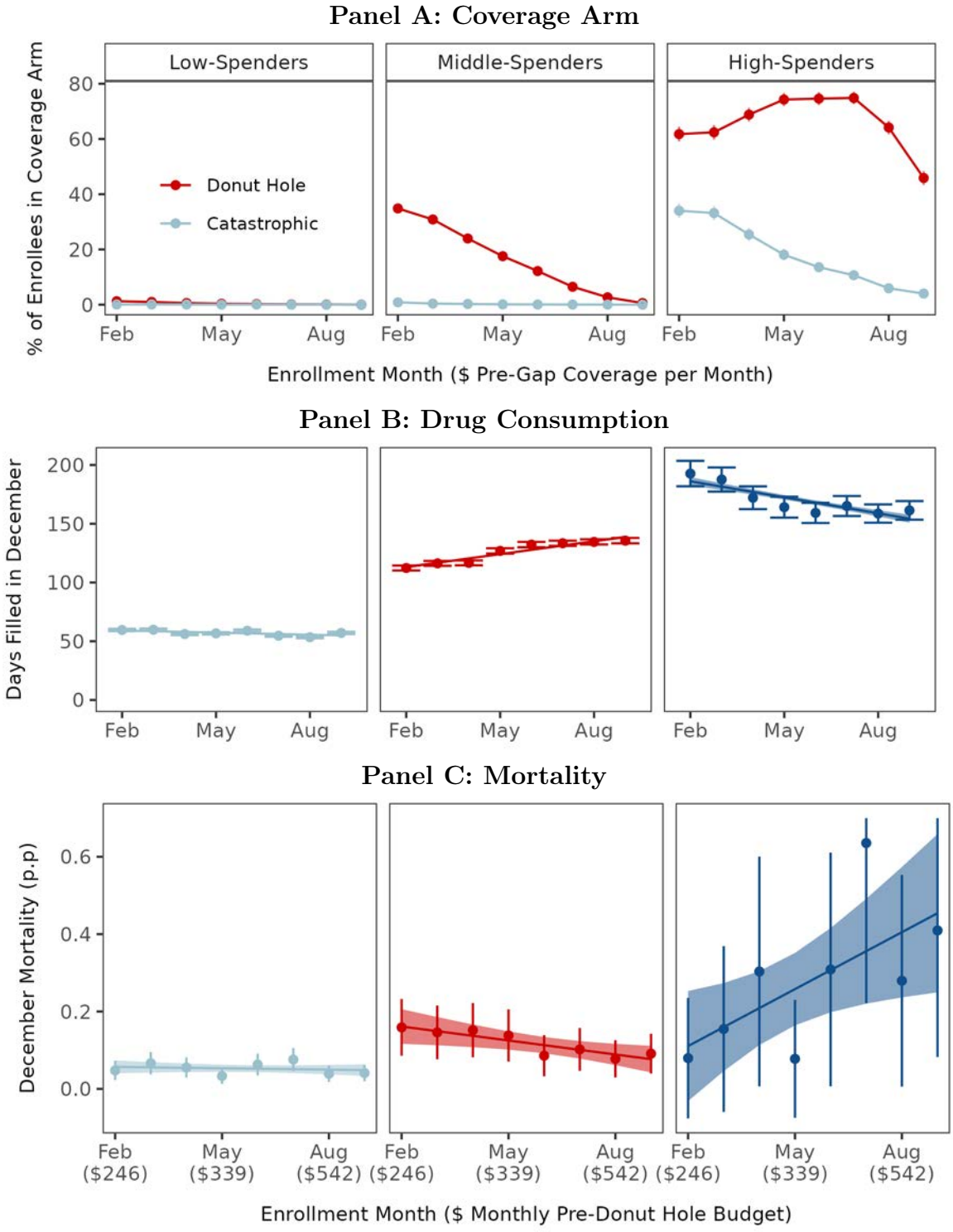


FIGURE III
Effect of Enrollment Month on End-of-Year Coverage Arm, Drug Consumption, and Mortality

Notes: Panel A: Fraction in the donut hole or catastrophic coverage at year-end (y-axis) by enrollment month (x-axis) and initial 90-day spending bin (horizontal sub-panels). Panel B: Drug-days filled in December (y-axis) by enrollment month and spending bin. Panel C: December mortality (y-axis) by enrollment month (translated into monthly pre-donut budget, in parentheses) and spending bin. 95% confidence intervals are shown; one (July, high-spenders) is truncated.

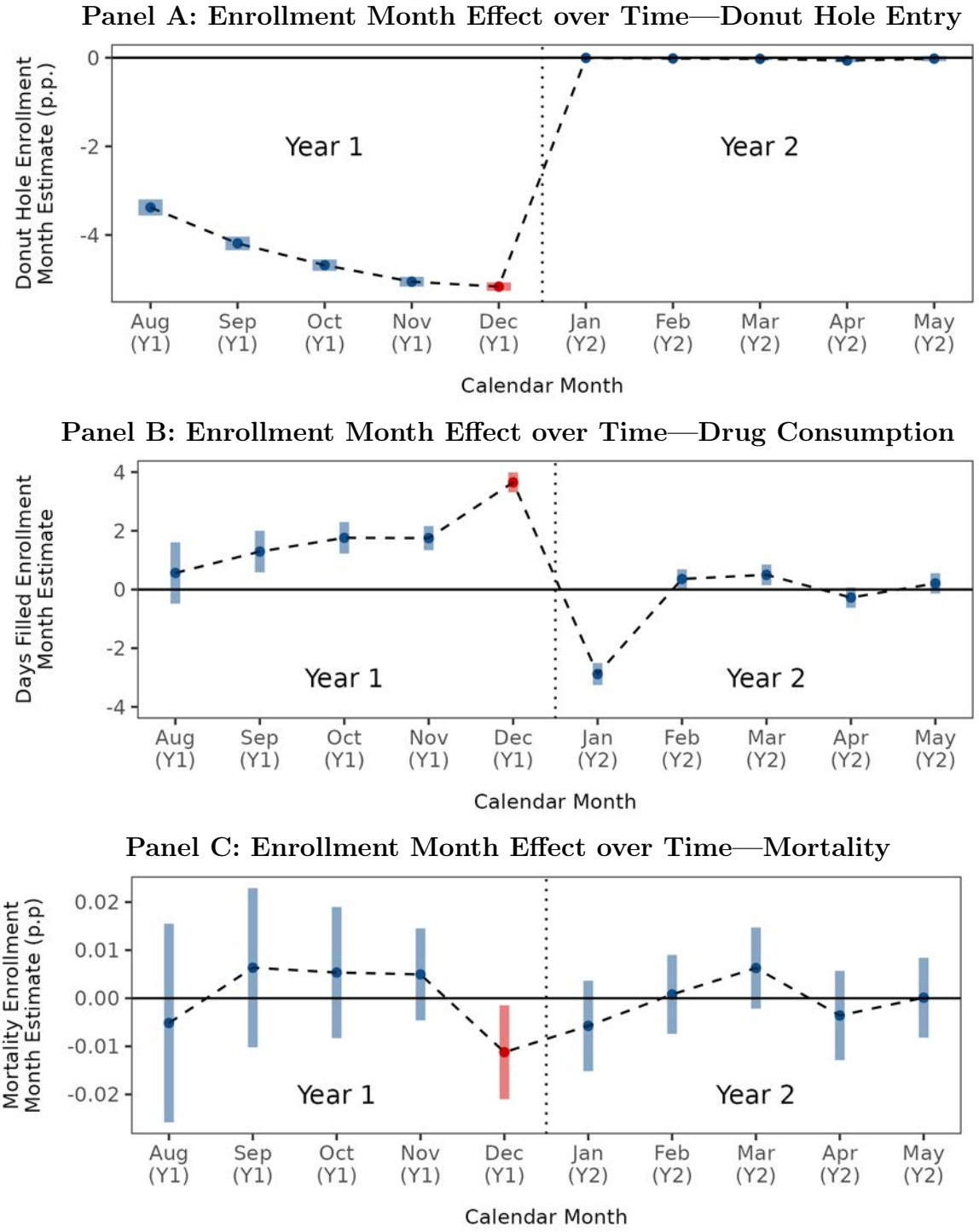


FIGURE IV
Evolution of Enrollment Month Effects Over Time, Before and After December (Year 1)

Notes: Each point shows the linear coefficient (and 95% confidence interval) measuring the effect of enrollment month over time, from August of calendar year 1 to May of year 2, for middle-spenders. Panel (A) shows the effect on donut hole entry, (B) drug consumption, and (C) mortality. The red points show enrollment month effects in the calendar month of December (i.e., each point summarizes one of the three middle panels of Figure III, as a linear coefficient). For example, Panel (C) shows γ_{1t} from the regression $Y_{it} = \gamma_{0t} + \gamma_{1t}Z_i + X_i\gamma_{2t} + \gamma_{year} + \gamma_{plan} + \epsilon_{it}$, for middle spenders in months $t = \{8, 9, ..., 12, 1, 2, ..., 5\}$, and the red point is $\gamma_{1,12}$.

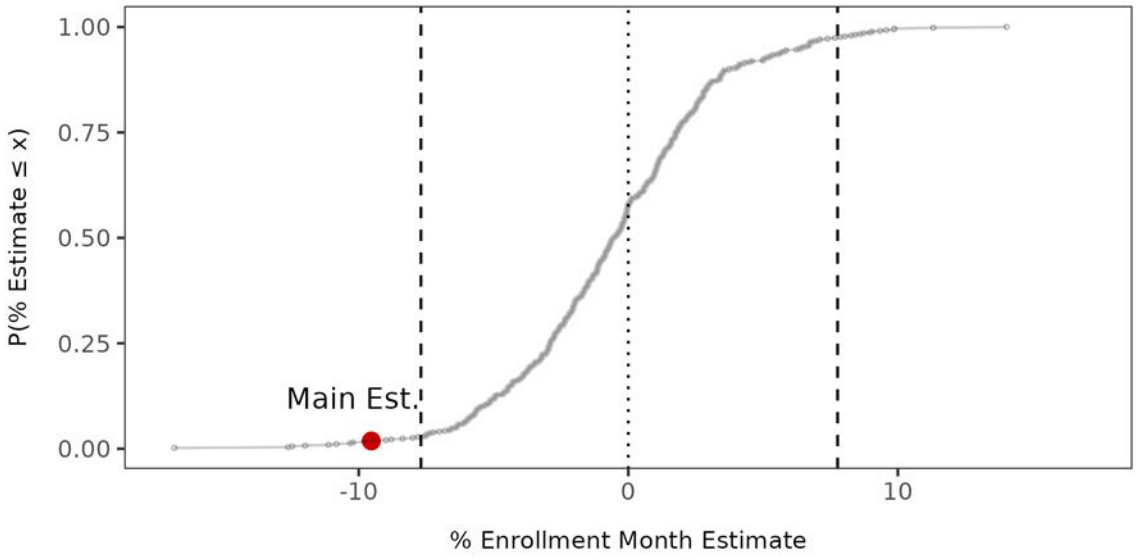


FIGURE V
Distribution of 'Placebo' Estimates of Enrollment or Birth Month on Mortality

Notes: Regression estimates of the effect of enrollment month or birth month on mortality, for middle-spenders, in a variety of settings lacking an enrollment month-drug budget link: 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 C.1 provides further detail on the separate types of tests that contribute. Vertical lines show the 2.5 and 97.5 percentiles. Our main (non-placebo) estimate from Table III is shown as a red dot.

Appendix A Prices, Part D Design, and Enrollment

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

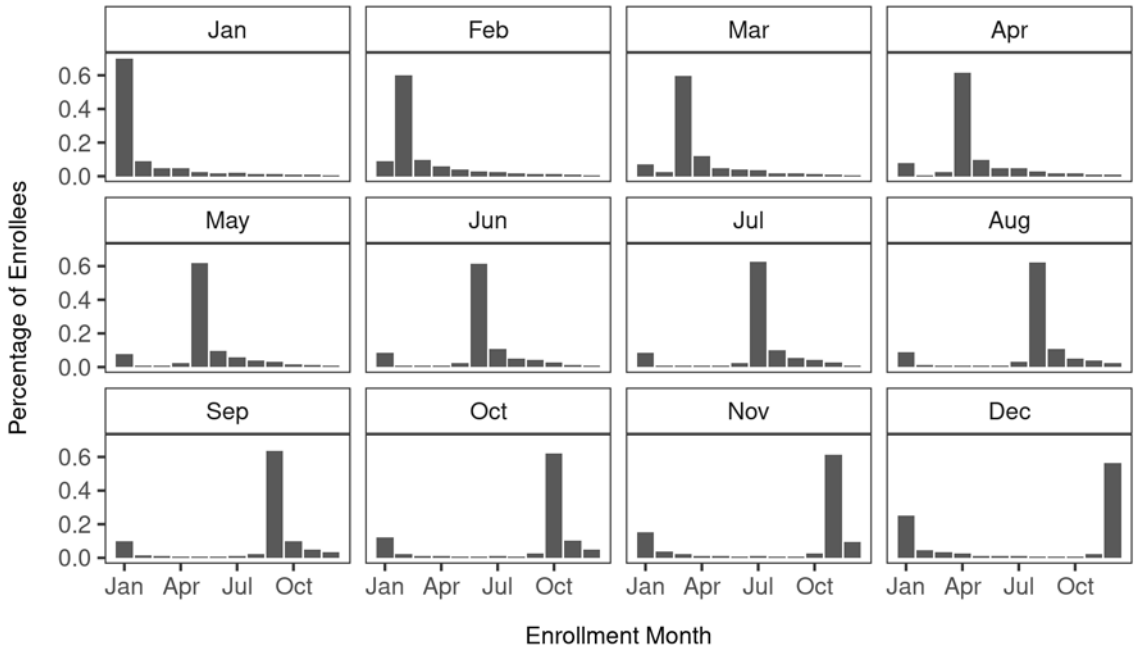


FIGURE A.1
Enrollment Timing by Birth Month

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.

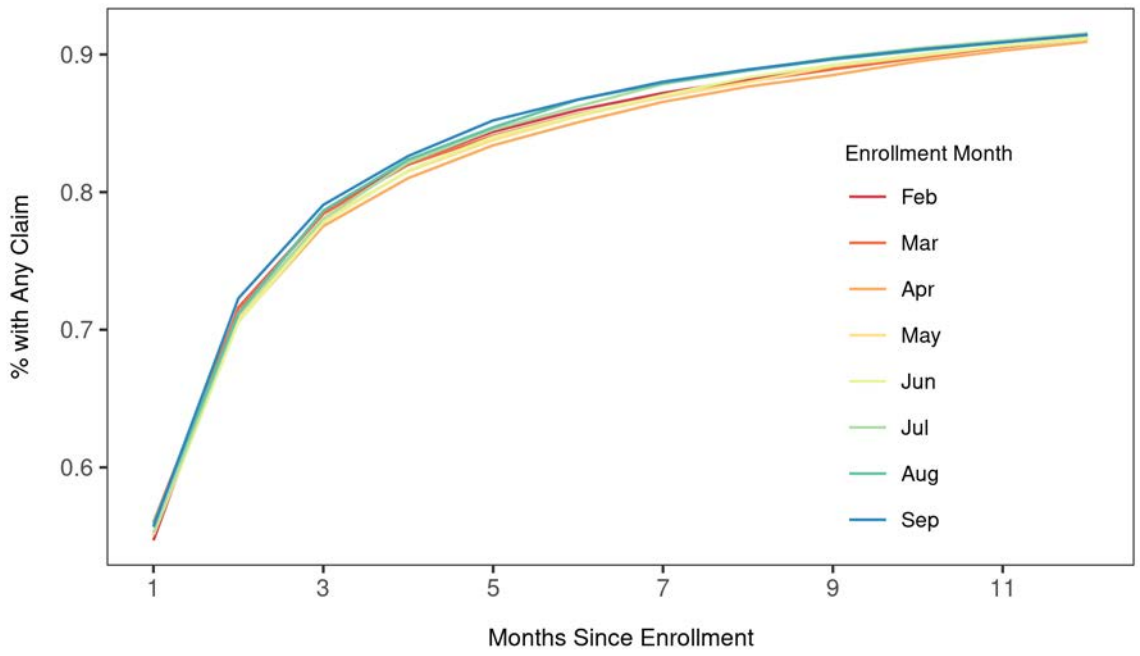


FIGURE A.2
Utilization Ramp-Up

Notes: This figure plots the percentage of beneficiaries in each enrollment month (shaded lines) with a claim for at least one Part D drug fill since enrolling (y-axis), by the number of months since enrollment (x-axis).

TABLE A.2
FORWARD-LOOKING BEHAVIOR AFTER ENROLLMENT

	(1)	(2)		(3)	(4)
	Total Spending (30 Day)			Total Spending (Cumulative)	
Days From Enrollment	Mean	Enrollment Month Est.	Days From Enrollment	Mean	Enrollment Month Est.
1-30	124.9	-0.161 (0.219)	1-30	124.9	-0.161 (0.219)
31-60	113.4	-0.0533 (0.206)	1-60	238.4	-0.214 (0.355)
61-90	118.1	0.395* (0.206)	1-90	356.5	0.181 (0.493)
91-120	121.6	1.14*** (0.215)	1-120	478.1	1.32** (0.643)

Notes: Columns (1) and (3) present the mean total spending by days since enrollment. Columns (2) and (4) present corresponding estimated coefficients on enrollment month from OLS regressions of spending on enrollment month.

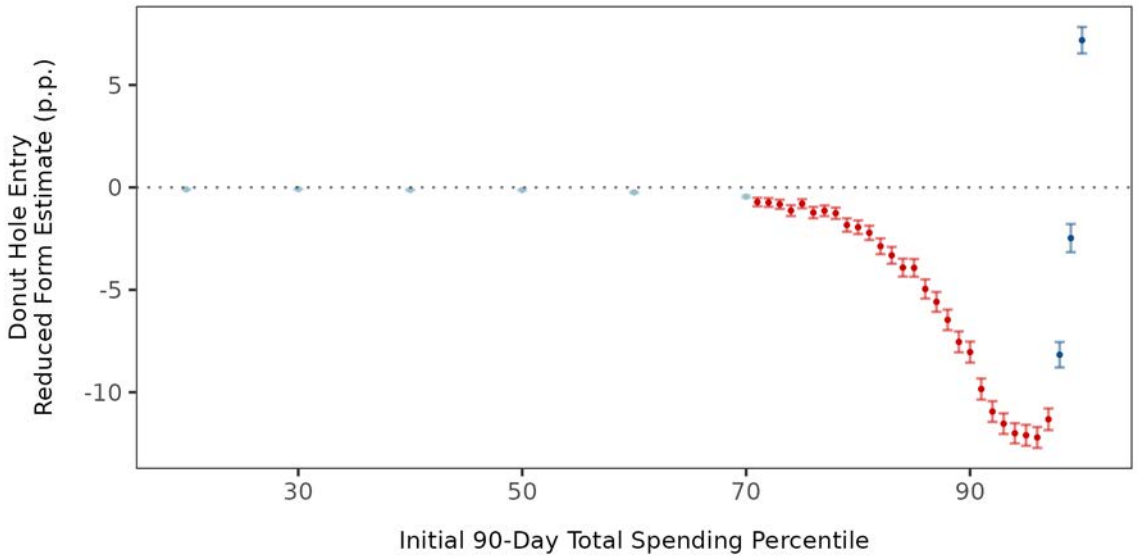


FIGURE A.3
Effect of Enrollment Month on Donut Hole Entry, by Fine Bin of Initial Spending

Notes: Each point represents the linear coefficient from a regression performed within a percentile of initial 90-day spending (ϕ_{1j} from equation (3)). 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). The first and second deciles of initial spending, as well as each subsequent decile below the 71st percentile, are grouped because there is not enough variation to define unique percentiles. 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.

Appendix B Rationale for Reduced Form Model and Bias in Two-Stage Least Squares

Our understanding of the causal chain linking the Medicare policy we study to patient health is shown in Figure B.1, Panel A.

(a) 'True' Model (b) Reduced Form Model

FIGURE B.1
Proposed Causal Relationships in (a) the 'true' model and (b) our ultimate reduced form model. Y denotes mortality; Q drug consumption; S drug spending; P drug prices; X, W observed or unobserved patient factors; and Z enrollment month. Function $g(S, Z, t; \theta)$ is the Medicare policy that compares cumulative spending from Z to period t to policy budget limits θ.

Given our interest in the effect of budget-induced drug consumption changes on mortality, a natural choice of model would be two-stage least squares, instrumenting for drug consumption using enrollment month. Unfortunately, this strategy suffers from several flaws. As the Figure B.1 shows, the effect of Z is mediated via a feedback loop connecting (Q, P, S), indicating the complex inter-temporal relationship of these variables with each other and with mortality Y. To illustrate the problem, consider a simple setting where we observe patients over two periods, $(t-1)$ and t, and model mortality as:

$$Y_{it} = \beta_1 Q_{it} + \beta_2 Q_{i,t-1} + \beta_3 X_i + \nu_{it} \tag{5}$$

Drugs can have both short- and long-term effects on health, so both present and past consumption may affect mortality. Clearly, choosing either Q_{it} or $Q_{i,t-1}$ as the endogenous variable for the first stage will violate the exclusion restriction; the same problem applies to prices or spending. We cannot easily get around this by simply redefining our endogenous variable as, for example, cumulative consumption $(Q_{i,t-1} + Q_{it})$: in any period t, the measurement of consumption and all other variables in $(t-1)$ is correlated with the instrument—enrollment month—so we will not

observe consumption for later enrollees as we go back in time. Of course, we can calculate cumulative consumption since period $(t-k)$—i.e., redefine the endogenous variable as $Q_{i,t-k:t}$—but only if we drop the latest k enrollment months from the sample, because we do not observe $Q_{i,t-k}$. In fact, in practice we must drop the latest $(k+2)$ months, because consumption ramps up over the first 1-2 months of enrollment and only reaches steady-state in month $(z+2)$.

It is worth noting one specific consequence of this problem: any effort, implicit or explicit, to estimate the effect of consumption Q_{it} on same-period mortality Y_{it} will be biased up. To see why, consider a model that tries to estimate the effect of Q_{it} on Y_{it} (i.e., β_1 from Equation 5) in isolation, without accounting for the effect of prior consumption $Q_{i,t-1}$ (i.e., β_2). The resulting estimate $\tilde{\beta}_1$ will be biased if $\beta_2 \neq 0$, unless $Q_{it} \perp\!\!\!\perp Q_{i,t-1} \mid X_i$. So a lot depends on the relationship between present and prior consumption (holding constant patient factors), which we can write as[47]

$$Q_{it} = \pi_1 Q_{i,t-1} + \pi_2 X_i + \mu_{it} \tag{6}$$

Substituting for $Q_{i,t-1}$ in Equation 5 (and dropping X_i for simplicity, since in any case we assume observables are balanced across enrollment months) yields

$$Y_{it} = \left(\beta_1 + \frac{\beta_2}{\pi_1}\right) Q_{it} + \left(\nu_{it} - \frac{\beta_2}{\pi_1}\mu_{it}\right) \tag{7}$$

In other words, our estimate of the effect of Q_{it} on mortality, $\tilde{\beta}_1 = (\beta_1 + \frac{\beta_2}{\pi_1})$, will be biased; and because of π_1, the bias can be quite large, even if the effect of prior consumption on mortality β_2 is small. The reason for this is that $0 < \pi_1 < 1$, due to inter-temporal substitution across periods (Aron-Dine et al., 2015). Past consumption $Q_{i,t-1}$ increases cumulative spending, causing some beneficiaries to enter the donut hole, increasing prices P_{it} and reducing Q_{it}. Even those beneficiaries who do not enter the donut hole are known reduce their consumption over time in anticipation. To summarize, the mis-attribution of mortality effect β_2 from $Q_{i,t-1}$ to Q_{it} is exaggerated by inter-temporal substitution between $Q_{i,t-1}$ and Q_{it}. To give a rough sense of the magnitude, running Equation 7 in our sample, using November and December as $(t-1)$ and t, yields an estimate of

[47] We emphasize that there are many possible mechanisms by which Q_{it} and $Q_{i,t-1}$ might be causally related: via the direct price effect, forward looking behavior, irregular timing of prescription fills, etc. This formulation is an empirical one and agnostic to the mechanism, but does serve to illustrate the consequences of the correlation for our analytical strategy in as transparent a way as possible.

$\pi_1 = 0.164$ (SE: 0.0024) pooled across all drugs, suggesting at least that the degree of bias could be large. While it would be interesting to perform a more formal exercise to bound the bias in $\tilde{\beta}_1$, our identification strategy becomes more tenuous the further we go back in time, as we become unable to observe later enrollment month cohorts, meaning any such exercise could not be comprehensive.

A final note is that we have used a highly stylized model of the relationship between $Q_{i,t-1}$ and Q_{it} to build intuition. Clearly, more sophisticated models are possible: for example, while Equation 7 represents the population average, the degree of mis-attribution is correlated with our instrument (earlier enrollees substitute more than later enrollees because they fall into the donut hole, meaning the bias would in fact increase proportional to $(\pi_1 \times Z_i)$. But we emphasize that the presence of bias is independent of any model specification: it results from (i) the health effects of prior drug consumption, and (ii) inter-temporal substitution, both of which are well-documented and uncontroversial in the medical and economics literature.

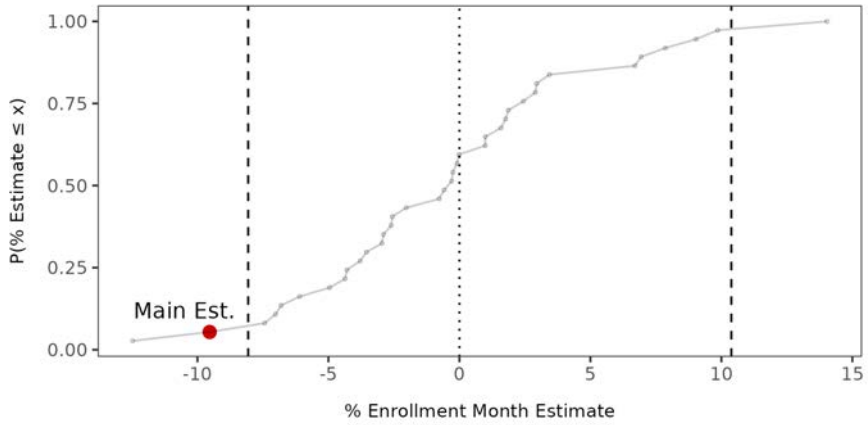
Appendix C Falsification Tests, Robustness Checks, and Power Calculation

We group our falsification checks into three categories. First, we extend the intuition of Figure IV and follow our main sample for four years after the first enrollment year, after the cost-sharing–enrollment month link is broken. Specifically, we estimate equation 4 replacing December year 1 mortality on the left hand side 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, and those who are right-censored (e.g. those enrolling in 2012 are excluded from the year 3 regression because we do not observe their mortality). 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), providing a stringent falsification test. These estimates are shown in Appendix Figure C.1, Panel A.

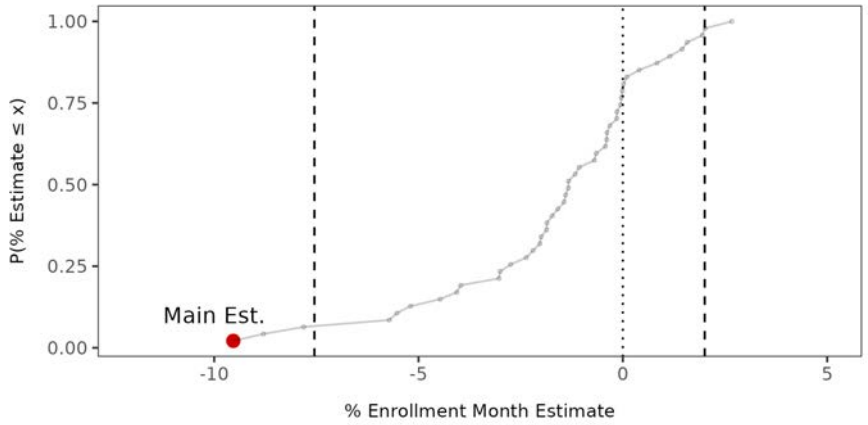
Second, we focus exclusively on the population of dual-eligibles, who face little-to-no cost sharing for prescription drugs, so there are no differences in cost-sharing by birth month. As a result, they provide a very useful comparison sample. Here, we pool all ages (66-85) together, but split the sample 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 C.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 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 focus on middle-spenders, 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 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 C.1, Panel C.

Panel A: Main Sample, January Year 2–December Year 5



Panel B: Dual-Eligibles, by Geography & Demography



Panel C: Older Non-Dual & Dual, Younger Disabled

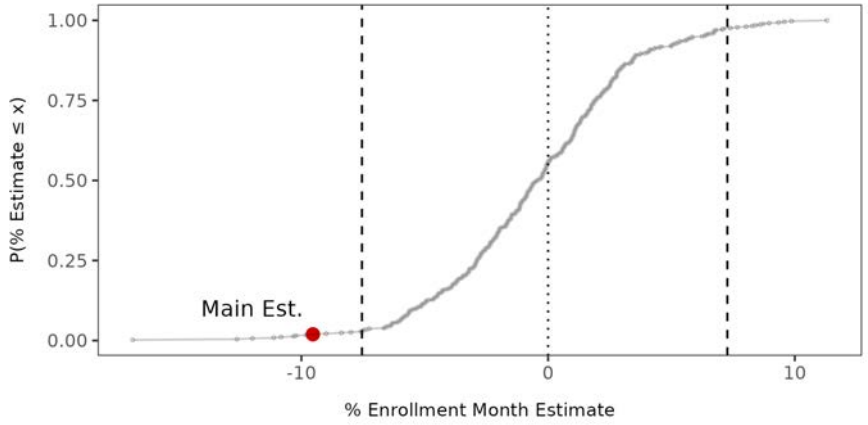


FIGURE C.1
Detail: Distribution of 'Placebo' Estimates of Enrollment or Birth Month on Mortality

Notes: Regression estimates of the effect of enrollment month or birth month on monthly mortality, for middle-spenders (initial spending in the 71-97th percentiles, measured in the first 90 days of enrollment). Estimates are divided by mean mortality in each sample/month to get a percentage change. Panel A: 36 enrollment month effects on monthly mortality in our main sample from January of year 2 to December of year 4. Panel B: 46 birth month effects on December 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 dual-disabled enrollees from age 50-64. For each age/sample group there are 9 estimates reported corresponding to monthly mortality in April-December. In Panels B and C, initial pending is calculated using claims in the fist 3 months of the year (January-March). Vertical lines show the 2.5 and 97.5 percentiles. The main (non-placebo) estimate is shown as a red dot.

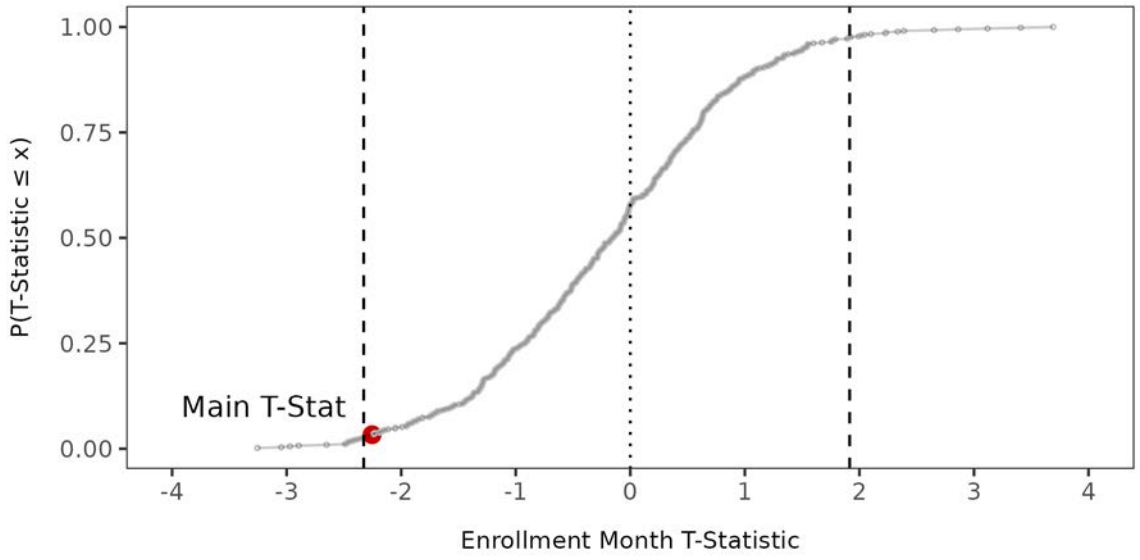


FIGURE C.2
Distribution of 'Placebo' T-statistics of Enrollment or Birth Month on Mortality

Notes: Estimated t-statistics of the effect of enrollment month or birth month on monthly mortality, for middle-spenders (initial spending in the 71-97th percentiles, measured in the first 90 days of enrollment). The same samples that comprise Figure V (and C.1) are used.

TABLE C.1
ALTERNATIVE MIDDLE-SPENDER DEFINITIONS

		(1)
Bottom Percentile Cutoff	Top Percentile Cutoff	Enrollment Month Estimate (p.p.)
Panel A: Main Specification		
70	97	-0.0112**
		(0.00498)
Panel B: Alternative Specifications		
70	96	-0.0114**
		(0.00509)
70	95	-0.0103**
		(0.00518)
65	97	-0.0115**
		(0.00457)
65	96	-0.0116**
		(0.00465)
65	95	-0.0108**
		(0.00472)
60	97	-0.0117***
		(0.00411)
60	96	-0.0118***
		(0.00417)
60	95	-0.0111***
		(0.00422)

Notes: Column (1) presents the estimated coefficient on enrollment month for the middle-spending bin from equation (4). Panel A mirrors the estimate in Panel A of Table III, where the middle-spenders bin is defined from the 71-97th within-enrollment month percentiles of initial 90-day spending. Panel B presents estimates from alternative definitions of the bottom and top within-enrollment month percentile cutoffs used to define the middle-spenders bin.

TABLE C.2
BIRTH MONTH ESTIMATES

	(1)	(2)
	Dec. Mortality Rate (p.p.)	Birth Month Estimate (p.p./mo)
Panel A: Entire sample (N = 358,706)		
Low-spenders	0.053	-0.000476
		(0.00188)
Middle-spenders	0.118	-0.00913*
		(0.00466)
High-spenders	0.288	0.053***
		(0.0185)
Panel B: Birth month enrollee sample (N = 274,102)		
Low-spenders	0.047	-0.000141
		(0.00228)
Middle-spenders	0.124	-0.013**
		(0.00591)
High-spenders	0.292	0.0563**
		(0.0229)

Panel C: Mortality by birth month (middle-spenders, birth month enrollee sample)

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: In Panel A, present estimates of equation (4) where enrollment month is replace with birth month. In Panel B we present the same but the sample is restricted to those enrolling "on time" in their birth month. In Panel C, we report the mortality rates by birth month for those the middle-spenders in the birth month enrollee sample.

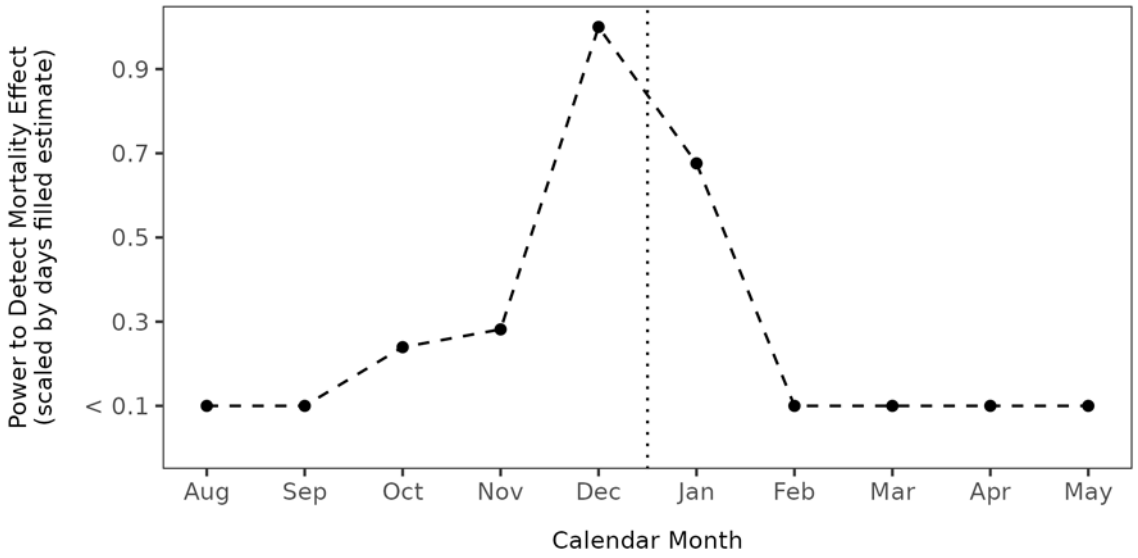


FIGURE C.3
Power to Detect Mortality Effect by Month

Notes: Each point is the estimated power to detect a statistically significant coefficient of enrollment month on mortality in a given month of calendar year 1 (August-December) and calendar year 2 (January-May). We normalize power in December to 1, and set the target effect size to the estimated effect of enrollment month on mortality in December (-0.0112 p.p., see Table III). For each month other than December, power is calculated following Dupont and Plummer (1998), as a function of changes in (i) effect size and (ii) sample size. To determine effect size in other months, we scale the December effect by the relative effect size of enrollment month on consumption (from Equation 2) in that month relative to the December's (shown in Figure IV panel B). For example, the estimated coefficient α_1 (drug-days filled on enrollment month) in November is 1.77 days filled/enrollment month, or 48% of the December coefficient (3.69). We therefore set the effect size of November as -0.0054 p.p., 48% of the effect size in December (-0.0112 p.p.). Power in each month prior to December also reflects the observed sample size, which reflects the loss of an additional month of enrollees. For example, the November analysis must drop September enrollees, whom we have not yet observed for three months to define spending bins, so the November sample size of middle spenders is set at 83,672 (86.4% of the December sample of 96,849). Note that here we assume only same period effects of utilization on mortality, and is therefore a lower-bound on power.

Appendix D Risk Prediction and Drug Consumption

For each of three drug categories M, we define a set of adverse events the drug category is prescribed to prevent: cardiovascular, diabetes, and pulmonary. We compile a list of observable adverse outcomes in each category by inspecting the patient's longitudinal claims record for the presence of ICD codes indicating an adverse event: 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. For each of the three categories, let Y_M be an indicator indexing whether a beneficiary experienced an adverse event in that category. We identify medications belonging to a category using ATC3 codes, and define T_M as an indicator for whether the beneficiary was prescribed any drug in that class.

In a sample of dual-eligible enrollees, where each observation is a beneficiary-year. We set $Y_M = 1$ if we observe an adverse event in category M (e.g., heart attack or stroke) over April to December of a given year. We set $T_M = 1$ if we observe a claim for a drug in category M over January-March, and form a set of predictors including race, sex, state and drug filling behavior (number of claims, total spending, ATC4 indicators) over the same period. In total this yields 1,770 features. We then define separate training samples for each drug–event pair, restricting to those for whom $T_M = 0$ (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 in the untreated, i.e., in potential outcomes notation, we wish to estimate $Pr(Y_M^0 = 1|X)$, not $Pr(Y_M^1 = 1|X)$. Naturally this choice of prediction target also induces selection bias: we form predictions on Y_M^0 in patients selected into treatment status $T_M = 0$, but then wish to generate predictions on in patients with arbitrary treatment status. In particular, this means our predictions are likely to underestimate risk on average, because doctors select patients into treatment $T_M = 1$ on the basis of higher risk. We do verify that, as a check of face validity, risk increases in predicted risk for both groups (Figure D.1). This is similar in spirit to Mullainathan and Obermeyer (2019), who predict the yield of testing for heart attack in the tested, and apply the model to generate predictions in the untested.

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 and gradient boosted models and 10% is used to create an ensemble, via no-intercept OLS, that predicts Y_M using both models' predictions. We then validate the model using the final 20% held-out sample. The model follows Mullainathan and Obermeyer (2019) closely. Finally, we apply this model to generate predictions in the main sample of 65 year-old beneficiaries (using the same predictors, similary measured in the first 90 days of enrollment). We make separate predictions for each outcome Y_M (but use the same predictors for each).

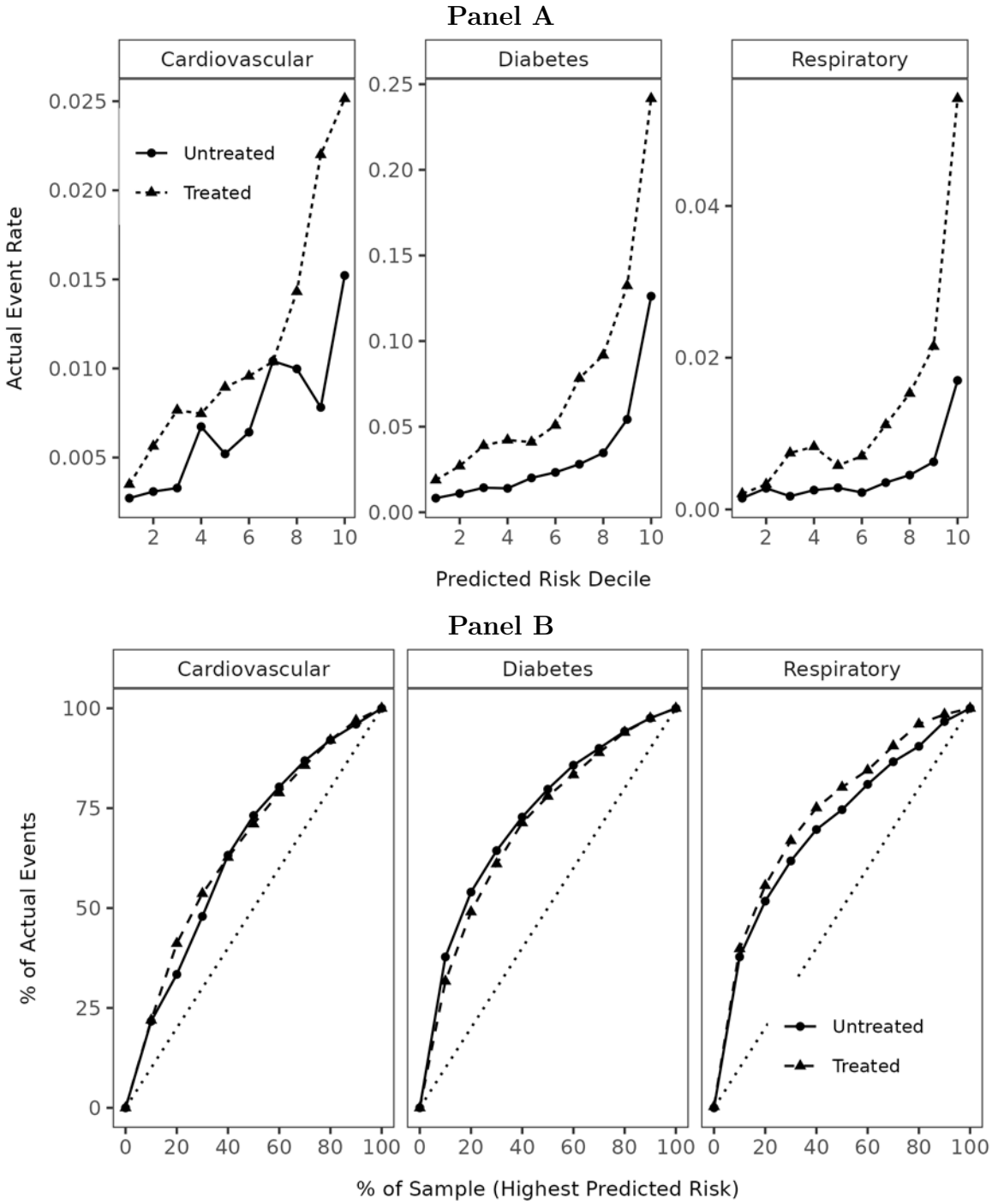


FIGURE D.1
Acute Event Risk Prediction Calibration and Capture

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.

TABLE D.1
CLASS-SPECIFIC CUTBACKS

		(1)	(2)	(3)	(4)
ATC 3 Code	ATC3 Label	% with Initial Claim	Mean Dec. Days Filled	Pre-Donut Budget Est. (Days/$100)	(%)
C10A	Statins	59.2	18.6	0.662*** (0.0762)	3.6
C07A	Beta-Blockers	31.4	9.6	0.166*** (0.0546)	1.7
C09A	ACE Inhibitors	30	9	0.17*** (0.0529)	1.9
C03A	Thiazides	26.7	7.9	0.157*** (0.05)	2.0
N06A	Antidepressants	24.8	7.9	0.273*** (0.0531)	3.5
A02B	GERD	24.3	6.5	0.435*** (0.0459)	6.7
D07A	Corticosteroids (Top.)	22.4	3.3	0.255*** (0.0304)	7.7
A10B	Diabetes (Oral)	20.9	8.7	0.447*** (0.0635)	5.2
R01A	Decongestants	20	3.3	0.261*** (0.031)	8.0
C09C	ARBs	19.6	5.5	0.267*** (0.0414)	4.9
R03B	Inhalants	17.9	3.4	0.296*** (0.0328)	8.7
S01A	Antiinfectives	17.4	0.9	0.0118 (0.0108)	1.3
C08C	CCBs	16.1	4.9	0.169*** (0.0396)	3.5
N02B	Other Analgesics	15.4	1.6	0.0047 (0.0167)	0.3
H03A	Thyroid	15	4.8	0.111*** (0.0409)	2.3
M01A	Antiinflamatory	14.4	2.7	0.101*** (0.0278)	3.7
C05A	Hemorrhoidal	13.4	1.9	0.0404* (0.0224)	2.1

ATC 3 Code	ATC3 Label	% with Initial Claim	Mean Dec. Days Filled	Pre-Donut Budget Est. (Days/$100)	% Est.
H02A	Corticosteroids (Oral)	12.5	1.1	0.0492*** (0.0162)	4.6
R05D	Cough suppressants	11	1.1	0.00399 (0.0133)	0.4
S01B	Anti-inflamatory (Oc.)	11	1.1	0.0316* (0.0162)	2.8
A01A	Dental Caries	10.6	0.9	0.0548*** (0.015)	5.9
N02A	Opioids	10.3	1.4	0.0105 (0.0166)	0.8
A07E	Anti-inflamatory (Intest.)	10.1	1.3	0.104*** (0.0195)	8.2
D07X	Corticosteroids (Top., Comb.)	9.2	0.8	0.0297** (0.0133)	3.5
J01M	Quinolones	9.2	0.3	0.0106* (0.00551)	3.2

Notes: This table presents average utilization and cutbacks by specific drug class (ATC3) for middle-spenders. Column (1) shows the percent of enrollees with a claim in the class in the first 90 days. Column (2) shows the average number of days filled in the class in December. Column (3) presents estimates of γ_1 from equation 4, where the dependent variable is the number of days filled in the class in December of year 1. Column (4) presents the percentage reduced form estimate (column (3) divided by column (2)).

TABLE D.2
UTILIZATION RESPONSE BY RISK AND INCOME

	(1)	(2)	(3)	(4)	(5)	(6)
	All		*Bottom 2/3 Income*		*Top 1/3 Income*	
	Mean	Est. (S.E.)	Mean	Est. (S.E.)	Mean	Est. (S.E.)
All Classes	126.4	4.93*** (0.289)	129.2	4.8*** (0.36)	121.3	5.2*** (0.493)
Cardiovascular	50.2	1.42*** (0.157)	51.0	1.4*** (0.193)	48.8	1.43*** (0.275)
Diabetes	10.2	0.618*** (0.0701)	11.0	0.623*** (0.0875)	8.6	0.567*** (0.118)
Respiratory	5.3	0.459*** (0.0453)	5.4	0.482*** (0.0564)	5.0	0.433*** (0.0776)

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

Notes: Column (1) presents the mean number of December days filled for all drugs and by broad class, for those in the 71-97th percentiles of initial spending. Column (2) presents 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 chronic pulmonary disease. 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
DROPPING ALL DRUGS

	(1)	(2)	(3)
	No Dec. Fills (%)	Pre-Donut Budget Est. (% Fill/ $100)	Est. (%)
Panel A: All middle-spenders			
All middle-spenders	16.6	-0.536*** (0.0877)	-3.23
Panel B: By initial fills (middle-spenders)			
Lowest third	32.1	-0.148 (0.22)	-0.46
Middle third	16.1	-0.737*** (0.137)	-4.58
Top third	6.9	-0.786*** (0.0993)	-11.47

Notes: Panel A shows an estimate of γ_1 (the effect of pre-donut budget changes) from equation 4, where the dependent variable is an indicator for filling zero drug prescriptions in December (mean and estimate are shown in percentage points). Panel B present similar estimates where pre-donut budget, in addition to being interacted with spending group, is also interacted with within-spending group tercile of number of fills in the first 90 days.

Appendix E Comparisons to Existing Literature

TABLE E.1
DRUG DEMAND RESPONSE: COMPARISON TO RECENT STUDIES

	(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 that uses the December price in dollars (either overall or class specific) per fill as the endogenous variable. (This assumes that the December spot price is the only driver of filling, which we know is not the case. So all the caveats—that enrollment month can affect utilization via many mechanisms, of which this is just one—still apply. This is likely why our estimates, while the same order of magnitude as other studies, are biased upwards because of correlations between P_t and other periods before and after.) 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).

TABLE E.2
HEALTH INSURANCE AND MORTALITY: COMPARISON TO RECENT STUDIES

Study & Setting	Age Mean	Mortality p.p./yr	Effect p.p./yr	Change %	Intervention (Change %)
Abaluck et al. (2021)					Full vs. partial donut
Medicare Advantage	77.5	4.70	-0.460	-9.8	hole coverage (IV)
Goldin et al. (2020)					Obamacare
All taxpayers	38.6	1.01	-0.063	-6.3	coverage (+4.6%)
Miller et al. (2021)					Medicaid
Medicaid, 55-64yo	59.3	1.40	-0.132	-9.4	coverage (+13.5%)
Present study, Part D					
Middle-spenders	65.0	1.43	-0.197	-13.8	Per $100/mo. pre-donut
All	65.0	0.94	-0.053	-5.68	hole drug budget (+24%)

Notes: Summary of effect sizes, both absolute and in percentage terms, from three recent studies of the effect of health insurance on mortality. Abaluck et al. (2021)'s design compares mortality in beneficiaries who switch Medicare Advantage plans, which have a number of correlated characteristics, so isolating the specific effect of donut hole coverage is challenging. Nevertheless, in a 2SLS design with donut hole coverage as the endogenous variable, mortality was reduced by 0.46 p.p. per year vs. a base rate of 4.70 p.p., or 9.8%. (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) who study partial vs. no donut hole coverage). Goldin et al. (2020) take advantage of randomly-assigned letters advising taxpayers of the penalty for lacking health insurance coverage under the Affordable Care Act, and find that this reduced mortality by 6.3% (reduced form); they estimate that the intervention concurrently increased the number of months of insurance coverage by 4.6% over the study period. Miller et al. (2021)'s differences-in-differences study of Affordable Care Act Medicaid expansion finds a 9.4% reduction in mortality (reduced form); a survey-based measure of uninsurance fell by 13.5% as a result of expansion. For comparison, we present estimates from the present study in the table. The first row presents estimates from middle-spenders, which translate the results in Table III into annualized numbers for comparability with the studies above. Note that the mortality figure differs slightly from the figure in Table I: this table presents annualized December mortality rather than mortality calculated over the entire year. The second row dilutes the middle-spender effect by the share of middle-spenders in the population (0.27), again for comparability to the studies above.