Title
Essays in Social Insurance

Author
Berman, Jacob

Publication Date
2021

UNIVERSITY OF CALIFORNIA

Los Angeles

Essays in Social Insurance

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Economics

by

Jacob Berman

2021

ABSTRACT OF THE DISSERTATION

Essays in Social Insurance

by

Jacob Berman
Doctor of Philosophy in Economics
University of California, Los Angeles, 2021
Professor Adriana Lleras-Muney, Chair

My first chapter explores the effect of income on health in the context of Social Security and Medicare. Income is a powerful predictor of health among the elderly, but existing research has struggled to identify a causal link. I estimate the causal effect of Social Security income on health care utilization and health outcomes among elderly men. Using Medicare administrative records and a regression discontinuity design, I exploit several changes in the Social Security benefit formula that vary abruptly by date of birth. This feature has been overlooked by prior research and causes workers born one day apart to receive positive and negative income shocks. Over my sample period, I estimate a \$100 increase in monthly Social Security benefits leads to a \$38 decline in monthly federal Medicare expenditures. To provide evidence the decline in spending is driven by improvements in health, I show income leads to reductions in diagnoses for chronic conditions and mortality. My results suggest cuts to Social Security benefits may have unintended social and fiscal costs. Overall, these findings highlight the importance of examining health outcomes when evaluating the costs and benefits of social insurance programs.

My second chapter examines the Qualified Medicare Beneficiary (QMB) program—a

means-tested benefit that exempts low-income Medicare beneficiaries from their cost-sharing obligations. Patients enrolled in the QMB program face zero prices for Medicare services, receive monthly premium exemptions, and subsidized Part D prescription drug coverage. However, because providers face administrative burdens and lower reimbursement rates, they may limit access for QMB patients. Given these offsetting features, it is unclear whether patients benefit from enrolling in the program. To examine the effect of QMB enrollment on patient access and outcomes, I study an expansion of QMB eligibility that occurred in Connecticut in 2009. By doubling the income limit and removing the asset test, the state more than doubled enrollment. Using a difference-in-difference design and Medicare administrative records, I estimate the effect of QMB enrollment on health care utilization and health outcomes. I find that the program appears to be successful at reducing beneficiary costs without limiting access to care. I show QMB enrollment leads patients to save $1,300 in annual outpatient cost-sharing liability without any change in the utilization of outpatient services. Overall, my results suggest that within the universe of Medicare, the program transfers a substantial share of producer surplus to consumers.

My third chapter studies how cuts to Social Security retirement benefits affect Social Security disability enrollment. Specifically, I exploit a policy change which caused an abrupt decline in the generosity of retirement benefits for workers near certain date of birth cutoffs. Using various regression discontinuity designs, I do not find evidence that cohorts affected by the policy change have higher rates of benefit receipt. My preferred specification estimates a precise null effect. My confidence interval implies the increase in the full retirement age caused disability enrollment to only change between -0.25 and 0.28 percentage points. Relative to existing work, my paper provides evidence that disability enrollment is less sensitive to policy changes. I find there are fewer marginal SSDI applicants implying the program's moral hazard effects are more modest than prior work would suggest.

The dissertation of Jacob Berman is approved.

Kathleen McGarry

Martin Hackmann

Laura Wherry

Adriana Lleras-Muney, Committee Chair

University of California, Los Angeles

2021

For Arian. Over the last six years, you were there to celebrate with every success and commiserate with every setback. I feel so lucky to have shared this journey with you.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I am grateful to Leslie McGranahan for first encouraging me to apply to graduate school. As a research assistant, she gave me tremendous freedom to pursue my own projects and helped spark my passion for economics research.

For the dissertation, I am most indebted to Adriana Lleras-Muney. She remained committed to my success over the many years it took to assemble the funding and data for my papers. It is hard to imagine what my research would have looked like without her feedback and guidance. Additionally, Kathleen McGarry, Martin Hackmann, and Laura Wherry provided outstanding advising. They consistently offered excellent comments and suggestions.

I also thank Marianne Bitler, Moshe Buchinsky, Ashvin Gandhi, Alex Gelber, Tal Gross, Jonathan Skinner, Meghan Skira, Gonzalo Vazquez-Bare, and Till von Wachter for their feedback and suggestions. My projects benefited from comments by participants at the UCLA Applied Microeconomics Proseminar, the CCPR Student Proseminar, ASHEcon, and the UC Davis Center for Poverty Research. I also acknowledge Dr. Ioana Popescu and Matthew Lahmann who provided invaluable assistance acquiring and managing the data.

This work was supported in part by the CCPR Population Research Infrastructure Grant from NICHD: P2C-HD041022, the Alfred P. Sloan Foundation Pre-Doctoral Fellowship on the Economics of an Aging Workforce awarded from the NBER, the Lewis L. Clarke Graduate Fellowship from UCLA, and a Dissertation Fellowship from the Center for Retirement Research at Boston College.

VITA

2009–2013	B.A. Economics, University of Chicago
2013–2015	Associate Economist, Federal Reserve Bank of Chicago
2017	Summer Research Associate, Congressional Budget Office
2018–2020	Pre-Doctoral Fellowship on the Economics of an Aging Workforce, National Bureau of Economic Research
2020–2021	Center for Retirement Research Dissertation Fellowship, Boston College

CHAPTER 1

Can Income Buy Health? Evidence from Social Security Benefit Discontinuities and Medicare Claims

1.1 Introduction

Although there is an intuitive connection between income and health, existing research has struggled to identify a causal link. Factors that influence income—like education or family background—also influence health, and disentangling the direct causal effect of income on health is challenging. Quasi-experimental research designs can solve these endogeneity problems, but results from these settings may not generalize to broader populations. For example, comparing health outcomes of lottery winners and losers can estimate a causal effect, but the relevance of this effect for policy design is unclear (Cesarini et al., 2016).

In the United States, a causal link from income to health has direct policy implications for older adults. This is because they receive the majority of their income and health benefits from Social Security and Medicare. A causal relationship would imply that changes in Social Security benefits could affect Medicare spending, but whether such a relationship exists is unknown. Federal spending for these two programs will soon exceed 10% of GDP, so even modest effects could have major budget implications.

In this paper, I estimate the effect of income on health outcomes in a setting which simultaneously features policy-relevant variation, a large population in the United States, and an opportunity for causal identification. Specifically, I exploit a previously unstudied anomaly in the Social Security formula that causes the generosity of monthly pension benefits to vary abruptly by exact date of birth. As a consequence of this policy, workers born one day apart receive different income. Assuming workers with birth dates near these cutoffs are similar along other dimensions, differences in Social Security income are as good as random. Because Social Security and Medicare cover nearly all adults over 65, the two programs provide an ideal setting to examine the causal effect of income on Medicare-related health outcomes.

I start by showing that the generosity of monthly Social Security income changes abruptly for workers born after January 2 among all recent birth cohorts. In my sample, the amount ranges from -2% to 4.5% depending on the particular cohort. Unlike Social Security changes studied in prior research (e.g. the 1917 Notch), workers are unlikely to know these changes exist. The calculations to identify the changes are complex, and in fact, this paper is the first to describe them explicitly. In my context, the lack of salience is useful because it minimizes the extent to which labor income offsets the change in Social Security income. Additionally, Social Security reforms have proposed benefit changes of a similar magnitude, so this setting has direct policy relevance.[1]

The primary outcomes are health care spending, diagnoses for chronic conditions, and mortality. I measure these using a 100% extract of restricted access administrative records from the full Medicare population from 2006 to 2011. Because the files cover essential health services across inpatient and outpatient settings (Medicare Part A and B), I can measure substitution effects between different types of care. The files also provide detailed hospital procedure and diagnosis codes which are useful for constructing measures of quality.

My estimation sample consists of men with an average age of 75. Thus, I observe them 13 years after the income shock becomes binding at age 62. While previous work has focused on the short-term effects of liquidity shocks (Evans and Moore, 2012) or labor force exit (Fitzpatrick and Moore, 2018), the panel structure of my data enable me to study outcomes

[1] For example, see Fiscal Year 2014 Budget Request (Office of Management and Budget) or Options for Reducing the Deficit: 2019 to 2028 (Congressional Budget Office).

long after the shock occurs. This helps detect effects that a short-term analysis would otherwise miss. I focus on men because the date-of-birth cutoff is linked to wage earners and men are the primary wage earners for these cohorts.

My empirical strategy relies on a regression discontinuity design with date-of-birth as the assignment variable. There are ten discontinuities corresponding to the ten birth cohorts in my sample. Compared to a single discontinuity, working with multiple discontinuities provides a valuable opportunity for specification and falsification tests. I test if effects are symmetric between positive and negative shocks, and that effects are null for cohorts with no shock. Because some of the discontinuities are modest, I consider several techniques for stacking the discontinuities to maximize statistical efficiency.

Neither theory nor empirical research provide clear predictions of how income will affect health and Medicare spending. To provide a framework for understanding the results, it is useful to distinguish between three distinct hypotheses. First, we might expect income to increase health status and decrease spending. Income may improve underlying health by leading to lower stress, a healthier lifestyle, or new investments in non-Medicare health services. To the extent that beneficiaries are healthier, they will require fewer Medicare services and spending will decline. Evidence for this pathway comes from research finding that Social Security payments lead to increased prescription drug use (Gross, Layton and Prinz, 2020) which can reduce costly hospitalizations (Chandra, Gruber and McKnight, 2010). Similarly, there is evidence that Social Security income reduces mortality among disability beneficiaries (Gelber, Moore and Strand, 2018).

Second, we might expect income to increase both health status and health care spending. Higher income may diminish sensitivity to cost-sharing; this would increase Medicare spending as wealthier beneficiaries demand more services. If the marginal dollar of health spending is effective at improving underlying health, then health status should increase as well. This pathway is consistent with structural models of health investment (Grossman, 2000; Hall and Jones, 2007). In these models, health care is an input to a health production function and the income elasticity of demand for health is positive.

Third, we might expect income to worsen health outcomes and increase health care spending. If the income elasticity of goods such as tobacco or alcohol is positive, then an income shock may lead unhealthy consumption to increase. To the extent that beneficiaries are in worse health, they will require more Medicare services and spending will increase. Several papers find evidence in this direction, at least in the short-term. For example, Evans and Moore (2012) find liquidity shocks cause greater drug and alcohol mortality; Gross and Tobacman (2014) find a similar pattern for hospitalizations.

My results are most consistent with the first hypothesis. I find that increases in Social Security income reduce health care spending and mortality. Specifically, I estimate a 1% increase in Social Security income causes a 0.93% decline in payments for Medicare covered services, and a 0.98% decline in mortality over a 6 year period. I also find evidence of declines in diagnoses for chronic conditions. In terms of fiscal spillovers, my findings suggest that 38% of the per-capita cost of increased Social Security benefits would be offset by lower Medicare spending. Because there is a positive externality, retirement insurance models that ignore spillover effects onto Medicare will underestimate the optimal Social Security benefit level.

This paper contributes to existing research on how income affects the health of the elderly by using population-level administrative data and a new source of causal identification. The most closely related research studies "the Notch," a policy change which abruptly cut Social Security income for retirees born after January 2, 1917. Researchers have used the Notch to study outcomes such as prescription drug utilization (Moran and Simon, 2006), mortality (Snyder and Evans, 2006), long-term care (Goda, Golberstein and Grabowski, 2011), and earnings (Gelber, Isen and Song, 2016).[2] The Notch cohorts are generally too old

[2]Other papers consider outcomes such as household structure (Engelhardt, Gruber and Perry, 2005), weight (Cawley, Moran and Simon, 2010), home health (Tsai, 2015), mental health (Golberstein, 2015), and cognitive function (Ayyagari and Frisvold, 2016). Because Gelber, Isen and Song (2016) uses SSA data to study labor force outcomes rather than Medicare data to study health outcomes, they observe exact date of birth.

to be observed in Medicare research files, so these papers are forced to measure health and health care utilization in survey data. As Handwerker (2011) argues in detail, this creates a challenge for identification. Because survey data reports birth dates (at best) at the quarterly level, variation between cohorts can overwhelm the variation in benefit amounts. This implies the quarter-of-birth instrument derived from survey data may not satisfy the necessary exclusion restrictions. In contrast, I consider cohorts born in the 1920s and 1930s for whom administrative data with exact date of birth is available. This allows me to focus on a narrow, 30-day bandwidth around the cutoff and minimize any bias arising from seasonal trends (Buckles and Hungerman, 2013). Additionally, my large sample provides statistical power to identify changes in health outcomes that would be undetectable in other data.

Outside of the Social Security context, these results build on a large literature studying the impacts of income on health and health care utilization (Grossman, 2000; Cutler, Deaton and Lleras-Muney, 2006; Hall and Jones, 2007; Chetty et al., 2016). Applied research on older adults has focused on wealth shocks due to changes in equity markets (Schwandt, 2018; McInerney, Mellor and Nicholas, 2013), inheritance shocks (Meer, Miller and Rosen, 2003; Kim and Ruhm, 2012; Van Kippersluis and Galama, 2014) or lottery winnings (Lindahl, 2005; Gardner and Oswald, 2007; Apouey and Clark, 2015; Cesarini et al., 2016). These studies show the effect of income on health can vary from positive to zero to negative depending on the population, statistical methods, and institutional details. The range of estimates makes it difficult to apply existing results to a policy context. Conversely, a valuable feature of my setting is that the size of the income-variation and the population affected are directly relevant for discussions of Social Security reform.[3]

This paper also contributes to literature studying spillovers between social insurance programs by providing the first empirical evidence of long-term spillovers from Social Security

[3]There are two exceptions to highlight. First, Gelber, Moore and Strand (2018), who use a regression kink design to show higher income reduces mortality among Social Security disability beneficiaries. Second, Fitzpatrick and Moore (2018), who show that eligibility for retirement benefits at age 62 leads to a large spike in male mortality. Their result matches other work finding liquidity shocks have negative short run-effects (Dobkin and Puller, 2007; Evans and Moore, 2012; Gross and Tobacman, 2014).

to Medicare.[4] Prior work focusing on short-term liquidity effects finds that some beneficiaries postpone filling Medicare drug prescriptions until they receive their monthly Social Security payment (Gross, Layton and Prinz, 2020). Other work on Social Security spillovers has studied the effect of changing the retirement age on disability applications (Duggan, Singleton and Song, 2007; Li and Maestas, 2008), substitution between disability insurance and unemployment insurance (Lindner, 2016; Mueller, Rothstein and von Wachter, 2016), disability and welfare assistance (Borghans, Gielen and Luttmer, 2014), or disability and Medicaid (Burns and Dague, 2017). Medicare spillovers have mostly been studied in the context of Medicaid. For example, Grabowski (2007) examines conflicting incentives for long-term care, and Carey, Miller and Wherry (2018) explore whether Medicaid expansions lead to changes in access for Medicare beneficiaries. Social Security and Medicare are unique because they account for over a third of federal spending and cover nearly 70 million people. Given their size, understanding fiscal externalities between the two programs is necessary for evaluating how policy changes can affect the long-term budget outlook.

The remainder of the paper is structured as follows. Section 2 describes my source of income variation in the context of Social Security and Medicare program rules. Section 3 describes the Medicare administrative data, and section 4 discusses my identification strategy. Section 5 provides aggregate results for spending, chronic conditions, and mortality. To explore the underlying mechanisms, Section 6 discusses treatment effect heterogeneity. Section 7 describes a back-of-the-envelope calculation to measure fiscal externalities and Section 8 concludes.

[4]See Philipson and Becker (1998) for a discussion of Social Security and Medicare in the context of optimal retirement benefits, and Zhao (2014) for a application of spillovers with an overlapping generation general equilibrium model.

1.2 Institutional Setting

Nearly all elderly Americans receive retirement income from Social Security and health insurance from Medicare.[5] The generosity of Social Security income varies abruptly by exact date of birth, but Medicare benefits do not. This provides a setting to examine how quasi-random income variation affects health care utilization and health outcomes for a large population.

1.2.1 Social Security Income

Social Security provides monthly retirement benefits to qualified retirees and their spouses. As a social insurance program, it insures the labor market income of workers against old-age risks. Workers pay into the program through mandatory payroll deductions and receive benefits as a function of their contributions. The benefit formula is progressive. Higher income workers receive higher benefits, but the marginal replacement rate declines with income. Workers can start claiming at age 62, or they can receive bonus benefits by delaying up to age 70.

Because workers make contributions over several decades, implementing the formula requires adjusting wages and benefits for inflation. The modern benefit formula indexes wages and prices separately. Nominal wage histories before entitlement are adjusted using the Average Wage Index (AWI), a time series the Internal Revenue Service computes using administrative records. Nominal benefit amounts after entitlement are adjusted using the Consumer Price Index (CPI), a time series the Bureau of Labor Statistics computes using survey data. The CPI and AWI base years vary by a worker's date of birth. For wage indexation, the base year is the calendar year a worker turns 60. For benefit indexation, the base year is the calendar year a worker turns 62. These features interact in a way that

[5] Workers and their spouse are eligible for both programs if the worker has at least 10 years of creditable labor market earnings. 97% of adults over age 65 meet this threshold. Infrequent workers with disabilities, late-arriving immigrants, and certain government employees account for the remaining 3% (SSA, 2015).

means two workers with identical nominal earnings histories will receive different benefits depending on their date of birth.

Consider a worker born in January of year b compared to an identical worker born a month before in December of $b-1$. The worker with a January birthday has a base year for wage indexation of $b+60$, and *does not* receive a CPI adjustment during his first year. The worker with a December birthday has a base year for wage indexation of $b+59$, but *does* receive a CPI adjustment during his first year. The percent change in benefits for January birthdays is the difference between the percentage growth in AWI at age 60 minus the growth in the CPI at age 61

$$\%\Delta Benefits \approx \%\Delta AWI_{b+60} - \%\Delta CPI_{b+61} \tag{1.1}$$

Appendix Section 3 derives this equation from the benefit formula. There are four features of the discontinuities to highlight. First, they affect every cohort born after 1917. Appendix Figure A.1 shows this includes future retirees who are not yet entitled. The relevant discontinuity occurs at January 2 instead of January 1 because under Social Security regulations an individual attains a particular age on the day preceding the anniversary of their birth.[6] Second, they are the same in percentage terms regardless of income level or claiming age. Consider a pair of workers with identical nominal earnings histories born on either side of the cutoff for a given cohort. The percentage difference comparing two low-income earners claiming at age 62, and the percentage difference comparing two high-income earners claiming at age 65 are the same.[7] This is because the wage index changes average indexed monthly earnings and the thresholds in marginal replacement rates by the same amount. Finally, recipients are unlikely to be aware of these discontinuities. The calculations involved

[6] POMS Regulation GN 00302.400

[7] Because half of the cohorts in the sample are affected by a 0.5% change in the delayed retirement credit, there are slight differences that arise when claiming after 65. Appendix Section 3 shows these changes are too small to threaten identification.

are opaque. They are not described on the Social Security website or in previous academic research. While other benefit changes like the Notch or changes in the Full Retirement Age are highly salient, these changes are effectively invisible. A beneficiary would only be aware of the shock if he calculated how his benefits would change under different hypothetical dates of birth.[8]

Table 1.1 summarizes the size of the income shock for the 10 cohorts in my sample. The benefit changes range from 4.5% to -1.9%. These are comparable in magnitude to changes studied in prior research. On the low end, Deshpande, Fadlon and Gray (2020) studies 2 month increases in the Full Retirement Age which cut benefits by 1.1%. On the high end, Gelber, Isen and Song (2016) studies the 1917 Notch which cut benefits by 7%. In my setting, the differences in the monthly dollar amount for the average worker ranges from $59 to -$24. For context, the average premium for Part D prescription drug insurance during this period was $30 per month.[9]

Most changes are positive. Equation (1) predicts this will occur if nominal wages rise faster than prices—a pattern we expect when there is long-run productivity growth. In some years (1993, for example) negative changes will occur when macroeconomic shocks cause prices to grow more quickly than wages. By chance, the wage and price parameters for the 1927/1928 and 1933/1934 cohorts nearly exactly offset each other. I consider these two cohorts unaffected and use them as placebo tests.

[8]If beneficiaries were aware of the shock, it would be "realized" when the Social Security Administration publishes the AWI and CPI computations in the Federal Register. This occurs in late October of the year a cohort turns 61.

[9]An alternate statistic for measuring the generosity of benefits is the replacement rate, the ratio of benefits at entitlement to the lifetime average of real wages. For the 1936/1937 discontinuity, the replacement rate moves from 45.3% to 47.4%. See Appendix Section 3 for additional details.

1.2.2 Health Coverage from Medicare

Social Security beneficiaries receive health insurance from Medicare. Non-disabled beneficiaries become eligible for Medicare at age 65 and those already claiming retirement benefits are automatically enrolled.[10] About 70% of individuals enroll in traditional Fee-For-Service (FFS) Medicare. FFS Medicare consists of Part A which covers inpatient hospital services, skilled nursing, hospice, and home health; Part B which covers physician, outpatient, and preventive services; and Part D which covers outpatient prescription drug benefits. Nearly all providers accept FFS Medicare patients and referrals are not required. Alternatively, roughly 30% of beneficiaries enroll in a privately run Medicare Advantage plan. Under this option, Medicare pays private insurers to provide benefits through a managed care regime. These plans offer lower out-of-pocket costs by restricting access to a narrower network of providers. In addition, they sometimes provide vision and dental benefits which traditional Medicare does not.

Although FFS Medicare has substantial cost-sharing, about 80% of FFS beneficiaries have these expenses paid by secondary insurance plans which cover some or all of the beneficiary's cost-sharing liability. The majority of secondary plans are provided by firms for their retired employees, but consumers can also purchase them directly (MedPAC, 2018). The remaining out-of-pocket health expenditures are for services not covered by Parts A and B such as prescription drug cost-sharing, nursing, long-term care, and non-medical services (Fahle, McGarry and Skinner, 2016). Low-income beneficiaries may have these services covered by Medicaid, a separate means-tested health insurance program that is administrated by state governments. Beneficiaries covered by both Medicaid and Medicare are also generally exempt from cost-sharing, so they do not have to purchase costly supplement plans.

[10] Receiving Social Security is a sufficient but not necessary condition to receive Medicare. Those who are entitled to Medicare but not Social Security OASI benefits include some Supplemental Security Income (SSI) recipients, Railroad Retirement Board beneficiaries, certain government workers, and individuals over 65 with less than 40 quarters of covered earnings.

1.3 Data

To examine the effects of the wage and price deflators on benefit levels, I use Social Security public use files. I measure health care utilization, chronic conditions, and mortality using Medicare administrative files.

1.3.1 Public Use Social Security Files

Benefit amounts are unobserved in the Medicare data, so I rely on the Public Use Benefits File to validate the actual benefits paid follow the pattern predicted by equation (1). This 1% anonymized extract from Social Security administrative records provides year of birth, sex, and benefit amount for the year 2004. I compute nominal benefit amounts over the 2006 to 2011 period by adjusting the 2004 amount using the SSA Cost-Of-Living Adjustment (COLA) time series.

The file also measures annual wages for all years between 1951 and 2003. This allows me to confirm that changes in benefit amounts are not offset by changes in labor income. Using this file, I define labor income as mean earnings between ages 62 and 66. I also consider any participation on the extensive margin by measuring the fraction with non-zero earnings. Because the treatment effect is likely to vary across the income distribution, I also use a separate dataset on mean benefit amounts by zipcode. The zipcode data is derived from administrative records covering the universe of elderly beneficiaries.[11]

1.3.2 Restricted Access Medicare Administrative Files

My primary dataset is derived from Medicare administrative files. I use the 100% full population panel from 2006 to 2011 as well as 2017.[12] Every individual enrolled in Medicare during

[11] See OASDI Beneficiaries by State and ZIP Code, 2011.

[12] Although this period includes the implementation of the Affordable Care Act, the law had only minor effects on Medicare. See CRS Report R41196.

these years is included. The dataset has four parts. First, the Master Base Summary File provides demographic data such as exact date of birth, sex, race, and most recent zipcode. There is also an entitlement code reported directly from the Social Security Administration. The code measures if someone claims benefits based on their own wage history, a spouse's wage history, as a survivor, or through another type of entitlement. On the enrollment side, I observe months of coverage in Part A, Part B, Part D, Medicare Advantage, and Medicaid.

When beneficiaries die, their exact date of death is recorded in that year's file and they are deleted from the next year's file. I assume that beneficiaries who appear in the 2011 file but not the 2017 file died during the interim.[13] Mortality rates computed using this method are nearly identical to the mortality rates in Social Security life tables and vital statistics. Cause of death is unobserved in these data.

Second, the Cost and Utilization segment provides health expenditure and utilization data. This file measures 11 categories of service (e.g. inpatient hospitalization, evaluation and management, imaging) and includes Medicare payments, cost-sharing liability, and visit counts. Each observation summarizes a beneficiary's utilization over the calendar year. Third, the Chronic Condition segment measures if a beneficiary has received treatment for any of 27 chronic conditions (e.g. hypertension, heart failure, or depression). Medicare constructs this file using a special algorithm developed by professional medical coders. The algorithm searches all recent Part A and B claim records to see if providers billed Medicare under a diagnosis code associated with a given chronic condition. They are imperfect measures of underlying health because those who fail to seek treatment are excluded.

Fourth, the MEDPAR segment provides procedure and diagnosis billing codes for inpatient claims. I use the file to measure avoidable hospitalizations and hospital quality. I define an avoidable hospitalization using the Prevention Quality Indicators. These identify patterns of billing codes for admissions which might have been avoided through access to

[13] Technically, a beneficiary could leave the panel by voluntarily terminating both their Medicare and Social Security benefits, but this is extremely rare.

high-quality outpatient care.[14]

The data have two key limitations. First, the utilization data exclude Medicare Advantage beneficiaries. Although spending for these individuals is unobserved, I do observe demographic and enrollment data. This allows me to rule out selection into the estimation sample around the discontinuities. Second, payments made by supplemental insurers are unobserved. Although Medicare records the cost-sharing due, it does not record whether that payment was made directly by the consumer, or by an insurer on behalf of the consumer. My main spending outcome will include spending from all sources, but I also present results for direct Medicare payments to providers.

1.3.3 Survey Data

To provide context for causal results, I also present descriptive results in Appendix Section 2. Using the Medical Expenditure Panel Survey, I measure the correlation between income and various definitions of health expenditures. Although other datasets could provide useful evidence on mechanisms, it is challenging to estimate changes in health and pension benefits using survey data. This is due in part to sample sizes (which are two orders of magnitude smaller) and in part to large, non-classical measurement error in reporting of retirement income (Bee and Mitchell, 2017).

1.3.4 Sample Construction and Summary Statistics

The 100 percent Master Beneficiary Summary File from 2006 to 2011 consists of about 60 million unique beneficiaries. I make several restrictions to create the estimation sample. Most importantly, I focus only on men. I do so for three reasons. First, for women at the margin of claiming on their own record or as a spouse, the income shock near the January

[14] The MedPAR files are only available for 2009 to 2011. I use the composite PQI 90 which includes admissions with diagnoses such as uncontrolled diabetes, bacterial pneumonia, or urinary tract infections. See AHRQ Prevention Quality Indicators Technical Specifications for more details.

2 cutoff may induce them to switch. That is, they may choose to claim based on earnings linked to their own date of birth, or earnings linked to their spouse's date of birth. Allowing women to pick the date of birth at which earnings are computed could bias the results by creating a selected sample around the cutoff. Since almost no men from these cohorts would have higher benefits by claiming on their spouse's wage history, focusing only on men ensures that wage-earner birth date remains fixed.[15] Second, combining wage-earning women with all men would not create a sample representative of a general population. This would limit the external validity of the results. Finally, the links between income and health differ by sex. Fitzpatrick and Moore (2018) show that the mortality effects of claiming early Social Security are much larger for men than women. Focusing only on men will help clarify the discussion of causal mechanisms.

My core sample includes cohorts born between 1927 and 1937. I exclude cohorts born after 1937 because they are subject to changes in the Full Retirement Age (FRA); cohorts prior to 1927 are also excluded because binding maximum taxable earnings thresholds complicate the interpretation of the income shock.[16] I also exclude persons who were originally entitled to Medicare before 65 due to disability, and those who do not receive Social Security Old-Age retirement benefits based on their own wage history. Because my primary outcome is Medicare spending, I exclude those who ever enrolled in Medicare Advantage. To ensure a balanced panel, I exclude people who die within the observation period.[17]

In most specifications I also exclude persons born on January 1 and January 2—a so-called donut hole RD design (Barreca, Lindo and Waddell, 2016). The density of birth dates spikes on January 1 raising concerns about manipulation of the running variable. However, Kopczuk and Song (2008) argue this is the result of clerical errors by the Social Security

[15]My Medicare files do not include the identifier which would allow me to link spouses. About 0.6% of male Medicare beneficiaries receive Social Security benefits as an "aged husband" or "aged widower."

[16]The mean age of my sample is already 75 and excluding older cohorts avoids pushing it even higher.

[17]See Appendix Section 3 for a discussion of how these discontinuities apply to disability benefits. I also exclude a small fraction of beneficiaries who were not continuously enrolled in Parts A and B.

Administration, not manipulation on the part of beneficiaries. Persons born on January 2 also may be selected because Social Security rules interact in a peculiar way that allow them to retire one month earlier than normal.[18]

Table 3.2 reports summary statistics for the 3,287,465 unique beneficiaries in the estimation sample. The majority (87%) are white and not enrolled in a Part D prescription drug plan. The mean age is about 75 years old—10 years after Medicare eligibility and 13 years after Social Security eligibility. On a monthly basis, mean Social Security income is $1,246 and mean payments for Medicare services is $648. Medicare paid providers directly for 84% of these services. The remainder was paid by supplement insurers or by beneficiaries directly out-of-pocket. 35% of the baseline sample alive at the end of 2011 is dead by the end of 2017.[19] Appendix Table A.1 shows that excluding Medicare Advantage beneficiaries leads to a sample that is somewhat whiter, but has a similar mortality rate and share with Medicaid coverage.

1.4 Empirical Strategy

Because my setting does not match a classic regression discontinuity (RD) design, I consider techniques for "stacking" multiple discontinuities. A classic RD assigns a binary treatment when a single running variable exceeds a known cutoff. As RD techniques have grown in popularity, researchers have explored how RD tools can be applied in non-classical settings. Examples include stacked RD designs, which collapse discontinuities from multiple cutoffs (Cattaneo et al., 2016); regression kink designs, which test for discontinuities in slopes (Card

[18]Dropping January 2 birthdates also allows for symmetry around the cutoff. That is, I drop one birthdate on the left and right of the cutoff. See Kopczuk and Song (2008) for a discussion of these issues in the context of Social Security administrative records. Appendix Table A.11 shows that including these birth dates makes no difference to the main results. For more recent cohorts, Jacobson, Kogelnik and Royer (2020) show that the use of Cesarean sections leads to "missing" births near major U.S. holidays.

[19]The mortality rate matches data from life tables which predict 35.4% mortality for men from these cohorts (United States Mortality Database, UC Berkeley).

et al., 2015; Gelber, Moore and Strand, 2018); and difference-in-discontinuities designs, which test for changes in discontinuities over time (Duggan, Gupta and Jackson, 2019; Persson, 2020). My setting does not match any of these categories. Because it includes placebo discontinuities as well as multiple cutoffs, it combines difference-in-discontinuities with stacked, ordered treatments. For this reason, I consider different specifications for aggregating the discontinuities. The goal in each case is to measure a single average elasticity across all cohorts.

1.4.1 Estimating Several Regression Discontinuities Separately

I first run the classic regression discontinuity design separately for each of the 10 cohorts. Following Gelman and Imbens (2019), I estimate a specification with varying linear trends in date of birth:

$$\log(Y_i) = \beta_0 + \beta_1 D_i + \beta_2 DOB_i + \beta_3 (DOB_i * D_i) + e_i \tag{1.2}$$

where i indexes date of birth, Y_i denotes the outcome of interest, DOB_i is a linear trend normalized as the distance in days from the cutoff, and D_i is a dummy for the January 2 cutoff. The coefficient of interest is β_1 which estimates the percentage change in the mean of the outcome at the cohort boundary.

The unit of observation is an average collapsed within a date of birth cell and regressions are weighted by the number of individuals in each cell. I take this approach to be consistent with Gelber, Isen and Song (2016). Using aggregate data estimates standard errors which are more conservative and accounts for correlated shocks at the date of birth level (Angrist and Pischke, 2009). Working with averages also ensures all observations are non-zero. This allows me to consistently use logarithmic specifications.[20]

[20]As a robustness exercise, Appendix Table A.7 also presents results in levels using the individual level data.

The β_1 coefficient identifies a causal effect of the benefit shock on the outcome if two assumptions are satisfied: (i) beneficiaries cannot precisely manipulate their date of birth around the cutoff, and (ii) no potential confounders are also changing discontinuously at the cutoff.

This specification provides a unique β_1 for each of the 10 cohorts in my sample. One technique for summarizing these results would be to compute the average coefficients for positive and negative treatments and to rescale by the average treatment size. However, this method discards valuable information about the ordering of the shocks. Within the group of positive shocks, the effect size should increase with the treatment size. The same pattern should hold in the opposite direction for negative shocks. To take advantage of this ordering, we can regress the $\hat{\beta}_{1,c}$ coefficients from equation (2) for each cohort against the income shock prediction, τ_c, from equation (1). That is,

$$\hat{\beta}_{1,c} = \alpha_0 + \alpha_1 \tau_c + e_c \tag{1.3}$$

Under this framework, α_1 estimates the elasticity of spending with respect to income. Plotting this relationship tests whether the pattern of discontinuities in the data match the pattern of discontinuities from the benefit formula. It also allows us to visually inspect if the effects are symmetric around zero and scale linearly.

1.4.2 Stacked and Scaled Regression Discontinuity

Extending the intuition of plotting coefficients from separate regressions, I next consider a specification which estimates multiple discontinuities in a single equation. By "stacking" multiple cohorts on top of each other I can maximize the efficiency of the estimator.[21]

[21] In Appendix Figure A.2, I present Monte-Carlo simulations which compare different estimators. If the identification assumptions are satisfied, then the stacked and scaled specification in equation (4) is most efficient.

Specifically, I estimate

$$\log(Y_i) = \beta_1(D_i * S_c) + \sum_{c\in C} \beta_{2,c} DOB_i + \sum_{c\in C} \beta_{3,c}(DOB_i * D_i) + \beta_c + e_i \tag{1.4}$$

where C is the set of cohorts over which the relevant discontinuities occur, β_c is a cohort fixed effect, and the summation signs permit slopes to vary by cohort on either side of the cutoff. S_c denotes the cohort-specific income shock. Because the outcome is measured in log points, I adjust the scale to be interpreted as a 1% shock. For example, under a 4% benefit shock $S_c = 0.04$ for all observations within the cohort. The coefficient of interest is again β_1. If the outcome Y_i is measured as the dollar amount of Medicare payments then β_1 estimates the elasticity of Medicare payments with respect to Social Security income.

In addition to the assumptions described above, this specification also assumes the elasticity is equal for all cohorts. That is, if a positive 4% income shock leads to a 4% spending decline, then a negative 1% income shock leads to a 1% spending increase. This is equivalent to plotting the coefficients against the income shock and constraining the linear fit to pass through the origin.[22] Appendix Section 1 describes a specification in the spirit of a difference-in-discontinuities design that allows the intercept to vary.

1.4.3 Bandwidth Selection

In my primary specifications I use a 30-day bandwidth. This is a natural bandwidth for several reasons. First, the beginning dates of Medicare and Social Security eligibility vary discontinuously for persons with birth dates on the second day every month.[23] Keeping the whole sample within a month of the cutoff avoids adding more discontinuities. Second, fertil-

[22]Individuals are unaware they have been "treated," so effects should be symmetric around zero. Non-symmetric effects would only arise from behavioral frictions. The complexity of the shock makes this unlikely.

[23]For most individuals who qualify by age, Medicare eligibility begins on the first day of the month an individual turns 65. For individuals with a date of birth on the first of the month, coverage starts the first day of the prior month.

ity patterns differ systematically by season (Buckles and Hungerman, 2013). Although these trends are smooth, they are not necessarily linear over the course of several months. Using a narrow 30-day bandwidth avoids modeling this seasonality and allows for a transparent, linear specification. Finally, the density of birth dates increases slightly at the first of the month. A 30-day bandwidth drops individuals born on December 1 and February 1 who may be somewhat selected.

In a standard RD setting, there are several procedures for selecting the optimal bandwidth. For example, the Calonico, Cattaneo and Titiunik (2014) method will select an automatic bandwidth and compute robust, bias-corrected confidence intervals. Although this tool was not designed for multiple ordered treatments of different signs, as a robustness check I explore how it could be adapted for my setting. These results are described in Appendix Section 1.

1.5 Aggregate Results

I present results in five subsections. I start by validating that Social Security income across cohorts differ exactly as predicted by equation (1). Next, I look at each cohort separately and provide graphical evidence that income shocks lead to declines in Medicare spending. I then combine all cohorts using a stacked RD specification to estimate a baseline elasticity. I validate the main result using several specification and placebo tests. Finally, I examine health directly by examining chronic conditions and mortality. Overall, the results suggest income can reduce health care spending and mortality.

1.5.1 Evidence of Discontinuities in Pension Income

Benefit amounts are not recorded in the Medicare data, so I use the Public Use Benefits File to confirm that benefit levels in the data follow the pattern predicted by equation (1). The absence of a date-of-birth variable in the public use data makes the preferred specification infeasible. As an alternative, I compare the difference in monthly benefits across cohorts by running a specification that includes no linear trends and a 12 month bandwidth.[24]

I first test if the pattern predicted by equation (1) matches the pattern in changes in the principle insurance amount, the full benefit entitlement before any claiming adjustments. The left panel of Figure 1.1 shows the slope of the fit line is almost exactly one ($\beta = 1.03$ with s.e. $= 0.08$). A related question is if beneficiaries adjust their claiming behavior in response to the shock. This would limit the effect of the discontinuities since beneficiaries might delay retirement in order to offset a decline in benefits. The right panel of Figure 1.1 tests this by comparing differences in cash amounts after claiming adjustments. The pattern is unchanged suggesting the shocks do not affect claiming behavior.

Another question is if the benefit shocks were offset by changes in labor market earnings. To test this, I use the same specification where each observation is a difference between birth cohorts. Figure 1.2 shows the pattern in benefit differences have no relationship to total earnings, or labor force participation on the extensive margin. This suggests the shocks for total income are similar to the shocks for Social Security income.[25]

The result that labor force outcomes do not change is consistent with recent work exploring how behavioral frictions influence retirement decisions. Factors such as incomplete information (Liebman and Luttmer, 2015), framing effects (Brown, Kapteyn and Mitchell, 2016), and cognitive biases (Brown et al., 2019) have large impacts on Social Security claiming choices. These results are incompatible with standard life-cycle expected utility models. In my setting, beneficiaries are unaware they have been "treated." The shocks have no salience, so it is not surprising that short-run labor force outcomes do not adjust. In contrast, when beneficiaries receive a highly salient "treatment," they are more likely to change

[24] This specification is only possible because Social Security benefits do not have a strong age gradient. This is not the case for health spending.

[25] Table 5 in Gelber, Isen and Song (2016) presents additional evidence that labor force outcomes are unchanged. Using SSA and IRS administrative data, they find no discontinuous changes in earnings around January 2 dates of birth for the 1928, 1930, 1932, 1934, and 1936 cohorts.

their behavior (Deshpande, Fadlon and Gray, 2020; Gelber, Isen and Song, 2016).

1.5.2 Graphical Evidence on Expenditures from Separate Cohorts

I next investigate how changes in Social Security benefits affect Medicare spending. To inspect the pattern in the underlying data, I plot total Medicare spending around the date of birth cutoff for negative, null, and positive treatments separately. To provide a consistent comparison across all 10 cohorts, I use equation (2) to residualize away the cohort-specific means and trends on either side of the cutoff. The bottom panel of Figure 1.3 plots spending residuals after this normalization. For context, the top panel plots the average income change for the three treatment types. The negative income shocks are smaller and less frequent, but there appears to be evidence of a symmetric effect: spending increases for negative shocks, remains unchanged for null shocks, and declines for positive shocks.[26]

To visually examine the cohorts separately, I plot the β_1 coefficients from equation (2) for each cohort against the predicted income shock. Figure 1.4 depicts these coefficients and their associated confidence intervals.[27] Several patterns emerge. First, the results are consistent with negative shocks leading to increases in Medicare spending, while positive shocks having the opposite effect. The slope of the OLS fit line, a measure of the elasticity, is -0.95 (standard error 0.43). Appendix Figure A.3 shows a nearly identical pattern when the bandwidth is selected using the Calonico, Cattaneo and Titiunik (2014) non-parametric estimator and robust bias-corrected inference. Second, the fit suggests the assumptions imposed by equation (4) are reasonable. The line intersects the origin which indicates there is, on average, no change for the placebo treatments. There is also no evidence of non-linear treatment effects. Finally, because the estimate for any single cohort is imprecise, it is necessary to estimate all cohorts simultaneously.

[26]A similar pattern is visible in Appendix Figure A.4 which shows the same plots without normalization.

[27]Appendix Figure A.5 plots the raw data for all 10 cohorts individually.

1.5.3 Evidence from Stacked Cohorts

Table 1.3 presents estimates from the stacked equation (4) with the log sum of all spending from 2006 to 2011 as the dependent variable. Column (1) presents the preferred specification with no controls. The elasticity estimates are centered around -0.9 implying that a 1 percent increase in Social Security benefits reduces total Medicare payments by 0.9 percent.

The result is similar across different specifications. Column (2) includes controls for race which may help absorb additional variation. I do not consider other potential demographic controls (Medicaid enrollment, zipcode characteristics) because these are likely to be affected by the benefit shock. Following previous literature I estimate robust standard errors at the date-of-birth level, but I also compute standard errors using day-month clustering (column 3) and jackknife sampling (column 4).[28] While the previous specifications include all 10 cohorts, columns (5) and (6) consider using only the positive treatments, or only the non-zero treatments. This is equivalent to estimating the slope in Figure 1 by omitting the four non-positive treatments, or the two null treatments. The coefficient remains unchanged suggesting particular cohorts are not driving the variation.

Table 1.4 shows the elasticities are similar for direct provider reimbursement payments and cost-sharing. To understand the fiscal implications of these estimates, it is useful to convert the elasticities into dollar values. I do this by replacing the log outcome variable in the main specification with levels and rescaling the treatment value such that a 1% change implies a change of $12.5 per month in Social Security income.[29] I find that if Social Security income increases by $100 per month, then payments from Medicare decline by $38 per month and cost-sharing payments decline by $6 per month.

While the results may appear large, they should be understood as a long-run treatment effect. On average, I observe beneficiaries 10 years after Medicare eligibility and 13 years after

---

[28] As an additional robustness check, Appendix Table A.7 presents results in levels using both the individual level microdata, and the date-of-birth collapsed file. The standard errors are similar in all cases.

[29] The average benefit amount from Table 3.2 is roughly $1,250 per month.

Social Security eligibility. Thus, the changes in health status I observe are the cumulative effect of benefit discontinuities compounded over this entire period. Rather than focusing on health status at the time of claiming, my setting measures changes in health which evolve slowly over time.[30]

1.5.3.1 Validating the Regression Discontinuity Design

The main specification uses linear trends to be consistent with previous work, but I also consider including quadratic terms. Appendix Table A.2 compares four specifications: a baseline linear model, linear with controls, a global quadratic on either side of the cutoff, and cohort-specific quadratic terms on either side of the cutoff. The linear specification with no controls minimizes the Akaike Information Criterion (AIC) and Bayes Information Criterion (BIC), so I use it as the baseline.

Although manipulation of the date-of-birth variable could threaten identification, the risk that beneficiaries deliberately use a fraudulent birth date is low. Not only are birth records difficult to falsify, but the complexity of the benefit formula makes it unlikely beneficiaries are aware these discontinuities exist. Appendix Figure A.6 plots the histogram for each treatment type and provides no evidence of manipulation. Appendix Table A.3 examines this by running regressions with the log count of observations as the dependent variable. Although January birth dates are more common than December birth dates, the pattern holds for all cohorts and does not vary by the sign of the treatment.

A related concern arises from endogenous enrollment into Medicare Advantage. If higher income reduces FFS enrollment this would create a selection bias because the sample would systematically differ on either side of the discontinuity. Although selection into Medicare Advantage would create bunching, I can also test for this directly using the enrollment file. Appendix Table A.4 presents OLS results from equation (4) with log share of the population

[30] I postpone a more detailed discussion of dynamic effects and age heterogeneity for Section 6.7.

as the dependent variable. I do not find evidence the shocks induce changes in coverage in Part A, Part B, or Medicare Advantage.[31]

A final threat to identification is discontinuous changes in other policies around the cutoff date. Potential confounders include school entry dates, tax liability, or strategic birth timing. Appendix Section 1 describes why these cohorts are unlikely to be affected by such policies. Additionally, Appendix Section 1 describes how to test this using a variation on the difference-in-discontinuities design. By assuming linear effects across cohorts, we can test if the pattern of treatment effects is consistent with a null effect for a null treatment. The constant term in Appendix Table A.6 is not different from zero, suggesting other policies are not changing around the cutoff.

1.5.3.2 Sensitivity and Placebo Tests

Given the importance of bandwidth selection to regression discontinuity designs, it is necessary to test how estimates vary under different assumptions. I rerun the regression above for various bandwidth lengths and plot the coefficients with their 95% confidence intervals in Figure 1.5. The coefficient consistently hovers around -1.0 regardless of bandwidth choice and is statistically significant at the 5% level for most bandwidths in the 10 to 50 day window.

As a placebo test, I rerun the regression with the predicted income shocks at placebo cutoff dates. I consider all dates between the months of February and November on either side of the cutoff. With a 30-day bandwidth, this will exclude any sample with the true January 2 cutoff. Figure 1.6 shows the histogram of these coefficients with the solid vertical line denoting the estimate from the baseline specification, and the dotted lines denoting 95% confidence intervals. The baseline coefficient is less than the placebo coefficients 98.8% of the time.

[31]Another potential concern relates to selection around the discontinuity due to mortality prior to the observation period. This is not necessarily a threat to identification, but does change the interpretation of the treatment effect. I postpone a discussion of these issues until section 7.

1.5.4 Regression Evidence on Diagnoses for Chronic Conditions

To examine more directly if income improves health, I consider diagnoses for chronic conditions as another outcome variable. To do this, I replace mean spending in equation (4) with the mean number of chronic conditions in a date-of-birth cell. This provides a simple summary measure of underlying health. For context, the average beneficiary has 4.1 chronic conditions. The two most common are hypertension (61% of the sample) and high cholesterol (57%). For consistency, I use the same specification with a 30-day bandwidth. Table 1.6 estimates the elasticity of the count in chronic conditions with respect to income is about -0.35. This suggests the declines in spending are due at least in part to improvements in health.

1.5.5 Graphical and Regression Evidence on Mortality

The final outcome I consider is an unambiguous measure of health: mortality. Unlike chronic conditions, which are only measured if beneficiaries seek treatment, mortality is measured for the entire sample. To be consistent, the main mortality results use the same baseline sample as the spending results. I measure mortality as either disappearance from the panel between 2012 and 2016, or a recorded death during 2017. For context, the cumulative mortality rate during this 6 year period is 35%.

To visually examine the cohorts separately, Figure 1.7 plots the β_1 coefficients from equation (2) for each cohort against the predicted income shock where the dependent variable of interest is the log fraction of the sample that has died by the end of 2017. The pattern for mortality matches the pattern in spending: larger income shocks are associated with larger declines in mortality. Appendix Figure A.9 shows an identical pattern when using Calonico, Cattaneo and Titiunik (2014) bandwidth selection and robust bias-corrected inference.

Table 1.7 presents regression results under a variety of specifications. Columns (1-2) use the same specification as the main spending results, and column (3) considers a difference-in-discontinuity approach described in Appendix Section 1. Columns (4-6) consider the same specifications under the Calonico, Cattaneo and Titiunik (2014) bandwidth selection procedure. Throughout all specifications, the elasticity is around -1.0. That is, a 1% increase in Social Security income reduces the mortality rate by 1%.

My finding that mortality declines with income is consistent with results from Gelber, Moore and Strand (2018) on Social Security Disability Insurance, but inconsistent with results from Snyder and Evans (2006) on the 1917 Notch. The discrepancy with Snyder and Evans (2006) can be explained by two factors. First, they use Census data with quarter of birth rather than exact date of birth. As Handwerker (2011) discusses in detail, this may threaten the exclusion restriction because variation between cohorts overwhelms variation from the instrument. Furthermore, in Appendix Figure 8 of Gelber, Isen and Song (2016) the authors are unable to replicate the Snyder and Evans result using full population Social Security records with exact date of birth.

Second, while the 1917 Notch changes both income and labor force participation, my setting only affects income. Gelber, Isen and Song (2016) show the Notch was highly salient. Beneficiaries responded by increasing labor force participation and postponing retirement. Given that early retirement can have adverse health impacts, the negative effect of the benefit cut offsets a positive effect from increased labor force participation (Fitzpatrick and Moore, 2018; Kuhn et al., 2019). Conversely, since my policy change has no salience, I find that claiming decisions and labor force outcomes are unchanged. Thus, the positive effect from income is not offset by a negative effect from retirement.

My mortality elasticity is large, but still within the range of estimates from other settings. The paper closest in methods and setting is Gelber, Moore and Strand (2018). They study the effect of income on mortality among Social Security Disability Insurance beneficiaries, another high-mortality, low-income group. Using administrative records and a regression kink design, they estimate an elasticity of -0.56 (s.e. 0.09). My estimate has a wide confidence interval, so I cannot reject that the two are different. Estimates from other contexts include

-0.94 for pension recipients in Russia (Jensen and Richter, 2004), -0.57 for Union Army pensions in the United States (Salm, 2011), and -0.18 for elderly recipients of conditional cash transfers in Mexico (Barham and Rowberry, 2013). In contrast, settings that consider healthier populations protected by generous safety net programs tend to find no effect. In their study of Swedish lottery players, Cesarini et al. (2016) find wealth shocks have no effect on mortality or most types of health care utilization.

1.6 Mechanisms and Heterogeneity

To explore the mechanisms through which income reduces Medicare spending and improves health status, I decompose the aggregate results into more detailed categories. Specifically, I explore how effects differ across demographic groups and health care settings. Several key findings emerge. First, I show the effect is largest for those in middle-income zipcodes and those without supplemental Medicaid or subsidized Part D coverage. This suggests nursing care and prescription drugs are two areas where the additional income makes a difference for health. Second, I find the spending changes are largest for the most expensive patients. Third, I show the decline in chronic conditions is concentrated in the conditions most sensitive to health care, environment, or behavior. Finally, I find limited evidence of improvements in quality or offsetting increases for certain categories of care.

1.6.1 Cost Sharing for Part A and B Services

In most structural models of health investment, higher income should lead to increases in both health and health care services (Grossman, 2000; Hall and Jones, 2007). While I do find that health improves, my finding for utilization is the opposite of what these models predict. Two pieces of background are useful for interpreting the result. First, preventive care accounts for a small share of Medicare spending. Services like immunization or cancer screenings are inexpensive. The majority of Medicare spending goes to acute care and managing chronic conditions.[32] In this sense, high levels of Medicare expenditures are best interpreted as an indicator of poor health. To the extent that Medicare expenditures are a type of health investment, they function asymmetrically. For people who are sick with many chronic conditions, medical spending mitigates rapid declines in the health capital stock. On the other hand, for people who are in good health, additional medical spending will do little to raise the stock of health capital. Once they are up to date on their immunizations and cancer screenings, productive health investment occurs mostly beyond the scope of Part A and B services through other medical services (prescription drugs, long-term care) or lifestyle changes (diet, exercise, stress, environmental factors).[33]

Second, the price at point of service for many beneficiaries is zero. Although beneficiaries can face substantial out-of-pocket costs for uncovered services (for example, long-term care and prescription drugs), supplement insurance pays for most Part A and B cost-sharing. The administrative files do not record the final payer, but data from the National Health Expenditure Accounts suggest only 4% of the cost of Part A and B services is paid out-of-pocket.[34] Within the context of Part A and B services, beneficiaries are minimally exposed to prices. Table 1.4 provides further evidence by comparing elasticities for Medicare provider reimbursement with Medicare cost-sharing payments. If cost-sharing were a binding constraint, then these elasticities should differ. The results indicate they are similar. Because beneficiaries have near full insurance for Part A and B services, health status plays a larger role in consumption decisions than income.

---

[32] Reid, Damberg and Friedberg (2019) show that even if preventative care is defined broadly (evaluation and management visits, preventive visits, care transition or coordination services, and in-office preventive services, screening, and counseling), it still accounts for only 2% of FFS Medicare spending.

[33] Appendix Section 4 sketches a model to formalize this intuition.

[34] Tables 8 and 12 from Age and Gender files National Health Expenditure Data, 2010. Part A roughly corresponds to Hospital Care and Part B roughly corresponds to Physician and Clinical Services.

1.6.2 Income Heterogeneity

The elasticity of spending with respect to income is likely to differ across the income distribution. To examine treatment effect heterogeneity by income, I split the sample into income quintiles based on beneficiary zipcode. I assign income quintiles using the mean Social Security benefits by zipcode and run equation (4) separately for each of the five subsamples.[35] Appendix Figure A.8 shows the effect is largest for the beneficiaries living in middle-income areas, while high- and low-income areas have smaller elasticities. A elasticity closer to zero for high-income beneficiaries is unsurprising given they have the fewest constraints on health investment. For low-income beneficiaries, the result may reflect the impact of additional safety net programs.

Beneficiaries near the federal poverty line are eligible for Medicare Savings Programs which exempt them from cost-sharing, Part B premiums, and many Medicare Part D costs.[36] This may eliminate a key barrier to care. The lowest-income beneficiaries are additionally eligible for full Medicaid benefits which includes services Medicare does not normally cover like nursing home and dental care. They also may be eligible for Supplemental Security Income (SSI) which effectively bounds the minimum benefit amount at 75% of the poverty line.[37] Appendix Figure A.8 also splits the sample by Medicaid enrollment and provides evidence the changes are largest for those not enrolled in Medicaid. Finally, because the benefit shock is the same in percentage terms across the income distribution, low-income beneficiaries have the lowest level change in the dollar value of benefits.

[35] One caveat to the results from this disaggregation is that the income shock may lead to sorting around the discontinuity if beneficiaries respond by moving neighborhoods.

[36] Eligibility depends on income, assets, and state of residence. See Data Book: Beneficiaries Dually Eligible for Medicare and Medicaid (MACPAC) for details. Appendix Table 2 does not find evidence the income shock influences Medicaid enrollment.

[37] 3% of OASI recipients over 65 also receive SSI payments (SSA Statistical Supplement 2010, Table 3.C5).

1.6.3 Quantile Treatment Effects

Focusing only on the mean treatment effect may overlook large changes in the distribution of health care spending. In particular, the spending distribution has a long right tail, so most of the effect could be driven by a small share of the sample. To compute quantile treatment effects, I estimate equation (4) replacing mean spending in a date-of-birth cell with the decile of spending in a cell. Figure 1.8 plots the coefficients for each spending decile and its associated confidence intervals. Although some estimates are noisy, there is a clear pattern that effects are largest at the top of the spending distribution.

The result is consistent with two explanations. First, the sickest patients with the highest spending may have the most to gain from additional income. This follows from a concave health production function where productive investments are made outside of spending on Part A and B services. For example, if an income shock reduces mental stress, the value of lower stress would be greater for someone in worse health. Second, higher income may induce patients to use more efficient providers. If providers differ in their supply elasticities, an income shock that induces patients to shift away from the most expensive providers would cause a large decline in spending.

1.6.4 Composition of Spending

To examine potential substitution effects across different types of care, I decompose aggregate spending into its subcomponents.[38] Figure 1.9 shows how each category contributes to the share of the variation in total spending. The percentages are normalized to sum to negative one. Nearly every category of spending declines with hospitals accounting for almost half of the total effect. This is consistent with income causing improvements in health and reducing the demand for acute care.

[38] The categories do not correspond neatly to definitions of preventive care, but evaluation/management and physician office visits are the closest.

Spending on physician-administered drugs (Part B drugs) also shows large declines. In contrast to normal Part D prescription drugs, Part B drugs include vaccinations, chemotherapy drugs, injections for rheumatoid arthritis, and biologics for autoimmune conditions. Although part of this decline can be explained by health improvements, supply side factors may also contribute. Medicare reimburses physicians for these drugs through a formula that provides incentives for prescribing expensive treatments (Chandra and Garthwaite, 2019). If providers for low-income patients emphasize profitability while providers for high-income patients emphasize clinical need, then provider switching would cause declines in spending. This would be consistent with evidence from randomized trials showing that patients who are assigned to higher-quality physicians receive less expensive treatment (Doyle, Ewer and Wagner, 2010). Similarly, hospitals in higher income regions are more likely to rapidly adopt cost-effective innovations (Skinner and Staiger, 2015).

1.6.5 Composition of Chronic Conditions

Like spending, the chronic conditions outcome can also be disaggregated at a more detailed level. Because precision is limited when considering each of the 27 conditions individually, I group the conditions into 3 categories that have a high, moderate, and low likelihood of being affected by the income shock. I do this by considering the variance of conditions across zipcodes. Intuitively, conditions which vary widely across geographies are most likely to be affected by health care, environment, or behavior. This is similar to how other papers define discretionary versus non-discretionary hospital procedures (Kaestner and Sasso, 2015). Specifically, I implement the following method: (i) compute mean prevalence of condition by zipcode; (ii) estimate residual mean after controlling for race and age; (iii) compute variance of residual across zipcodes; (iv) rank 27 chronic conditions by variance and divide into 3 equal sized categories.

The procedure creates groups corresponding to standard etiologies. Hypertension and diabetes are classified as high sensitivity conditions; COPD and depression are moderate sensitivity; prostate cancer and hip fracture are low sensitivity. To test the hypothesis that income affects the high sensitivity conditions more, I rerun the same specification for each of the three sensitivity groups. The top panel of Figure 1.10 shows how each category contributes to the share of the change in the total condition elasticity. As expected, the decline is driven by high sensitivity conditions. The bottom panel shows elasticities for the top 5 most common chronic conditions. Although the precision for individual conditions is limited, the point estimate is consistently negative.

1.6.6 Hospital Visits and Quality

To better assess how income affects the demand for hospital care, I use the hospital claims data to construct measures of provider quality and service intensity. Table 1.5 presents the elasticity for total acute stays as well as readmissions and avoidable admissions, two commonly used proxies for quality-of-care.[39] The elasticity for total acute stays is precisely estimated and shows a clear decline. Additionally, there is some evidence that readmissions and avoidable admissions decline as well. This is consistent with higher income beneficiaries receiving higher quality care in both inpatient and outpatient settings.

To investigate hospital quality directly, I decompose the change in hospital spending into high- and low-quality hospitals. I define a high-quality hospital as those with an above average risk-adjusted mortality rate. These ratings were developed by CMS and have been validated by (Doyle, Graves and Gruber, 2018). Specifically, they use quasi-random assignment of ambulances to hospitals to show the CMS ratings are causally associated with clinical outcomes. Columns (4-5) of Table 1.5 shows results of the quality decomposition. The confidence intervals for both categories are large and we can not reject they are equal. One challenge to estimating hospital quality is that claims data are available only for 3 years instead of the 6 years for more aggregate utilization measures.

[39] Following guidelines from AHRQ, I define "avoidable admissions" as admissions for one of 16 ambulatory care sensitive conditions using the PQI 90.

1.6.7 Dynamics and Age Heterogeneity

An open question in the literature is how treatment effects evolve over time. For example, increased spending on preventive care could lead to decreased spending on acute care several years later. To test this, Appendix Figure A.7 plots the elasticities and their associated confidence intervals for each of the six years in the data. Although we can rule out large positive effects in the earlier years, there does not appear to be a clear pattern. Appendix Section 1 presents another technique for measuring age heterogeneity using a date-of-birth by sample-year level regression. Appendix Table A.8 does not find an effect, but the results are imprecise. I do not observe the sample until several years after the benefit shock, so data from ages closer to 65 may tell a different story.

1.6.8 Using Survey Data to Test Other Mechanisms

Because my results rely on a narrow 30-day window around the date-of-birth cutoff, it is not possible to investigate other health inputs using survey data. In addition to a narrow bandwidth, allowing the slope of the age gradient to differ by cohort is key to identification. This is not possible using public use data with anonymized birth dates. Although information on exact date of birth is available in restricted access versions of datasets like the HRS or NHIS, power calculations suggest the sample sizes would be at least an order of magnitude too small.

1.7 Implications for Fiscal Policy

How would a 1% increase in Social Security benefits affect total federal outlays? Beyond the direct effect of paying Social Security benefits, the estimates in this paper suggest there are two indirect effects. First, per-capita Medicare spending on Part A and B services would decline. Second, life expectancy would increase slightly, the number of beneficiaries would go up, and spending on both Social Security and Medicare would grow.

Quantifying these effects requires several assumptions. To start, it is necessary to model how spending and mortality elasticities vary over time. The baseline estimates are for cohorts that, on average, have been treated for 10 years. Because the results on age heterogeneity are too imprecise to draw conclusions about dynamic treatment effects, I consider three potential scenarios. The first is a phase-in assumption where treatment effects are zero at age 65, increase linearly until the period we observe in the data, and then remain constant. Second, I consider a more aggressive scenario where the elasticities are the same for all ages before and after the observation period. Finally, I consider a more conservative scenario where the elasticities are zero before the observation period, and then constant afterwards.

To model base levels of spending and mortality over the life-cycle, I use the 2011 distribution of average Medicare spending by age and the 2011 SSA Life Tables. I also assume that those who are on the margin of dying—that is, those whose deaths are postponed due to the extra income from higher benefits—have the same Medicare spending as those who are always alive. In practice, they are likely to have higher spending since they are in worse than average health, but quantifying the magnitude is challenging.

Table 1.8 summarizes the costs relative to baseline Social Security spending for the different scenarios. For the phase-in scenario, the spending effect from more beneficiaries is larger than the savings from lower per-capita Medicare spending. The net effect for this scenario implies that increasing Social Security benefits by $1 would lead government expenditures to increase by $1.09. However, under the third scenario with more conservative assumptions, a $1 increase would lead government expenditures to increase by only $0.95. The net fiscal effect is sensitive to these assumptions because per-capita Social Security spending is much higher than per-capita Medicare expenditures near age 65.[40]

Several caveats apply. First, these estimates apply only to men. Evidence from Fitz-

[40]The per-capita Medicare externality differs from the $0.38 estimate earlier in the text because this exercise extrapolates the effect over the entire post-65 life-cycle.

patrick and Moore (2018) suggest that causal effects for women may be lower. Second, these only apply for individuals in FFS Medicare plans. The extent to which they apply to Medicare Advantage plans depends on how much health investment is unobserved. Third, they assume no change in labor force participation. Although they may be relevant to policy proposals that change benefits gradually (i.e. by converting to chained price indexes for inflation), more salient policy changes that affect retirement decisions are likely to have different effects.

From the perspective of optimal policy design, cost-benefit analysis should not be limited to the federal budget. A more complete accounting would include the value of quality-adjusted longevity, the insurance value of annuitization, and deadweight losses associated with tax financing. A final caveat is that my results represent a partial equilibrium effect. Predicting the effect of income shocks large enough to induce changes in technology or aggregate supply is beyond the scope of this paper.

1.8 Conclusion

I have argued that Social Security income reduces health care spending and mortality for elderly men. My evidence is based on comparing health outcomes for individuals born just before versus just after cutoffs which lead to discontinuous changes in Social Security income. I showed these income changes are in fact binding, and provide evidence that their lack of salience leads to no offsetting changes in labor force income. Next, I showed cohorts with positive income shocks experience declines in spending on Medicare covered services. I also found reductions in chronic conditions and mortality, supporting the view that underlying health is improving.

Although my ability to identify the mechanisms behind the effect is limited, the results indicate which channels might be at work. First, the income shock occurs starting at age 62, so early-life factors such as education or family background can be ruled out. Similarly, the results in section 5.1 suggest labor force decisions also do not play a role. Second, spending declines across nearly all settings for all years in the data. This suggests that income generates health investment mainly outside Medicare Part A and B services. Furthermore, the effect appears to be largest for those in the middle of the income distribution and those not covered by Medicaid. This suggests that Medicaid and subsidized Part D drug coverage protect beneficiaries from some of the challenges associated with low-income. Given that take-up rates for these programs are low, policymakers may consider simplifying eligibility rules or expanding auto-enrollment. Additionally, the income effect is largest for the sickest and most expensive patients. Targeted income transfers on those with the worst health may be an effective strategy for reducing the growth rate of health spending.

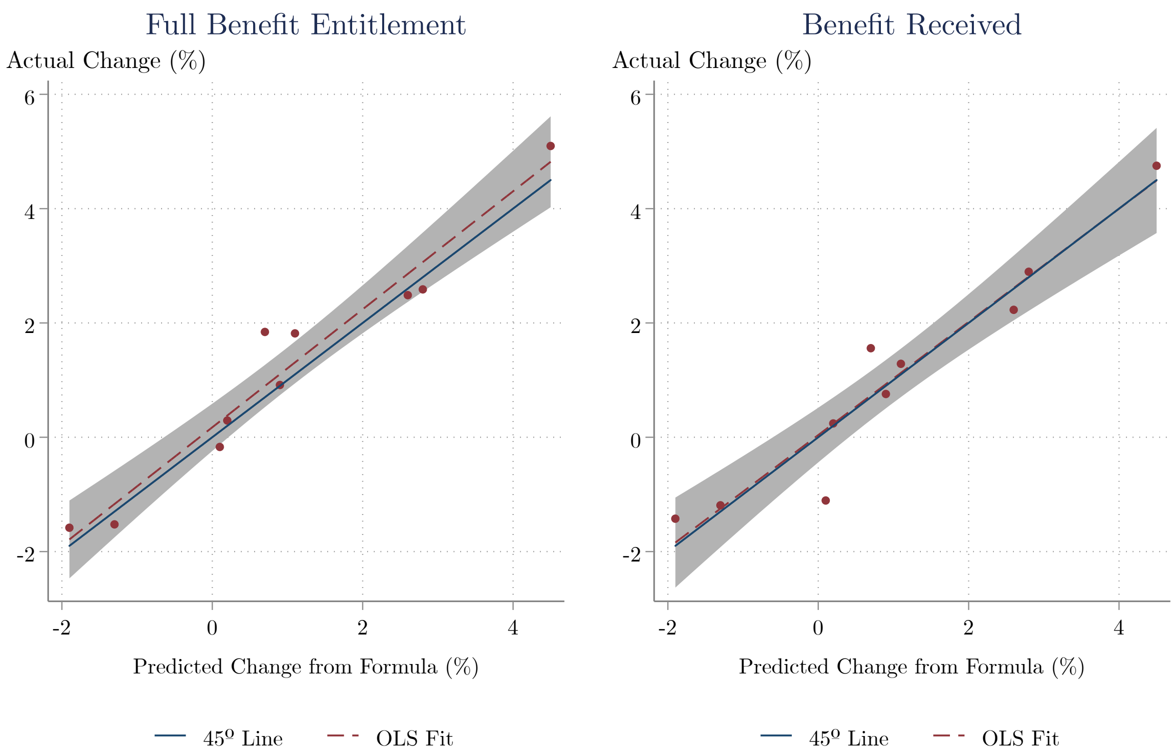
Examining my results in the context of existing research suggests there are other potential channels. For example, income is likely to improve mental and emotional health. Evidence suggests less financial strain can reduce stress, improve decision making ability, and reduce the burden of physical disease (Ridley et al., 2020). Given that the elderly suffer from high rates of anxiety, mood disorders, and depression, additional income may help ease the psychic burdens of aging (Golberstein, 2015).

Another potential mechanism is that income may provide protection from environmental risks. For example, there is evidence pollution has a large, negative effect on the health of older adults. Using Medicare claims, Deryugina et al. (2019) find exogenous changes in airborne particulate matter lead to increased hospitalizations and mortality. Similarly, Bishop, Ketcham and Kuminoff (2018) provide evidence that pollution can cause greater risk of Alzheimer's disease and related dementias. Another environmental risk is extreme temperatures. Deschênes and Moretti (2009) show heat waves and extreme cold temperatures lead to increased mortality for the elderly. This research suggests that additional spending on an air purifier, an air conditioner, or home insulation could have meaningful impacts on mortality.

Future research could explore these mechanisms by using a similar research design and other large databases. The holy grail would be to link the universe of Medicare claims with the universe of Social Security benefits and earnings records. Such a dataset would provide extraordinary opportunities to study the relationship between income and health using quasi-experimental research designs.[41] A less ambitious extension would be to explore how the income shock affects Medicaid, Medicare Advantage, or Medicare Part D claims. Because these programs cover other health services under unique cost-sharing structures, the income effect may differ. Another extension could use credit bureau data to provide evidence on how income affects household finances. A richer understanding of these mechanisms would help design more effective social insurance programs.

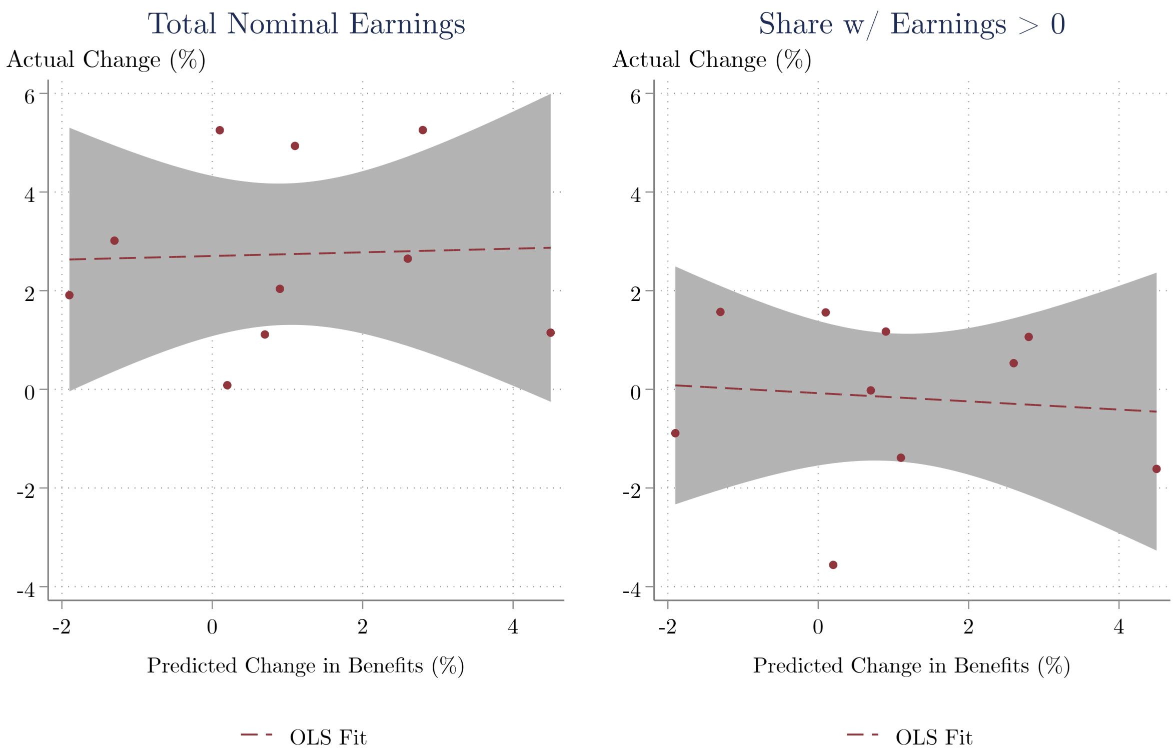
[41]For example, one obvious extension would be to apply the regression kink design of Gelber, Moore and Strand (2018) with health spending as an outcome. Although these linkages are available in some survey data, to my knowledge, SSA and CMS have never merged these datasets at a large scale.

Figure 1.1: Differences in Benefits by Cohort



Notes: Each observation is a percentage difference between birth cohorts aggregated at the annual level. The x-axis denotes the predicted difference implied by equation (1) assuming two identical earners born on either side of the cohort boundary. The left panel shows the full retirement amount before any early claiming penalties, and the right panel shows the actual benefit amount credited. In both cases, the difference in benefits is similar to the difference predicted by the benefit formula. The sample includes males receiving retirement benefits as a primary earner and born between 1927 and 1937. Source: SSA Benefits Public-Use File, 2004.

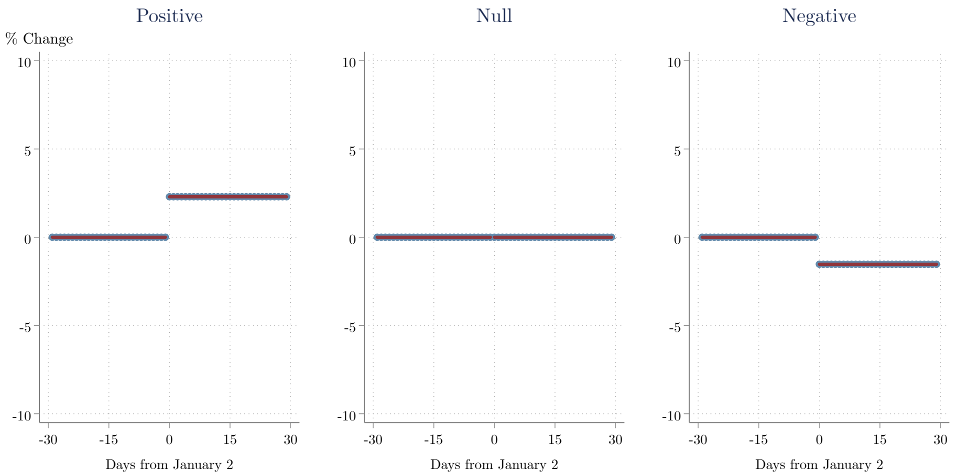
Figure 1.2: Labor Force Outcomes by Cohort



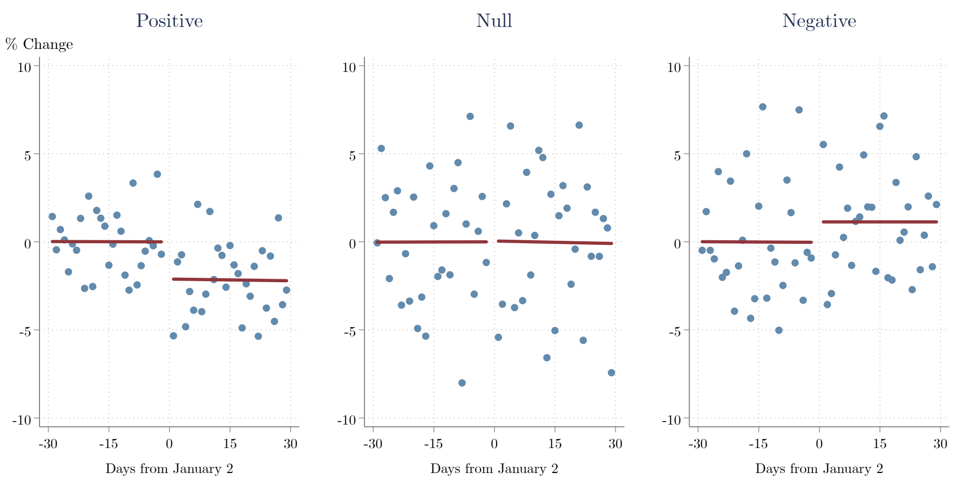
Notes: Each observation is a percentage difference between birth cohorts aggregated at the annual level. Earnings for all cohorts are observed from ages 62 to 66. The left panels shows there is no relationship between the changes predicted by the benefit formula and total nominal earnings. The right panel tests for changes along the extensive margin by examining the share with any labor earnings. Again, there is no clear relationship. This suggests the benefit discontinuities do not influence labor market outcomes.

Figure 1.3: Comparing Benefits and Medicare Spending by Treatment Type

(a) Social Security Benefits

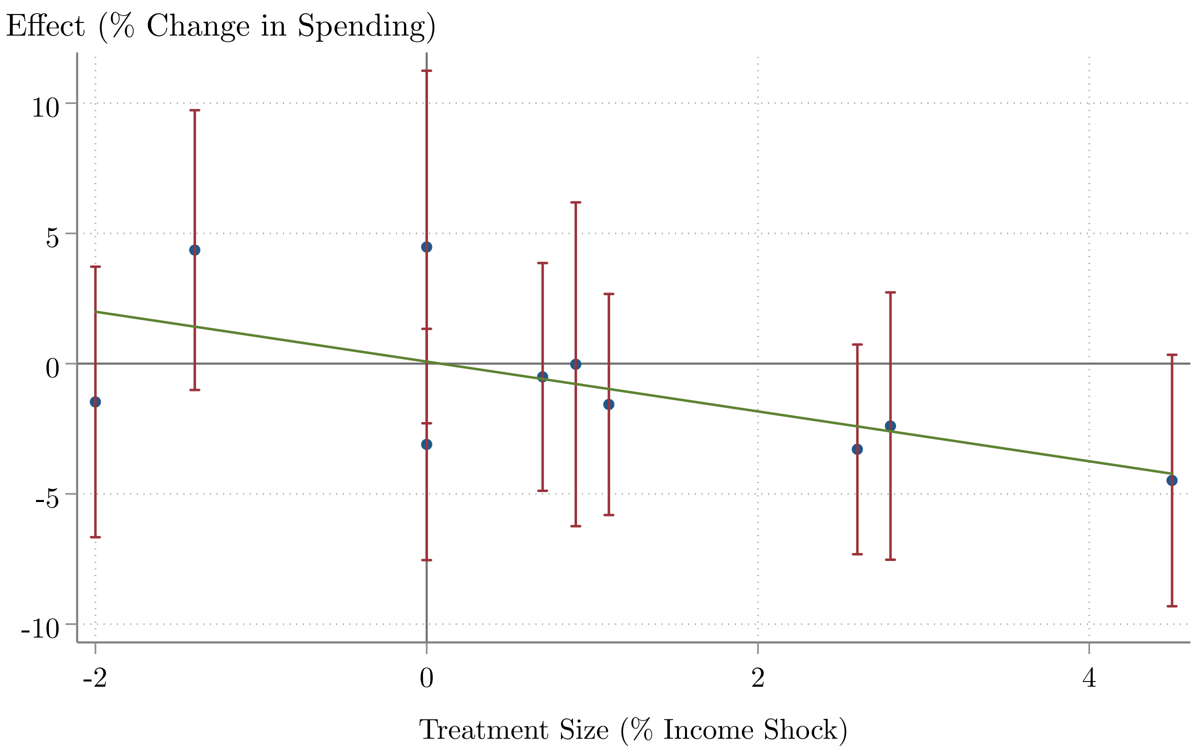


(b) Total Medicare Spending (2006-2011)



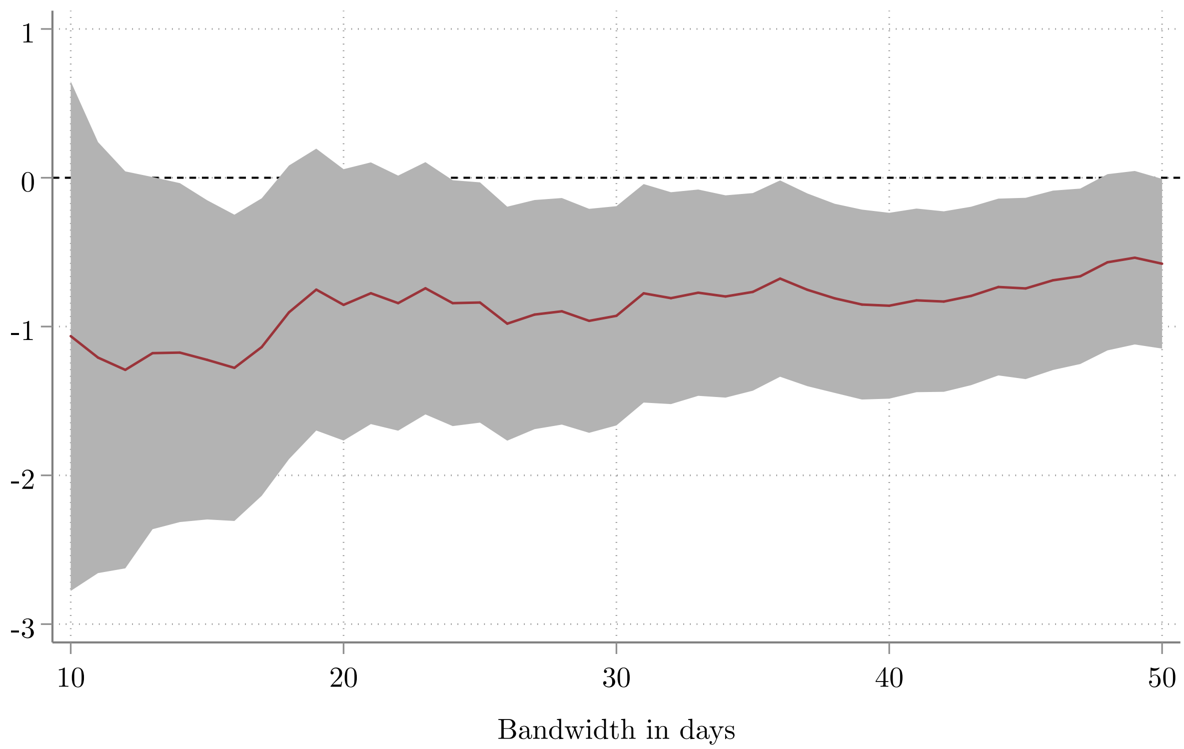
Notes: Columns from left to right show the six cohorts with positive shocks, the two with no shock, and the two with negative shocks. The top row (a) uses equation (1) to compute the change in relative Social Security benefits for individuals with identical nominal earnings. The bottom row (b) shows residualized total log Medicare Part payments where I use equation (2) to remove cohort-specific means and trends on either side of the cutoff. Each observation is an average within a date of birth cell. For positive shock cohorts, average benefits increase by 2.1% and spending declines 2.2%. For negative shock cohorts, average benefits decline by 1.7% and spending increases 1.1%. See text for details on sample restrictions.

Figure 1.4: Parametric Spending Effects for Each Cohort



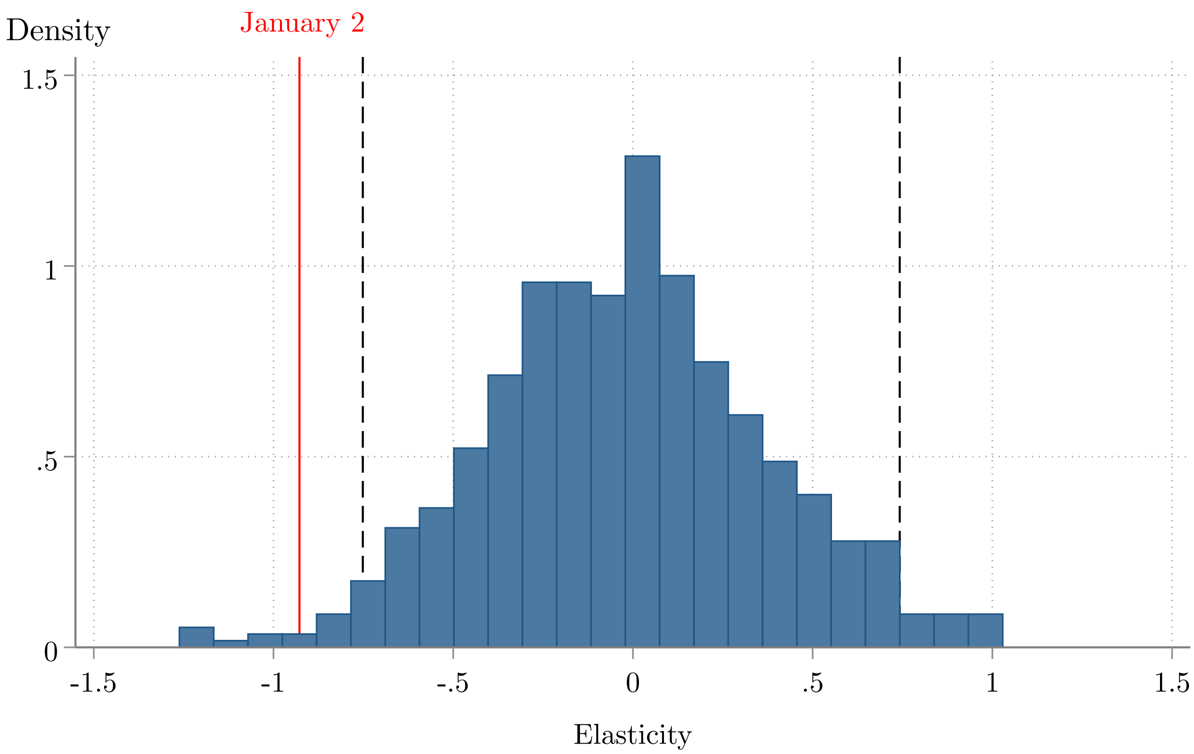
Notes: Each observation is the β_1 coefficient from equation (2) for a given cohort. Confidence interval are computed using robust standard errors. The x-axis denotes the predicted income difference implied by equation (2).

Figure 1.5: Bandwidth Sensitivity Test



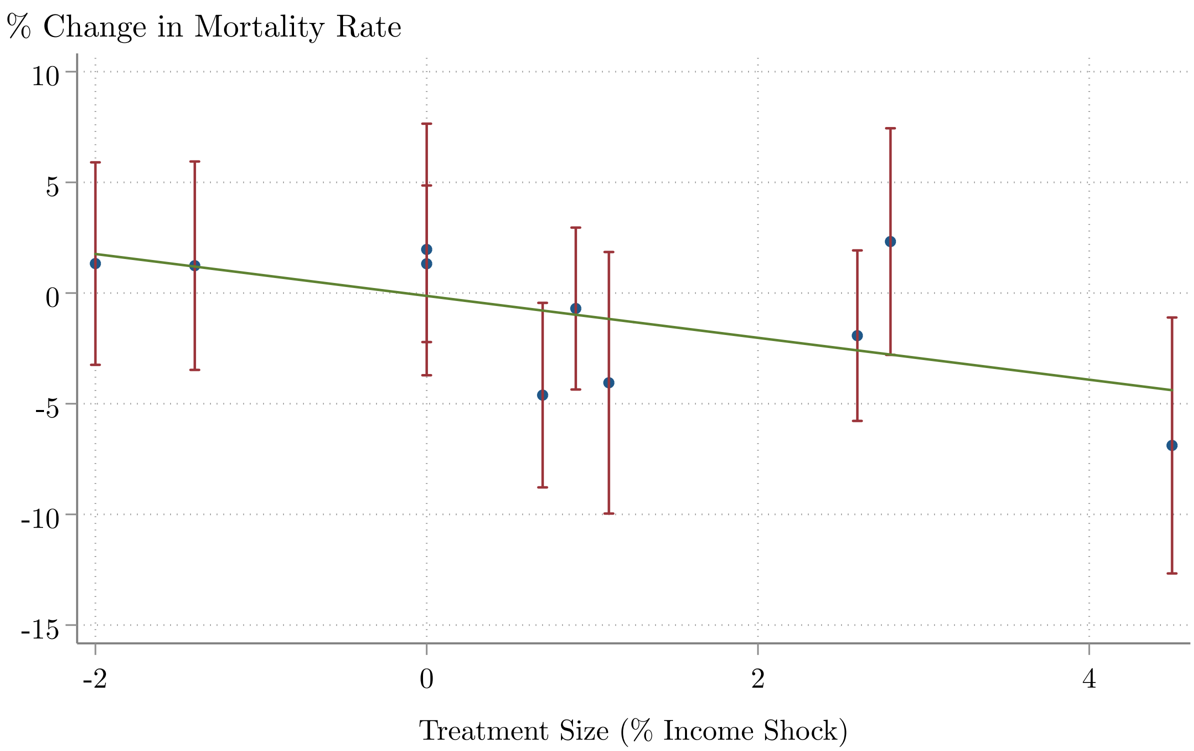
Notes: The solid line depicts the point estimate for the β_1 coefficient from equation (4). The shaded gray area depicts the associated 95% confidence interval. The point estimate consistently hovers around -1 suggesting the results are not sensitive to bandwidth choice.

Figure 1.6: Distribution of Coefficients for Placebo Discontinuity Dates



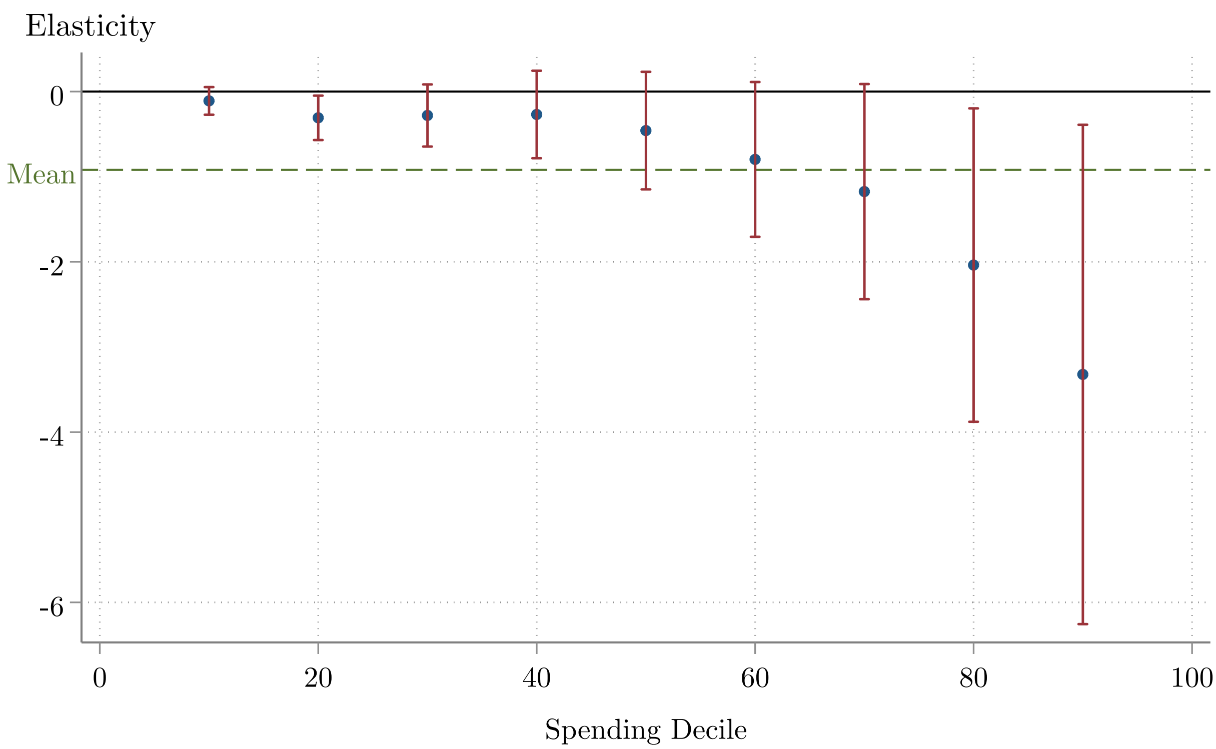
Notes: Coefficients are estimated using 606 placebo dates from months of February to November on either side of the cutoff. Dashed black lines denote the 95% confidence interval. The solid red line denotes the preferred elasticity estimated using equation (4).

Figure 1.7: Parametric Mortality Effects for Each Cohort



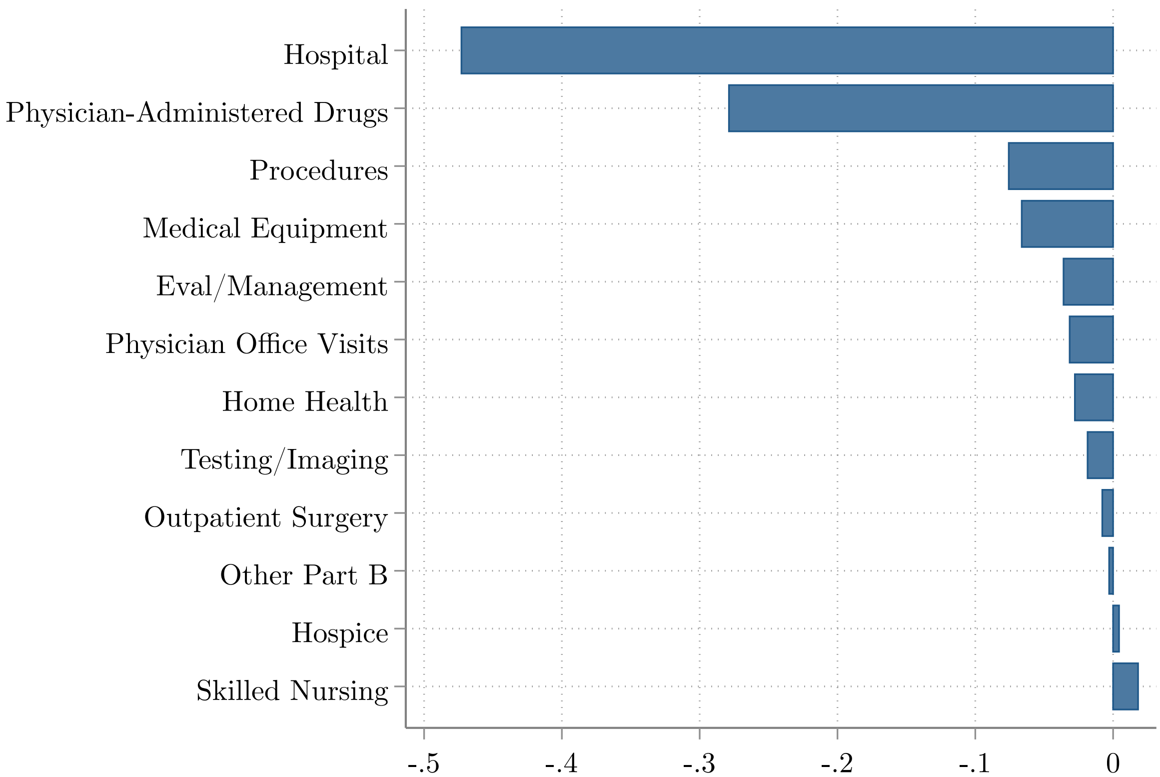
Notes: Each observation is the β_1 coefficient from equation (2) for a given cohort where the dependent variable is the log fraction of the baseline sample that has died by the end of 2017. Confidence interval are computed using robust standard errors. The x-axis denotes the predicted income difference implied by equation (2).

Figure 1.8: Elasticity by Decile of Total Spending



Notes: Coefficients are from a level regression of equation (4) where the dependent variable is the corresponding decile of total Medicare spending. The dashed green line depicts the mean elasticity for the full sample.

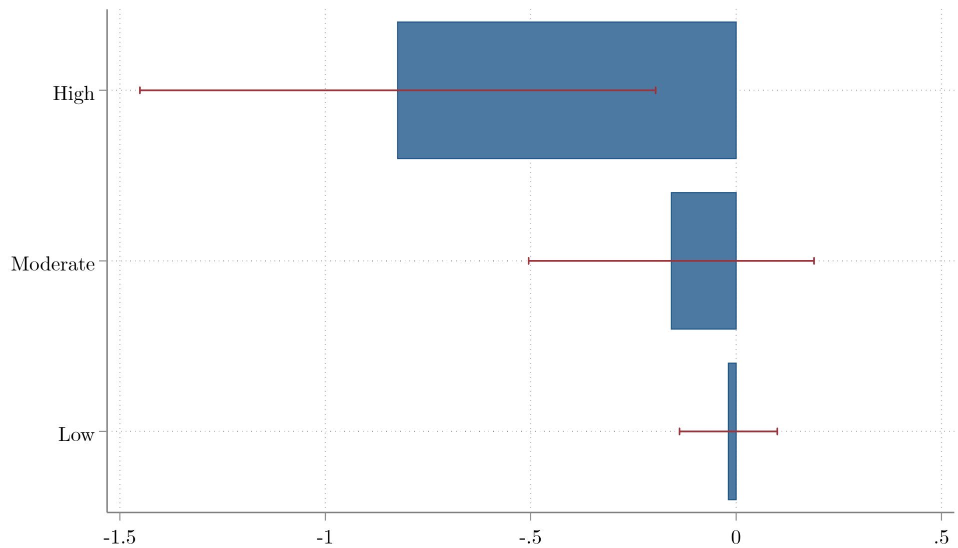
Figure 1.9: Decomposing the Aggregate Effect by Spending Type



Notes: Spending type in levels are estimated using equation (4). Contributions to percent change in total spending are normalized to sum to negative one. Hospital category includes Inpatient Part A and Outpatient Part B. Physician-Administered Drugs refers to Part B drugs. Examples of procedures include, endoscopy, hip replacement, pacemaker insertion, or angioplasty. Other Part B includes payments for anesthesia, dialysis, ambulance, chiropractor, and other unclassified services.

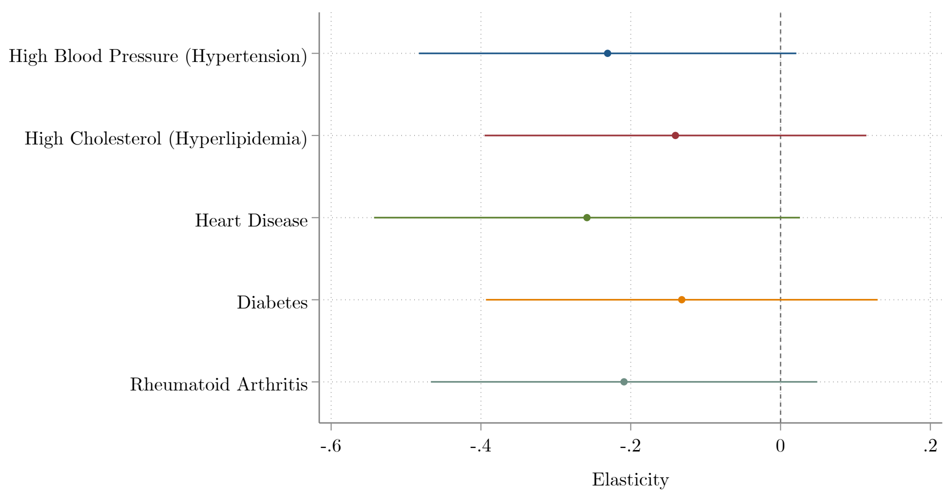
Figure 1.10: Elasticities of Chronic Conditions

(a) Percent Contribution to Chronic Condition Elasticity by Income Sensitivity



Notes: Conditions are classified as high, moderate, or low sensitivity to income shocks based on their variance across zipcodes. Conditions counts in levels are estimated using equation (4). Percent contribution to change in all condition elasticity are normalized to sum to negative one.

(b) Elasticities for Top 5 Chronic Conditions



Notes: Coefficients and their confidence intervals are computed using equation (4). Each point estimate denotes the percentage change in the fraction of the population with a chronic condition for a 1% increase in Social Security income. Negative coefficients imply income reduce disease incidence.

Table 1.1: Treatment Sizes by Cohort

Cohort Discontinuity	Treatment Size (%)	Monthly Income Difference ($)	Age	Years Exposed	Lifetime Income Difference ($)
1927/1928	0.2	3	81	21	720
1928/1929	-1.3	-16	80	20	-3,840
1929/1930	0.9	11	79	19	2,640
1930/1931	0.7	9	78	18	2,160
1931/1932	2.6	31	77	17	7,440
1932/1933	-1.9	-24	76	16	-5,760
1933/1934	0.1	1	75	15	240
1934/1935	1.1	14	74	14	3,360
1935/1936	2.8	35	73	13	8,400
1936/1937	4.5	59	72	12	14,160

Notes: Monthly income difference compares benefits for two workers born a month apart. We assume they have identical earning histories of the average wage index for their whole career. See Appendix for details. Dollar values are inflation-adjusted using SSA COLAs to 2011 levels. Age, years exposed, and lifetime income difference is computed as of 2011.

Table 1.2: Estimation Sample Summary Statistics

	Mean	Standard Deviation
Restricted Use Medicare Data		
White	0.88	
Dual Eligible ($\leq$ 135% FPL)	0.08	
Part D Enrolled	0.43	
Age	75.5	2.9
Monthly Beneficiary Cost-Sharing Liability	105	112
Monthly Payments Made by Medicare	543	717
Total Monthly Medicare Payments	648	818
Number of Chronic Conditions	4.1	
Died before 2018	0.35	
Public-Use SSA Benefits File		
Monthly Benefit Amount	1,246	408
Age at Claiming	63.6	1.63

Notes: Medicare data is computed from 100% Master Beneficiary Summary File 2006 to 2011. The estimation sample consists of men who are born from 1927 to 1937, never-disabled, receiving Social Security benefits on their own wage history, continuously enrolled in Parts A and B, never enrolled in Medicare Advantage, and alive at end of 2011. The final Medicare sample includes 3,287,465 unique beneficiaries. Social Security data is from the 1% Benefits Public-Use File in 2004. All dollar amounts are nominal spending between 2006 and 2011.

Table 1.3: Log Total Medicare Spending

	(1)	(2)	(3)	(4)	(5)	(6)
Elasticity	-0.928**	-0.903**	-0.928**	-0.928**	-0.979**	-0.911**
	(0.376)	(0.377)	(0.458)	(0.390)	(0.402)	(0.375)
Observations	570	570	570	570	342	456
R^2	0.841	0.842	0.841	0.841	0.849	0.847
Controls		Y				
Treatment Sign	All	All	All	All	Positive	Non-zero
Cluster	DMY	DMY	DM	Jacknife	DMY	DMY

Notes: OLS regressions from equation (4) using a 30 day bandwidth. Each observation is a date-of-birth cell. The dependent variable is log total spending for Medicare Part A and B from 2006 to 2011. Controls include percents white, black, hispanic, and asian. "All" treatments include all 10 cohorts including placebos. "Positive" treatments include the six positive income shocks. "Non-zero" treatments exclude the two placebo cohorts. DM denotes day-month level clusters and DMY denotes day-month-year level cluster. The number of individuals when using all cohorts is 455,986. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.4: Log Spending By Payer and Category

	Medicare Payments	Cost-Sharing Payments	Total Part A	Total Part B
Elasticity	-0.943**	-0.847***	-0.699	-1.075***
	(0.397)	(0.303)	(0.587)	(0.327)
Observations	570	570	570	570
R^2	0.839	0.831	0.822	0.761

Notes: OLS regressions from equation (4) using a 30 day bandwidth. Medicare payments are Medicare reimbursements directly to providers. Cost-sharing payments are paid to providers by either supplement insurers or beneficiaries.

Table 1.5: Log Spending / Admission Counts

	Acute Stays	Readmissions	Preventable Admissions	High Quality Hospitals	Low Quality Hospitals
Elasticity	-1.157***	-1.151	-1.774	-1.299	-1.483
	(0.430)	(1.153)	(1.364)	(1.311)	(1.251)
Observations	570	570	570	570	570
R^2	0.841	0.566	0.603	0.603	0.603

Notes: OLS regressions from equation (4) using a 30 day bandwidth.

Table 1.6: Log Count of Chronic Conditions

	(1)	(2)	(3)	(4)	(5)
Elasticity	-0.396**	-0.377**	-0.396**	-0.296	-0.376**
	(0.185)	(0.173)	(0.179)	(0.185)	(0.174)
Observations	570	570	570	342	456
R^2	0.927	0.928	0.927	0.928	0.928
Controls		Y			
Treatment Sign	All	All	All	Positive	Non-zero
Cluster	DM	DMY	Jacknife	DMY	DMY

Notes: OLS regressions from equation (4) using a 30 day bandwidth. Each observation is a date-of-birth cell. The dependent variable is log count of chronic conditions from the 2011 chronic condition summary file. See main text and notes on Table 3 for additional details.

Table 1.7: Log Percentage Dead by End of 2017

	30 Day Bandwidth			CCT Bandwidth Selection		
	(1)	(2)	(3)	(4)	(5)	(6)
Elasticity	-0.982**	-0.988**	-0.951**	-1.071***	-1.091***	-0.979***
	(0.420)	(0.423)	(0.457)	(0.309)	(0.309)	(0.337)
Intercept			-0.136			-0.408
			(0.868)			(0.667)
Observations	570	570	570	936	936	936
R^2	0.968	0.968	0.968	0.966	0.966	0.966
Race Controls		Y			Y	

Notes: Columns (1, 2, 4, 5) report regressions from equation (4) and columns (3, 6) report regressions equation (5). Columns (1-3) use a uniform 30 day bandwidth for all cohorts and columns (4-6) use separate CCT bandwidths for each cohort. The dependent variable is the log fraction of the baseline sample alive at the end of 2011 that has died by the end of 2017.

Table 1.8: Fiscal Costs

			Dynamic Treatment Assumptions		
Program	Type	Source of Change	Phase-in	Same	Zero
Social Security	Direct	Per-capita benefits	1.00	1.00	1.00
	Indirect	Total beneficiaries	0.25	0.34	0.15
Medicare	Indirect	Per-capita benefits	-0.46	-0.56	-0.36
	Indirect	Total beneficiaries	0.30	0.42	0.16
Net Effect			1.09	1.20	0.95

Notes: Costs are reported relative the direct effect of increasing per-capita Social Security benefits. A net effect of 1.09 implies that raising per-capita Social Security benefits by $1 would increase total federal expenditures across Social Security and Medicare would increase by $1.09.

CHAPTER 2

Unintended Effects of Medicare Cost-Sharing Exemptions: Evidence from the QMB Program

2.1 Introduction

The implementation of the Affordable Care Act has prompted a debate among policymakers about how future expansions of public health insurance should be structured. One view is that any new expansion should be modeled after Medicare, a program with some cost-sharing and a wide network of providers that receive moderate reimbursement rates. An alternate model is Medicaid, a program with almost no cost-sharing, but modest provider reimbursement. Designing a system that preserves access and minimizes moral hazard depends on setting the right balance between consumer and supplier prices.

To study this tradeoff, I examine the Qualified Medicare Beneficiary (QMB) program—a means-tested benefit that exempts low-income Medicare beneficiaries from their cost-sharing obligations. Patients enrolled in the QMB program face zero prices for Medicare services, receive monthly premium exemptions, and subsidized Part D prescription drug coverage. However, because providers face administrative burdens and lower reimbursement rates, they may limit access for QMB patients. Given these offsetting features, it is unclear whether patients benefit from enrolling in the program.[1]

To examine the effect of QMB enrollment on patient access and outcomes, I study an

[1] A broader program evaluation would consider externalities on non-Medicare patients. Because private payers are not included in my data, any analysis of spillover effects is beyond the scope of this paper.

expansion of QMB eligibility that occurred in Connecticut in 2009. By doubling the income limit and removing the asset test, the state more than doubled enrollment. Using a difference-in-difference design and Medicare administrative records, I estimate the effect of QMB enrollment on health care utilization and health outcomes. I find that the program appears to be successful at reducing beneficiary costs without limiting access to care. I show QMB enrollment leads patients to save $1,300 in annual outpatient cost-sharing liability without any change in the utilization of outpatient services. Overall, my results suggest that within the universe of Medicare, the program transfers a substantial share of producer surplus to consumers.

This paper contributes to prior work on access to care among Medicare dual-eligibles by providing the first quasi-experimental evaluation of expanded QMB eligibility. Although many papers have studied barriers in access to care for QMBs and other dual-eligibles, previous work has struggled to control for unobserved confounders. For example, Zheng, Hoover and Feng (2014) finds that dual eligibles have more outpatient evaluation and management visits in states with higher dual reimbursement rates. CMS (2015) provides anecdotal evidence that QMBs struggle to access high-quality care. Additionally, some providers are not familiar with the basic QMB program rules.

Roberts et al. (2021) is one of the only papers that uses quasi-experimental methods to estimate the effect of QMB enrollment on health care spending. Using a regression discontinuity design with income as the assignment variable, they find QMB enrollment reduces out-of-pocket spending and increases evaluation and management visits. However, because beneficiary covariates and eligibility for other safety-net programs also change discontinuously around the QMB income cutoff, it is difficult to attribute their result solely to the QMB program. In contrast, my ability to follow a balanced panel of beneficiaries across states over time provides a cleaner setting for causal inference.

Outside of the dual-eligibles context, these results build on a large literature studying how consumer and provider prices impact utilization. On the consumer side, researchers have used variation in cost-sharing to estimate demand elasticities for a number of health care services. Brot-Goldberg et al. (2017) explore the impact of high-deductible plans on utilization. They find that switching to a plan with a $4,000 deductible leads spending to decline by about 12%. They show the entire effect was due to declines in quantities as opposed to prices. Similarly, Chandra, Gruber and McKnight (2010) studies a change in a Medicare supplement plan that raised patient cost sharing for physician visits and prescription drugs. They show that utilization of both services declined with elasticities similar to previous studies. Kaestner and Sasso (2015) also find that outpatient spending is causally associated with more hospital admissions.

On the provider side, Clemens and Gottlieb (2014) exploit a change in Medicare payments that varies across regions to measure how much the quantity of care supplied varies with respect to payment rates. They estimate a long-run average elasticity of supply and show that treatment patterns adjust along several margins. In the context of Medicaid, there is evidence that higher reimbursement rates leads to improved educational outcomes among children (Alexander and Schnell, 2019), significant improvements in the quality of nursing care (Hackmann, 2019), and higher Medicaid enrollment (Chen, 2014). My setting is unique because the shocks to consumer and supplier prices occurs simultaneously. This accounts for potential interaction effects, and has direct relevance for policy reforms which pay for expanded consumer access by cutting provider prices.[2]

2.2 Background

Medicare is a federal program that provides health insurance to the elderly and disabled. Nearly all adults over the age of 65 as well as Social Security disability insurance beneficiaries are eligible. About 70% of individuals enroll in traditional Fee-For-Service (FFS) Medicare. FFS Medicare consists of Part A which covers inpatient hospital services, skilled nursing,

[2]See Shepard, Baicker and Skinner (2020) for a discussion of some recent policy proposals.

hospice, and home health; Part B which covers physician, outpatient, and preventive services; and Part D which covers outpatient prescription drug benefits.[3] Like private sector health plans, Medicare beneficiaries pay monthly premiums and are responsible for a portion of their medical expenditures through cost-sharing provisions. For all Part B services in 2009, these provisions included a $96 monthly premium, a $135 deductible, and 20% co-insurance. Unlike private sector health plans, there is no cap on out-of-pocket spending.[4] Although many beneficiaries purchase supplemental insurance plans (Medigap) to cover some or all of the cost-sharing, these plans are not affordable for the poorest households. Instead, beneficiaries near the poverty line can have their premiums and/or cost-sharing paid through either the Qualified Medicare Beneficiary (QMB) program or the Specified Low-Income Medicare Beneficiary (SLMB) program.

Both QMB and SLMB are administered by states with minimum eligibility standards set by the federal government. States have the option to expand eligibility, but most set it at the federal minimum. QMB covers all Medicare premiums, deductibles and co-insurance for individuals below 100% of the federal poverty line (FPL). SLMB covers only Medicare premiums for individuals below 135% of FPL.[5] Both programs impose a cap on liquid assets. In 2009 this amount was set at $4,000 for individuals and $6,000 for couples. Beneficiaries in either program also receive subsidized Part D prescription coverage through the low-income subsidies (LIS) or "Extra Help" program. Although many QMB/SLMB beneficiaries are also enrolled in Medicaid to receive nursing care, Medicaid eligibility is neither necessary nor

[3] Alternatively, roughly 30% of beneficiaries enroll in a privately run Medicare Advantage plan. These plans offer lower out-of-pocket costs by restricting access to a narrower network of providers. Because I do not observe utilization for Medicare Advantage beneficiaries, I exclude them from the analysis.

[4] Part A coverage generally requires no premium, but has a $1,068 hospital stay deductible. Effective in January 2011, the Affordable Care Act removed cost-sharing for a small number of Part B preventive services such as vaccinations and cancer screenings.

[5] Technically, there are two separate programs which cover only Medicare premiums: SLMB which is for incomes below 120% of the FPL, and the Qualifying Individual (QI) program for incomes below 135% of the FPL. These programs have separate names because the federal funding mechanisms differ slightly, but from the perspective of beneficiaries they provide identical benefits.

sufficient to receiving QMB/SLMB benefits.[6]

Compared to ordinary FFS patients, providers have less incentive to treat QMB patients since their effective reimbursement rate is lower. For QMB payments, states have the choice to pay either the full Medicare cost-sharing, or the amount by which the Medicaid rate for the same service exceeds what Medicare has already paid. The majority of states, including Connecticut, pay the lesser of these two amounts.[7] Since Medicaid rates are usually more than 20% less than Medicare rates, providers may not receive any additional payment. By law, providers may not attempt to collect any cost-sharing from QMB enrollees.

Prior work has shown that take-up rates for the QMB and SLMB program are low. Using the Survey of Income and Program Participation and linked administrative data, Caswell and Waidmann (2017) estimate only 53% of eligible beneficiaries are enrolled in the QMB program. This drops to between 32% and 15% for the SLMB program. Although some beneficiaries may rationally avoid the QMB program because of concerns about provider access, there is no reason to avoid the free premiums in the SLMB program. Lack of awareness, confusing applications, recertification requirements, and stigma are all potential factors for low participation rates.

In 2009, Connecticut implemented a major change in QMB and SLMB eligibility. In an effort to reduce spending on an existing state-run prescription drug program, Connecticut raised QMB eligibility from 100% to 207% of FPL and raised SLMB eligibility from 135% to 242% of FPL. They also removed the asset test from both programs. The change in policy affected program eligibility in three different ways: (i) existing SLMB beneficiaries with income from 100% to 135% FPL were auto-enrolled in the QMB program; (ii) previously unsubsidized beneficiaries between 135% and 207% FPL became newly eligible for QMB benefits; and (iii) previously unsubsidized beneficiaries between 207% and 242% FPL became

[6]For example, beneficiaries who qualify for Medicaid through the "medically needy" pathway are not generally eligible for QMB/SLMB benefits.

[7]See Table 4-4 of MACPAC Report to the Congress on Medicaid and CHIP, March 2013.

newly eligible for SLMB benefits.

2.3 Data

2.3.1 Restricted Access Medicare Administrative Files

My primary dataset is derived from the 100% full population panel of Medicare administrative records from 2006 to 2011. The file records basic demographic data such as date of birth, sex, race, and state of residence. On the enrollment side, I observe months of coverage in Part A, Part B, Part D, Medicare Advantage, and a detailed dual-status code. I define a beneficiary as enrolled in QMB if they are coded as QMB-only or QMB with Medicaid benefits. SLMB enrollment is defined as SLMB-only, SLMB with Medicaid, and Qualifying Individuals. I also observe if beneficiaries received their original Medicare entitlement through the Social Security disability program. This provides a useful measure of lifetime health outcomes.[8] A key limitation to highlight is that income is unobserved. This makes it impossible to identify everyone who is eligible for the QMB program.

For each Fee-For-Service (FFS) beneficiary I observe utilization for 11 categories of service (e.g. inpatient hospitalization, evaluation and management, imaging) aggregated over the calendar year. Each category records Medicare payments to providers, the corresponding amount of cost-sharing charged, and number of claims. Because enrollment in Medicare supplement plans is unobserved, it is not possible to differentiate cost-sharing paid directly by the beneficiary versus cost-sharing paid by a supplement insurer. I set cost-sharing liability for QMB beneficiaries to zero. I assume Connecticut's Medicaid program rejects any Medicare claims under the "lesser-of" payment rule and compute provider revenue as the sum of direct Medicare payments and the cost-sharing liability from non-QMB beneficiaries.[9]

[8] The history of SSI enrollment is also unobserved.

[9] Using linked Medicare and Medicaid claims Zheng, Hoover and Feng (2014) finds that for office evaluation and management visit in Connecticut in 2009, only 11% of of Medicare cost-sharing is paid for by Medicaid.

The main utilization outcome is total number of Part B outpatient claims. This follows prior literature which focuses on outpatient services because it is the main setting for preventative care. Additionally, rejected Medicaid claims for Medicare inpatient cost-sharing are sometimes recoverable through the Medicare bad debt program. Although this imposes administrative costs on providers, the bad debt program reduces the disincentive to treat QMBs as inpatients.

2.3.2 Sample Construction and Summary Statistics

I consider two separate samples. The first sample includes all Medicare beneficiaries regardless of dual-status. This allows me to estimate a policy-relevant treatment effect which accounts for endogenous take-up and any spillover effects across beneficiaries. Because income is unobserved, considering the whole universe is the only way to ensure all eligible beneficiaries are included. However, this broad coverage may come at the expense of precision. The treatment and control groups include beneficiaries across the income distribution. Because only a small fraction of the population will be eligible, variation among the ineligible population could make it difficult to measure a causal effect.

My second sample includes only Medicare beneficiaries who were enrolled in the SLMB program in the month prior to the policy change. After the policy change in Connecticut, these beneficiaries were newly eligible for QMB benefits and the state auto-enrolled them in the program. An advantage to using this narrow sample is the ability to measure precise effects. Take-up is mechanically near 100% and the incomes of the treatment and control groups fall within a narrow range of 100-135% FPL. Conversely, a disadvantage of this sample is the estimated treatment effect is less policy-relevant. There is no general mechanism to auto-enroll newly-eligible QMB beneficiaries and a large fraction of beneficiaries above 135% FPL are excluded from the analysis.

This is the lowest of any state in their sample suggesting the difference between Connecticut Medicare and Medicaid reimbursements is large.

In both samples I impose several restrictions to work with a balanced panel. Because the Medicare data only record claims from FFS beneficiaries, I exclude any beneficiary ever enrolled in Medicare Advantage. I also require beneficiaries to be continuously alive and enrolled in Parts A and B.[10] Table 2.1 compares the summary statistics of these samples between Connecticut and all other states. In both samples, Connecticut is more white and somewhat older than the rest of the United States. For the main utilization outcomes, cost-sharing liability is comparable to other states, but they are more intensive users of outpatient services.

2.4 Empirical Strategy

My empirical strategy relies on a standard two way fixed effects difference-in-difference model

$$y_{it} = \delta D_{it} + \alpha_s + \gamma_t + e_{it} \tag{2.1}$$

where y_{it} denotes an individual i outcome in a given year t, α_s is a state fixed effect, γ_t is a year fixed effect, and D_{it} is an interaction term for Connecticut during or after 2009. Because I worked with a balanced panel, all demographic characteristics are absorbed into the state fixed effect. The δ coefficient will identify a treatment effect if, in the absence of the policy change, the trends in outcomes in Connecticut would have remained parallel to the trends in outcomes in the control states. To assess the plausibility of this assumption, I estimate coefficients from an event study design

$$y_{it} = \sum_{y=2006, y\neq 2008}^{2011} \delta_t (S_i = CT) * (y = t) + \alpha_s + \gamma_t + e_{it} \tag{2.2}$$

[10]These restrictions create a sample that is likely to be healthier and wealthier than the overall Medicare population. On the demand side better health implies more sensitivity to cost-sharing, but more financial resources implies less sensitivity. On the supply side, provider elasticities may be higher for sicker patients with more complex conditions. I remove this restriction when I consider mortality as an outcome variable.

where each δ_t represents the difference in y_{it} between Connecticut and control states relative to 2008, the year before the expansion.

Difference-and-difference designs with a single treated unit lead to challenging inference problems. The most popular way to compute standard errors in difference-in-difference settings where the treatment unit is an entire state is to cluster errors at the state level. This adjusts the usual variance-covariance matrix estimate by accounting for errors that may be correlated for all individuals within a state. However, this formula relies on an asymptotic approximation that assumes the number of treated units is large. Conley and Taber (2011) show this can lead to test statistics that over-reject the null when the number of treated units is small. Although wild cluster bootstraps can solve the problem in many cases, inference will still be misleading when there is a single treated cluster (MacKinnon and Webb, 2017).

To correct for small cluster bias I compute confidence intervals using a permutation test. Specifically, I drop Connecticut from the data and re-estimate equation (2) several times with every other remaining state defined as the "treated" state. I then compare the distribution of these placebo coefficients from the correct version of equation (2) where Connecticut is included and defined as the treated state. To be significant at the 5% level under a two-tailed test, Connecticut would have to rank at the top or bottom of the placebo distribution.

As an additional check, I also consider the inference method described by Ferman and Pinto (2019). Their method is appropriate for difference-in-difference settings with a single treated unit. The procedure they propose is a cluster residual bootstrap that corrects for the heteroskedasticity that arises from different underlying sample sizes within a cluster. Specifically, the technique rescales the residuals to remove heteroskedasticity. Next, it resamples from linear combinations of each state's adjusted residuals such that the linear combination is the within-state difference in outcomes. This provides a bootstrapped distribution of placebo treatment effects that can be used to estimate an exact confidence interval.

2.5 Results

I present results on four key outcomes: enrollment, number of outpatient claims, cost-sharing liability, and provider revenue. To provide an understanding of how the QMB program affected all eligible beneficiaries, I start by considering the full sample. Next, to provide more precision I consider the SLMB-only sample which was auto-enrolled in the first month of the policy change.

2.5.1 Full Sample Utilization and Spending

Figure 2.1 plots the event study coefficients from equation (2). To provide a baseline for inference, I also plot the 95% interval of placebo coefficients from the permutation test. The pre-trends are parallel for all four outcomes which suggests the main causal identification assumption is reasonable.

The top left panel shows that raising the income cutoffs for the program successfully increased enrollment. The expansion occurred in October of 2009 and enrollment increased gradually for the two following years. By the end of 2011, enrollment doubled from 10% of the Medicare population to 20%. The top right panel shows this expansion had no discernible effect on outpatient utilization. Across all FFS beneficiaries, the confidence interval for total number of Part B claims ranges from than -7% to 4%.

The bottom left panel shows substantial declines in cost-sharing liability for Part B services. The result follows directly from increased enrollment. After a modest Part B deductible, cost-sharing is almost always a uniform 20% co-insurance of the Medicare reimbursement rate. Because QMBs are exempted from these costs, greater enrollment implies less cost-sharing. The difference-in-difference coefficient for enrollment is 5.8 percentage points and the coefficient for cost-sharing is roughly -75 dollars. These estimates imply enrolling in the program QMB reduces cost-sharing liability by $1,288.

The final panel in the bottom right shows provider revenue for Part B services per beneficiary. This is the full 80% rate paid directly by Medicare plus cost-sharing for non-QMBs. There is some evidence of a decline in overall revenue, but it is not significantly different from the underlying variation in placebo states. The point estimate suggests a 2.5% decline or revenue ($67 per enrolled beneficiary). Despite the expansion, cost-sharing from QMBs is a small fraction of total payments.

Figure 2.2 compares confidence intervals for the difference-in-difference coefficient. Both methods rely on bootstrapping, so the intervals are not centered around zero. The Ferman-Pinto method occasionally produces more precise intervals, although in many cases the differences are modest. Regardless of the method used, the results show a sharp increase in enrollment, no change in utilization, and a decline in consumer cost-sharing liability.

2.5.2 SLMB Only Sample Utilization and Spending

Figure 2.3 plots the same event studies for the SLMB-only auto-enrolled sample. The results are broadly similar and somewhat more precise. QMB enrollment is 25% in 2009 because the policy was implemented for 3 months of the year. In 2010, enrollment hit nearly 100%. The point estimate for the change in utilization is almost exactly 0 with a confidence interval ranging from -6% to 5%. In contrast to the result for the full sample, provider revenue now shows a sharp decline close to 20%. Again, Figure 2.4 shows the results are not sensitive to the choice of inference method. Dividing the change in cost-sharing liability by the change in enrollment implies the program QMB reduces cost-sharing liability by $1,328—nearly identical to the full sample estimate.

2.5.3 Mortality

To understand how the policy change affected welfare, I consider an unambiguous health outcome that is well-measured in my data: mortality. To measure the probability of death, I expand my baseline samples by removing the requirement that beneficiaries remain continuously alive. The average annual mortality rate during the pre-period is roughly 4% for the full sample and 7% for the SLMB-only sample. Figure 2.5 plots the event studies and does not show evidence of an effect. For the full sample, the parallel trends assumption may not hold which makes it difficult to reach any conclusion. For the SLMB-only sample there is a small drop in mortality, but the effect is not precise. The results suggest that any changes to unobserved quality of care do not lead to large changes in health outcomes.

2.6 Discussion

Standard models of competitive markets provide limited value in interpreting the result. Because prices for Medicare covered services are set by the government, any adjustment must happen on the quantity margin. Furthermore, it is unclear at any given time whether the quantity of Medicare services are in excess supply, excess demand, or equilibrium. An alternative approach is to compare the observed point estimate to a predicted estimate based on demand and supply elasticities from existing literature.

For the elasticity of demand, the most widely used estimates come from the RAND Health Insurance Experiment. Aron-Dine, Einav and Finkelstein (2013) review the original experimental data and show that elasticities vary significantly depending upon which type of insurance plans are compared. Conveniently, they report an elasticity of -0.18 comparing an insurance plan with 25% co-insurance to one with 0% co-insurance—nearly exactly the QMB treatment.[11] Thus, because the price cut is 100%, the predicted change in utilization from the demand side is 18%.

For the elasticity of supply, Clemens and Gottlieb (2014) provide an estimate relevant for my setting. They study a policy change in 1997 which redefined the geographic regions Medicare used to adjust the reimbursement formula for differences in the local price level. Using this quasi-experimental variation, they estimate a short-run elasticity of about 0.8. In

[11] Chandra, Gruber and McKnight (2010) finds a similar estimate in the context of Medicare beneficiaries.

my context provider prices are cut by 20%, thus the predicted change in utilization from the supply side is -16%. On net, these two elasticities predict a 2% increase in utilization—a value I can not reject is different from my point estimate.

A key limitation to highlight is that any spillover effects on private payers are unobserved.[12] One possibility is that providers engage in "cost-shifting" and increase the rate they demand from private insurers to offset lower Medicare rates (Frakt, 2011). Alternatively, a bargaining framework would suggest that by changing physicians outside options, lower Medicare rates allows insurers to demand lower private rates (Clemens and Gottlieb, 2017). Although data limitations prevent any analysis of private payers, my results do not provide evidence the expansion had spillover effects on non-QMB Medicare patients. Any spillover would imply the full sample and SLMB-only sample would generate different utilization results. Instead, I find no effect for both samples.

2.7 Robustness Checks

I consider two robustness checks to evaluate how different estimation techniques may affect my result. First, I use the synthetic control method (Abadie, Diamond and Hainmueller, 2010). To construct the synthetic untreated state I use controls for race, gender, age, and disability history. To provide a source of inference, I rerun the synthetic control estimation individually for all the placebo states and plot the difference between the synthetic control and treated. Using this method for the SLMB-only sample, Figure 2.6 shows a sharp clear increase in enrollment. Figure 2.7 shows no clear change in utilization, though the inference is imprecise.

The synthetic control method may not preform well in my setting because I have a short panel with only 3 years of pre-period data. Most applications of the method have many

[12]For context, in Connecticut Medicare covers 16% of the population and accounts for 20% of the spending on physician and clinical services. CMS, Health Expenditures by State of Residence.

more observations prior to treatment which can help improve the fit of the control group. Relatedly, because all of my state controls are fixed over time, it is difficult to construct a control group that has similar outcomes in levels.

Another estimation technique I explore is difference-in-difference with exact matching. Within each sample, I take every Medicare beneficiary in Connecticut and find a beneficiary in another state with the same gender, race, disability history, and age. I use the algorithm developed by Stepner and Garland (2017) to implement a one-to-one caliper matching without replacement where the caliper width for age is 180 days from the exact date of birth. This produces equal sized treatment and control samples with identical covariates which can be run through the difference-in-difference specification. The same inference problems apply in this setting because the single Connecticut cluster may have correlated errors. To account for this, I implement the same matching procedure for each of the placebo states and plot the distribution of coefficients. Figures 2.8 and 2.9 shows the exact matching estimation for both samples does little to improve precision.

2.8 Conclusion

The purpose of the QMB program is to provide low-income Medicare beneficiaries with greater financial protection and expanded access to care. My results suggest it is partially successful at this goal. On the one hand, the program sharply reduces cost-sharing liability. Beneficiaries save an average of \$1,300 in Part B cost-sharing which is between 5% and 10% of their income. Total savings would be even greater if the exemptions from Part B premiums, Part A cost-sharing, and Part D cost-sharing were included. Although the out-of-pocket savings may be lower for beneficiaries with supplemental coverage, this would be offset by higher premium payments. Robst (2006) shows that premiums for supplement plans are close to actuarially fair. This suggests the \$1,300 figure is a reasonable estimate of population average savings if premium expenses are also included.

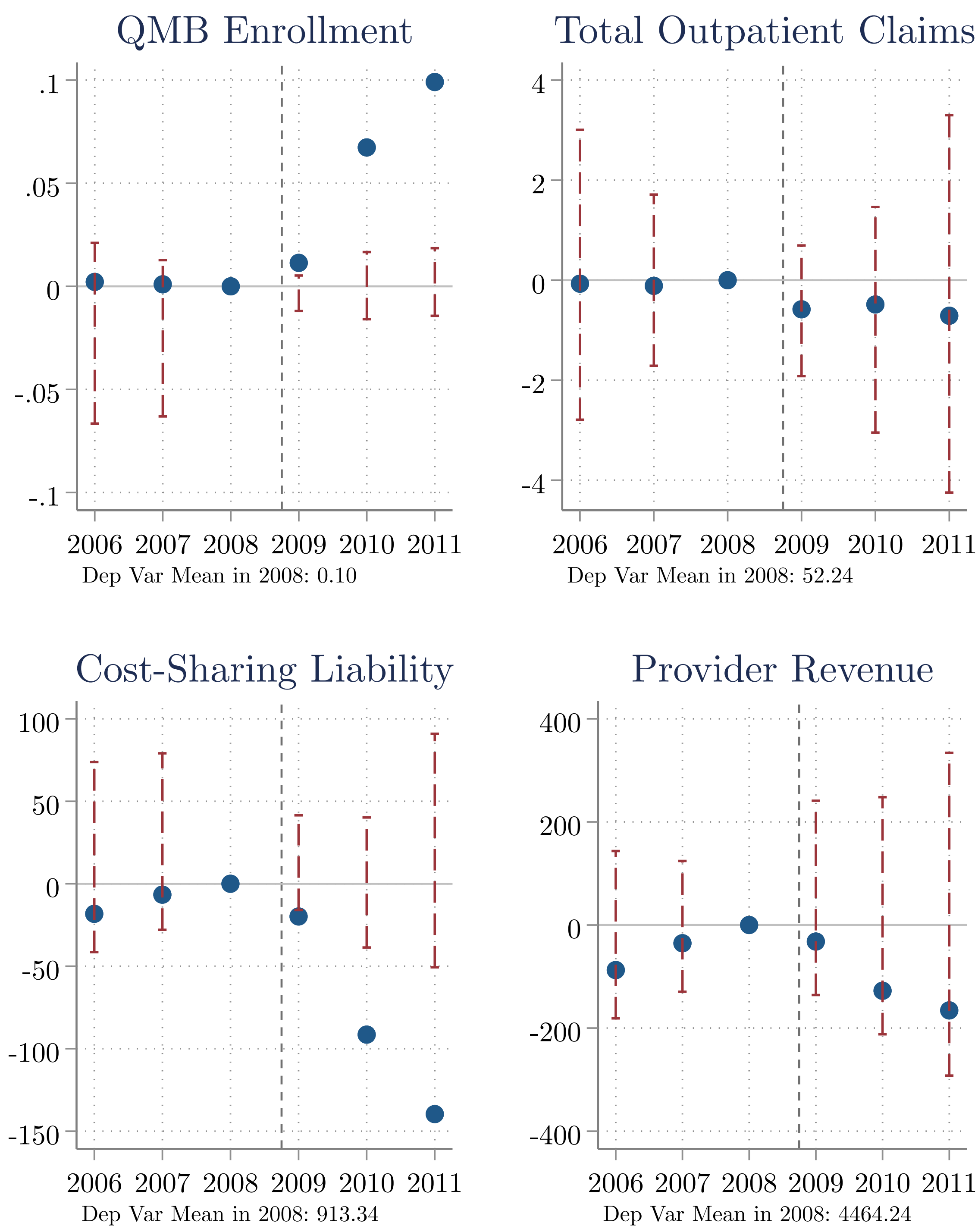
On the other hand, it is not clear the program improves access to care or health outcomes. If beneficiaries avoided necessary outpatient care prior to enrollment, removing cost-sharing obligations was not sufficient to increase the consumption of health care services. The fact that utilization and mortality are both unchanged is consistent with the supply and demand elasticities from prior work. Lower provider reimbursement from the "lesser-of" payment rule appears to offset any improvements in consumer access on the demand side. Using detailed claims data to measure how the effect differs across providers would offer insight into the underlying mechanism.

Table 2.1: Summary Statistics

	All Beneficiaries		SLMB Only	
	Connecticut	Rest of U.S.	Connecticut	Rest of U.S.
White	0.89	0.83	0.84	0.76
Black	0.06	0.08	0.09	0.16
Female	0.59	0.57	0.66	0.59
Ever Disabled	0.19	0.23	0.48	0.49
Age	76.8	75.3	72.3	68.9
Outpatient Claims	49.17	46.7	56	51
Cost-Sharing Liability	954	920	1087	1064
Share QMB	0.10	0.12	0	0
Observations	267,250	18,678,756	8,632	476,029

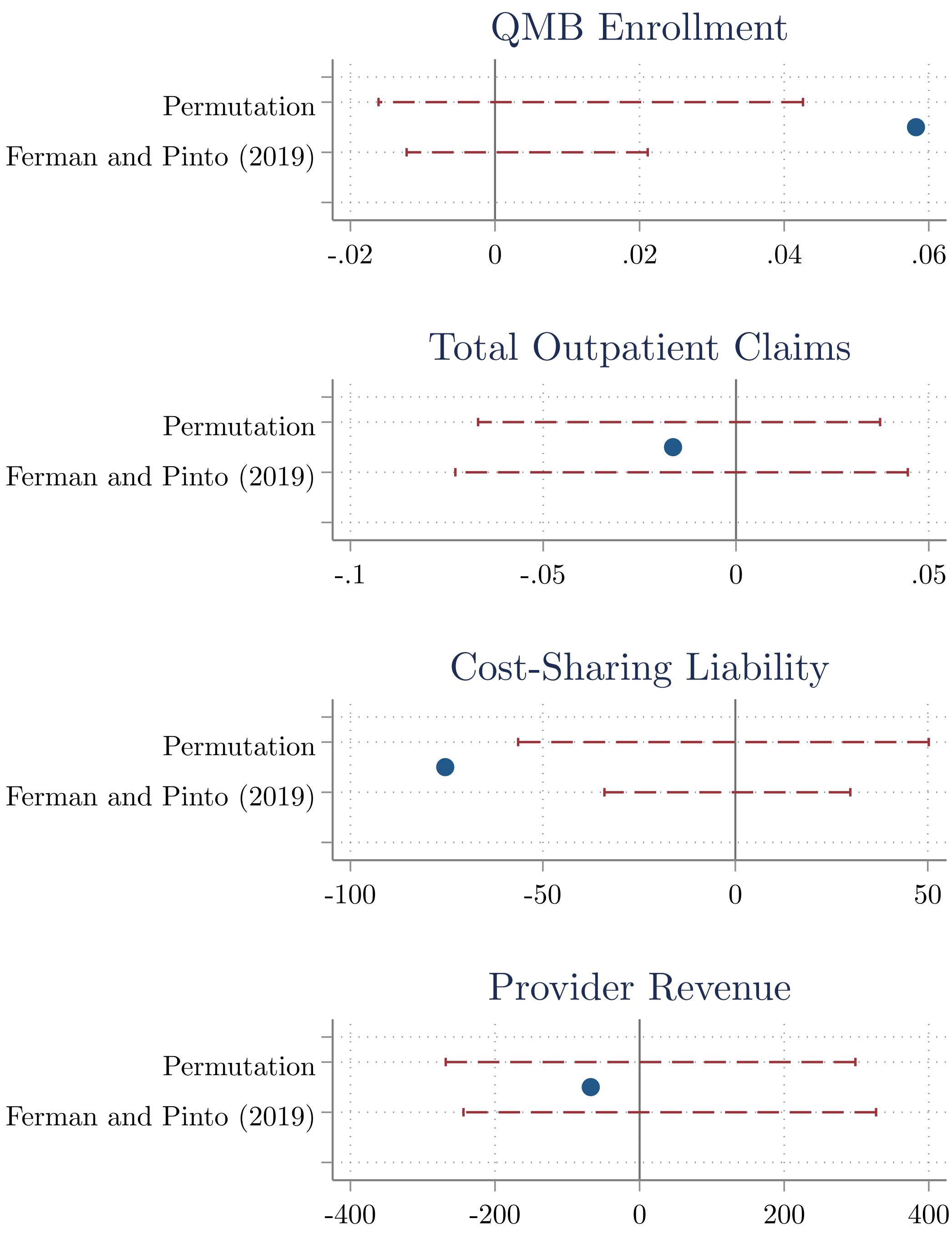
Notes: Sample is balanced panel of continuously alive, Part A and B FFS beneficiaries from 2006 to 2011 in Master Beneficiary Summary File. "SLMB Only" sample includes any beneficiary enrolled in the SLMB or QI program in September 2009. The means between Connecticut and the Rest of the U.S. are significantly different at the 5% level for all variables.

Figure 2.1: Event Studies for Full Sample



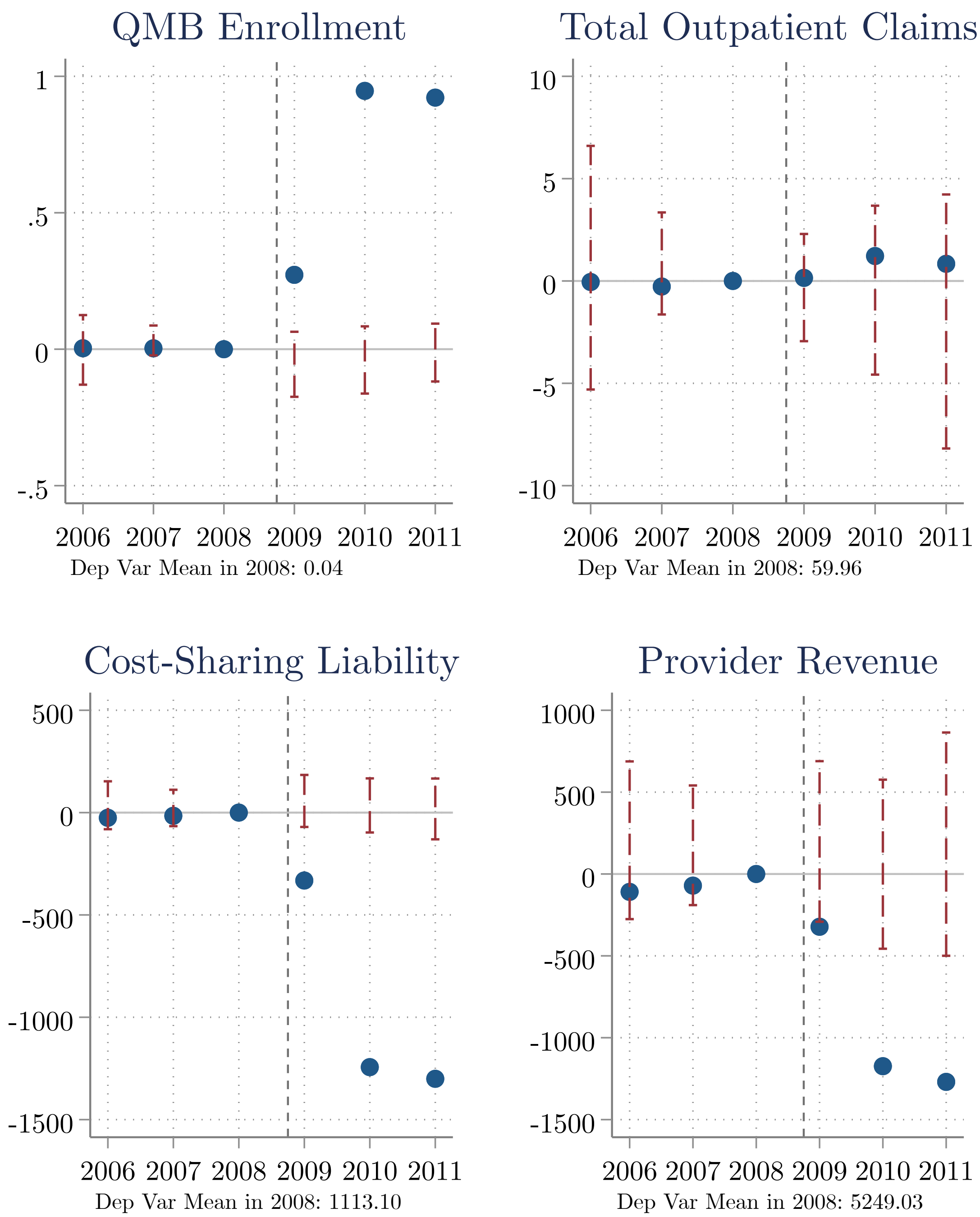
Notes: Data is a balanced panel of all FFS beneficiaries. Dashed red lines are show the distribution of placebo coefficients from permutation tests.

Figure 2.2: Full Sample Diff-in-Diff Treatment Coefficient with Inference



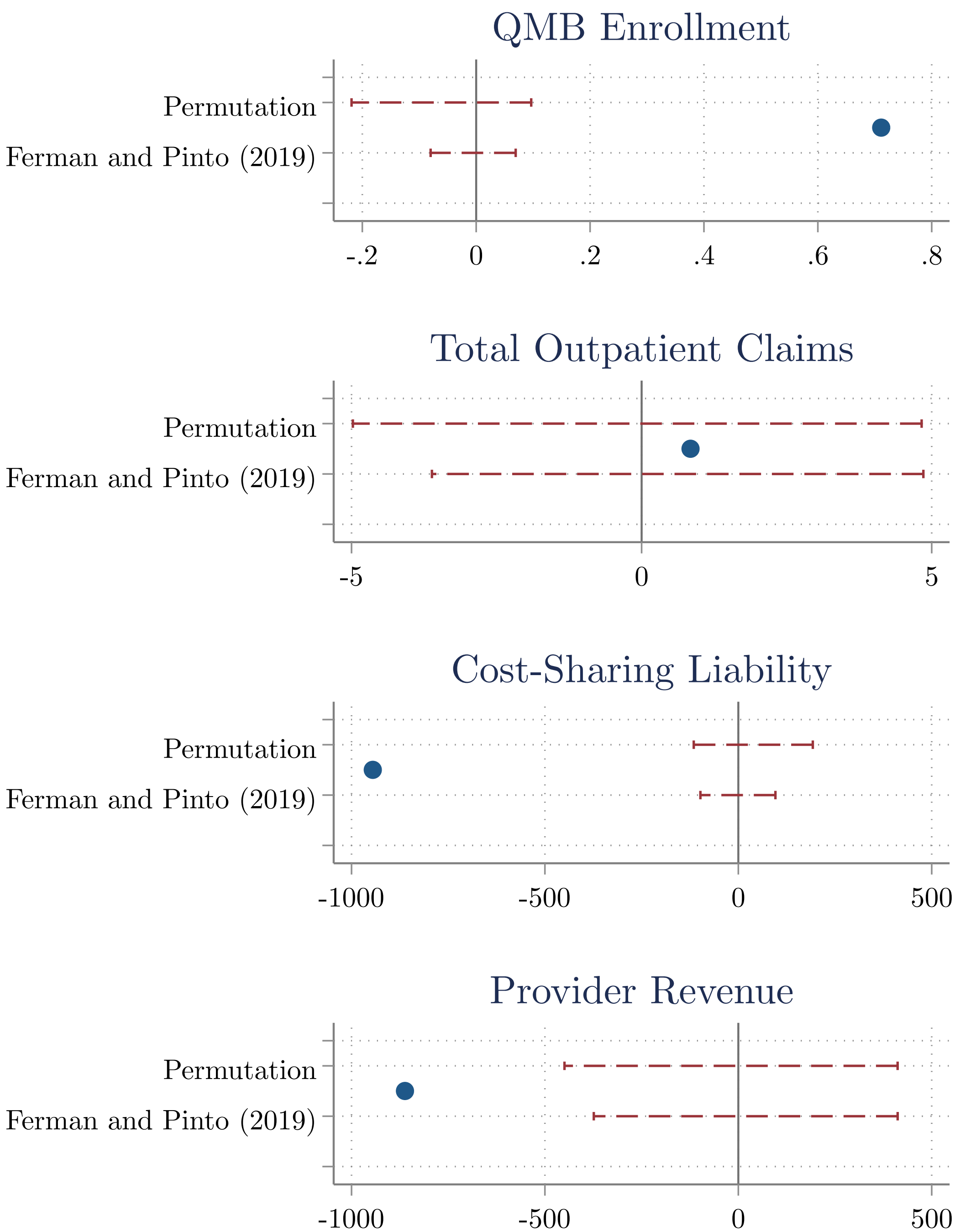
Notes: The blue dot is the interaction point estimate from the difference-in-difference specification and the dashed red liens compare 95% confidence intervals.

Figure 2.3: Event Studies for SLMB Only Sample



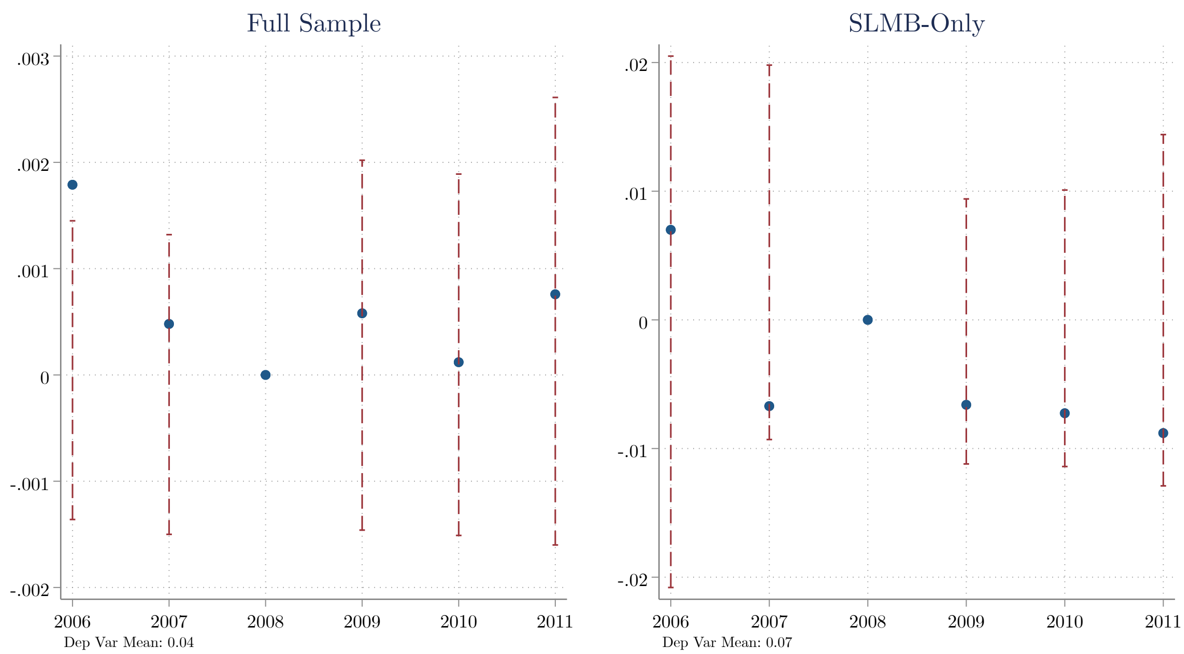
Notes: Data is a balanced panel of all beneficiaries in the SLMB program in October 2009. Dashed red lines are show the distribution of placebo coefficients from permutation tests.

Figure 2.4: SLMB Sample Diff-in-Diff Treatment Coefficient with Inference



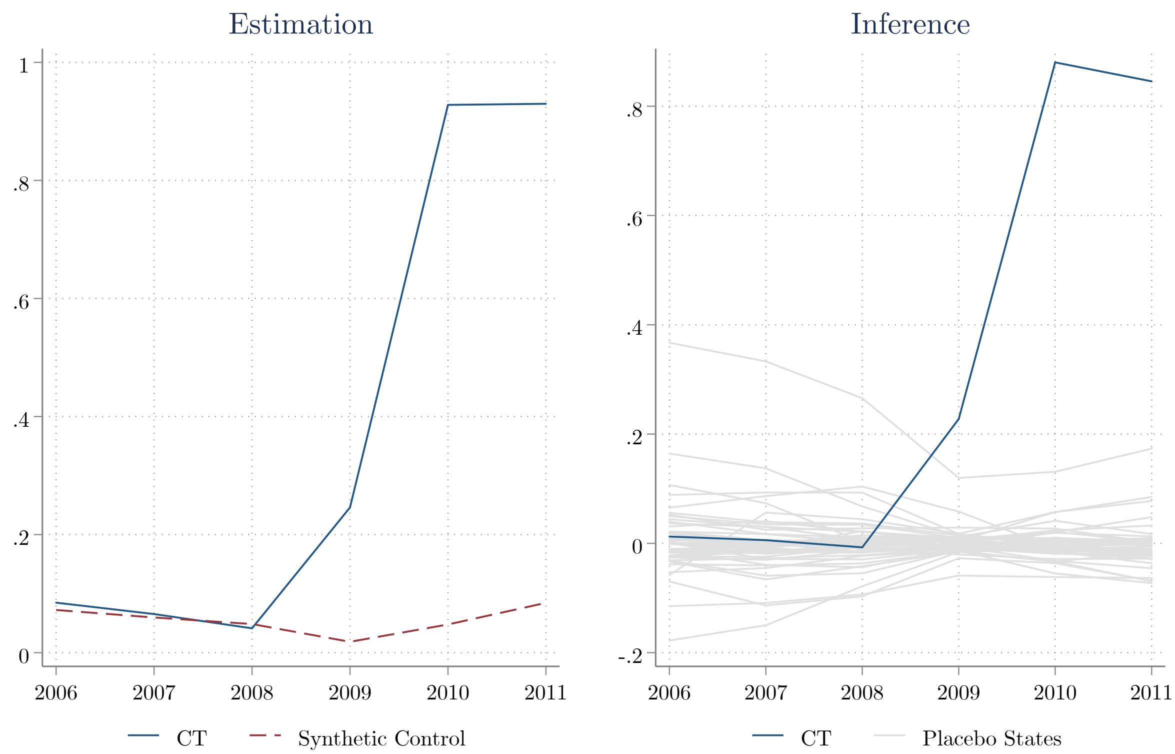
Notes: The blue dot is the interaction point estimate from the difference-in-difference specification and the dashed red lines compare 95% confidence intervals.

Figure 2.5: Event Studies for Mortality



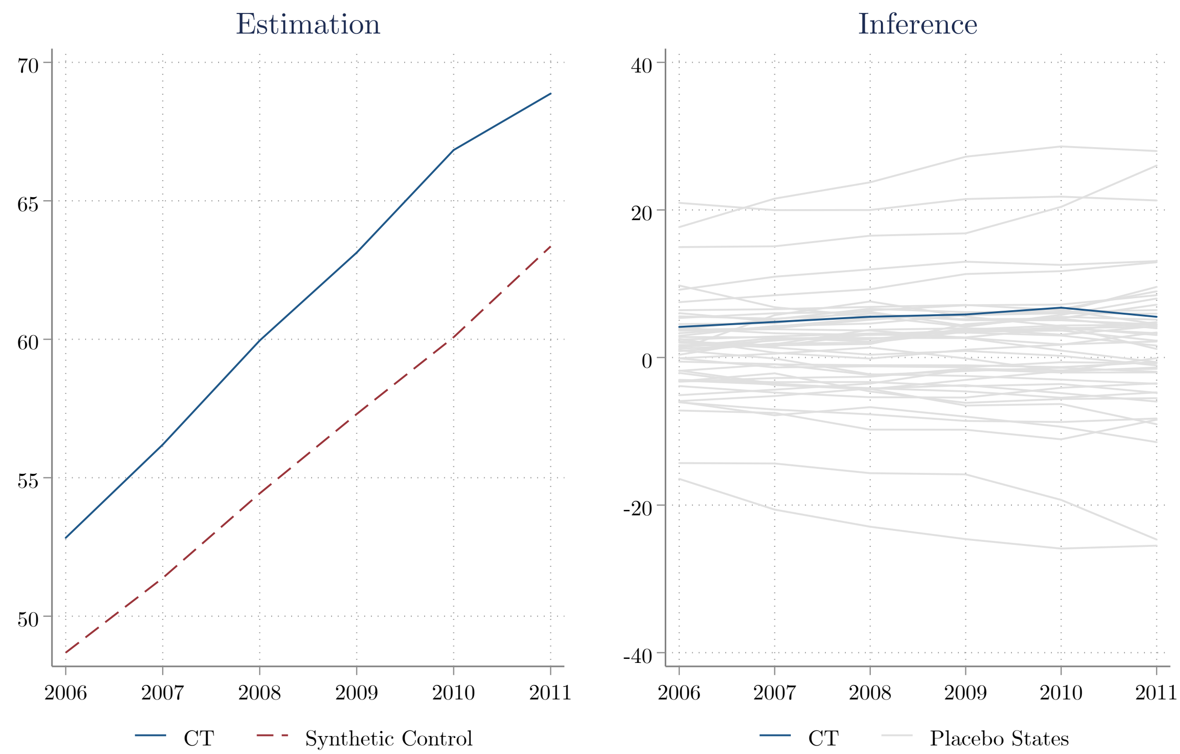
Notes: Sample includes both FFS and Medicare Advantage beneficiaries. Dashed red lines are show the distribution of placebo coefficients from permutation tests.

Figure 2.6: SLMB Sample Enrollment Synthetic Control



Notes: The left panel depicts the difference in outcomes between the treated and synthetic control state. Faded gray lines denote the placebo states.

Figure 2.7: SLMB Sample Claims Synthetic Control

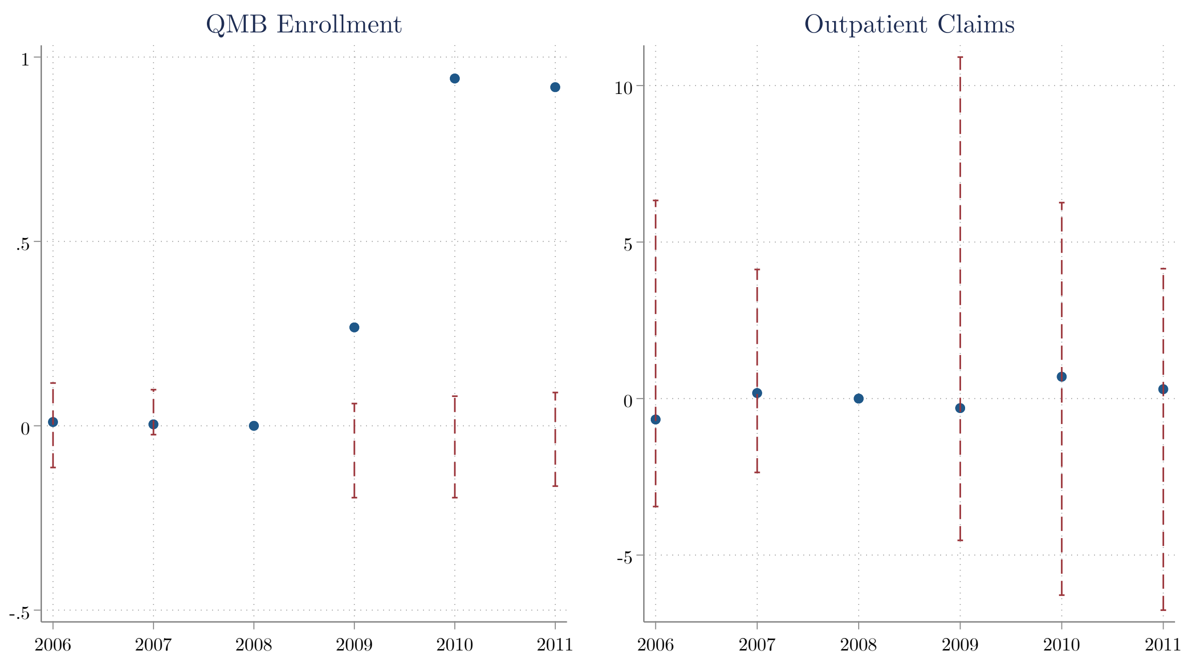


Notes: The left panel depicts the difference in outcomes between the treated and synthetic control state. Faded gray lines denote the placebo states.

Figure 2.8: Full Sample Exact Matching

Notes: Difference-difference event study using exact one-to-one matching on gender, race, disability history, and age. Confidence intervals are computed using distribution of coefficients from placebo matched states.

Figure 2.9: SLMB-Only Sample Exact Matching



Notes: Difference-difference event study using exact one-to-one matching on gender, race, disability history, and age. Confidence intervals are computed using distribution of coefficients from placebo matched states.

CHAPTER 3

Substitution Between Social Security Retirement and Disability Benefits

3.1 Introduction

A widely accepted view in the literature on Social Security Disability Insurance (SSDI) is that relative benefit generosity is a key determinant of program participation. Theory suggests a worker's decision to apply for SSDI depends on their potential income from labor market earnings or other social insurance programs. To understand the extent of moral hazard in SSDI, it is necessary to measure how a worker's outside option affects enrollment.

In this paper, I study how cuts to Social Security retirement benefits affect Social Security disability enrollment. Specifically, I exploit a policy change which caused an abrupt decline in the generosity of retirement benefits for workers near certain date of birth cutoffs. Assuming workers with birth dates near these cutoffs are similar along other dimensions, differences in benefit generosity are as good as random. To measure disability rates, I use the universe of Medicare enrollment data for selected years between 1999 and 2017. Because nearly everyone who lives to 65 will receive Medicare through either age or disability, this provides an ideal setting to precisely measure disability rates by cohort.

Using various regression discontinuity designs, I do not find cohorts with lower retirement benefits have higher rates of disability benefit receipt. My preferred specification estimates a precise null effect. My confidence interval implies the increase in the full retirement age caused disability enrollment to only change between -0.25 and 0.28 percentage points. The results provide evidence that disability enrollment is not affected by modest changes in relative benefit generosity. I find there are fewer marginal SSDI applicants implying the program's moral hazard effects are more modest than prior work would suggest.

This project is closely related to Duggan, Singleton and Song (2007, hereafter DSS). They study the same question and policy change using a different dataset and identification strategy. They estimate a specification that uses variation in the present value of Social Security benefits by age and year-of-birth. To compute enrollment rates, they use a 10% sample of the Master Beneficiary Record (MBR) and single year population estimates from the Census Bureau. Their results imply the policy change caused an additional 0.6% of men and 0.9% of women between the ages of 45 and 64 to receive disability benefits.

More broadly, this paper contributes to a growing literature studying spillover effects from disability insurance to other social insurance programs. For example, Mueller, Rothstein and von Wachter (2016) explore whether the marginally disabled increase disability applications when unemployment insurance becomes less generous. By exploiting variation in unemployment benefit duration across states, they find no evidence of substitution across programs. Similarly, Schmidt, Shore-Sheppard and Watson (2020) find expanded access to Medicaid under the Affordable Care Act had no impact on disability applications. These results are consistent with my finding that any spillover effects are likely to be modest.

This project also contributes to research on the determinants of growing SSDI enrollment.[1] Several papers have studied how changes in the demographic characteristics of the insured population affect disability receipt. Liebman (2015) finds that increasing coverage rates from greater female labor force participation as well as population aging account for about half of the growth in enrollment since 1984. Other work has explored whether the underlying incidence or severity of disability has increased. Autor and Duggan (2010) do not find evidence for this theory and show that population health is improving over the relevant

[1] See Maestas (2019) for a review of the most recent research.

period. Alternatively, some researchers have focused on changing features of the SSDI program. Because the benefit formula is linked to average rather than median wages, growing wage inequality has led SSDI replacement rates to gradually increase. Duggan and Imberman (2009) estimate this change in benefit generosity can account for 24% of the growth in male DI receipt.

3.2 Institutional Setting

Social Security is a federal social insurance program that provides monthly cash benefits to workers and their families in the event of retirement, disability, or death. Workers pay into the program through mandatory payroll deductions and receive benefits as a function of their contributions. To qualify for benefits workers must contribute a minimum amount through payroll taxes on Social Security covered employment. To be insured for disability, workers must have covered employment for about a quarter of a worker's adult life and for at least five of the 10 years prior to the onset of disability. For retirement, 10 years of covered employment is sufficient. In the case of retirement, insured workers are eligible to claim benefits at any age between 62 and 70. In the case of disability, workers must demonstrate they have a "medically determinable physical or mental impairment" which prevents them from earning a sufficient amount of labor income. Known as substantial gainful activity (SGA), this earning threshold is $1,260 per month for 2020.[2]

Workers entitled to Social Security are also eligible for health insurance from Medicare. All retirement beneficiaries are eligible for Medicare at age 65 and those receiving benefits prior to 65 are auto-enrolled. Disability beneficiaries are also eligible for Medicare, but benefits do not start until 29 months after the disability onset date.

Social Security retirement and disability programs use a similar benefit formula. The

[2]SSDI should not be confused with Supplemental Security Income (SSI), a separate Social Security disability program for those with limited work histories.

Social Security Administration (SSA) observes a worker's real earnings history and computes their average index monthly earnings for the years prior to entitlement. Benefits are a concave function of average earnings so that replacement rates are progressive. For retirement beneficiaries, benefits are actuarially adjusted with reference to the full retirement age (FRA). For the 36 months preceding the FRA, benefits are reduced by five-ninths of 1% for each month of early claiming. For months earlier than 36 months before the FRA, the monthly rate of reduction is five-twelfths of 1%. To adjust for inflation, benefit amounts after entitlement are adjusted annually by the CPI.

In 1983, a worsening fiscal outlook prompted Congress to made several changes to the Social Security retirement program which cut benefits and increased taxes. The most salient change was a gradual increase in the FRA from 65 to 67. By changing the base age from which actuarial adjustments are computed, this reduced benefits for nearly every claiming age. For workers claiming at 62, the earliest eligibility age, the benefit reduction shifted from 20% to 30%. Table 1 summarizes the exact parameters for each cohort. In contrast, the generosity of disability benefits did not change. Because the changes in the FRA are computed based on a worker's calendar year of birth, a worker born in January of a given year has abruptly less generous retirement benefits than an identical worker born in December the year before. This discontinuous change in benefits by date of birth provides the quasi-experimental variation I need to estimate a causal effect.[3]

3.3 Data

3.3.1 Restricted Access Medicare Enrollment Data

My primary dataset is the 100% full population panel of Medicare enrollment data. Specifically, I use the Master Base Summary File for the following years: 1999, 2002, 2004, 2006 to

[3] See Chapter 1 for a description of other Social Security benefit discontinuities.

2011, 2014, and 2017. Every individual enrolled in Medicare at any point during these years is included. The file provides demographic data such as exact date of birth, sex, race, and most recent zipcode. The file also records whether a beneficiary received Medicare eligibility through age or disability, as well as the date their coverage was effective. I compute the date of the first disability onset using the coverage start date. Benefit amounts are not recorded in the Medicare data.

3.3.2 Public Use Disability Analysis File

For additional detail on disability beneficiary characteristics, I rely on the Disability Analysis File. This 10 percent random sample covers all beneficiaries who have received SSDI at any time between March 1996 and December 2018. It records information such as the date of application, benefit amount, category of disability, and outcomes of continuing disability reviews. Unlike the Medicare data, this file has the advantage that beneficiaries can be observed immediately once they qualify for SSDI benefits. It also includes beneficiaries who died before 1999. Because it is a public use file, date variables are only observable at the month level.

3.3.3 Sample Construction and Summary Statistics

The 100 percent Master Beneficiary Summary File from my selected years between 1999 and 2017 consists of about 98 million unique beneficiaries. I restrict the baseline estimation sample to individuals born between 1932 and 1949. This includes 6 cohorts before the policy change, 6 cohorts during the change, and 6 cohorts after the change. I also drop individuals born on January 1 of every year. This accounts for the fact the relevant discontinuity occurs at January 2 and that individuals with missing birth dates are sometimes coded as January 1 birth dates.[4]

[4]See Kopczuk and Song (2008) for more detail.

Table 3.2 reports summary statistics for both datasets. The estimation sample consists of 49,222,887 unique beneficiaries in the Medicare data and 757,288 in the Disability Analysis File. To visualize the long-run trends in the data, the top panel of Figure 3.1 shows rates of disability receipt by cohort and sex. The pattern matches prior work showing increased disability receipt over time and higher rates for men than women. However, unlike the data shown in DSS, there does not appear to be any break in the trend around the 1938 cohorts when the policy change was first phased in.[5] The bottom panel plots the change across cohort. There is not evidence disability rates in the affected 1938-1943 cohorts are growing more quickly than in unaffected cohorts.

Figure 3.2 plots disability rates for both sexes by birth date at the monthly level. The data reveal peculiar seasonal patterns. Disability rates increase sharply from November to March and then gradually decline through the rest of the year. The pattern is clearest prior to 1943 after which it moderates somewhat. Part of this can be explained by differential fertility patterns across seasons (Buckles and Hungerman, 2013). Despite this seasonality, there does not appear to be a discontinuous change in disability around the January cutoff for affected cohorts.[6]

3.4 Empirical Strategy

To estimate the causal effect of benefit generosity on disability receipt, I use a regression discontinuity design with date of birth as the assignment variable. Because my setting includes multiple treated and untreated cohorts, I consider various techniques for "stacking" the discontinuities to maximize statistical efficiency.

[5] In the the DSS data, the majority of the sample had not yet reached 65. This is the primary reason the trends differ, although measurement error in their population denominator may also be a factor.

[6] There is also seasonality in the date of disability onset, but this not affect my identification strategy.

3.4.1 Estimating Several Regression Discontinuities Separately

I first run the classic regression discontinuity design for each cohort separately. Following Gelman and Imbens (2019), I estimate a specification with varying linear trends in date of birth:

$$Y_i = \beta_0 + \beta_1 D_i + \beta_2 DOB_i + \beta_3 (DOB_i * D_i) + \beta_4 X_i + e_i \tag{3.1}$$

where i indexes date of birth, Y_i denotes the outcome of interest, DOB_i is a linear trend normalized as the distance in days from the cutoff, and D_i is a dummy for the January 2 cutoff. In some specifications, I also include controls X_i for gender and race. The coefficient of interest is β_1 which estimates the change in the mean of the outcome at the cohort boundary. In all cases, the unit of observation is an average collapsed within a date of birth cell and regressions are weighted by the number of individuals in each cell. This approach accounts for correlated shocks at the date of birth level and is consistent with prior work that estimates discontinuities around date of birth cutoffs (Gelber, Isen and Song, 2016).

The β_1 identifies a causal effect under two assumptions: (i) beneficiaries cannot precisely manipulate their date of birth around the cutoff, and (ii) no potential confounders are also changing discontinuously at the cutoff. If these and other standard OLS assumptions are satisfied, each β_1 is a normally distributed random variable. We can average the means and variances of these variables to compute a single coefficient and standard error for all six policy changes.

3.4.2 Stacked Regressions

An alternative technique for computing a single treatment effect is to "stack" each regression in a single specification. Specifically we can estimate:

$$Y_i = \beta_1 D_i + \sum_{c \in C} \beta_{2,c} DOB_i + \sum_{c \in C} \beta_{3,c} (DOB_i * D_i) + \beta_c + e_i \tag{3.2}$$

where C is the set of cohorts over which the relevant discontinuities occur, β_c is a cohort fixed effect, and the summation signs permit slopes to vary by cohort on either side of the cutoff.

Although equation (2) requires similar identification assumptions to equation (1), we can relax these somewhat using a difference-in-discontinuities design. Intuitively, a difference-in-discontinuities is estimating the change between two distinct regression discontinuity estimates over time. A standard difference-in-difference compares outcomes between a treatment and a control group before versus after a policy change. A difference-in-discontinuities makes the same comparison except treatment and control groups are defined within a narrow bandwidth on either side of a cutoff (Duggan, Gupta and Jackson, 2019).

In my setting, I can difference out any discontinuity occurring around January 2 for the untreated cohorts. Specifically, I estimate:

$$Y_i = \beta_0 D_i + \beta_1 (D_i * FRA_c) + \sum_{c \in C} \beta_{2,c} DOB_i + \sum_{c \in C} \beta_{3,c} (DOB_i * D_i) + \beta_c + e_i \tag{3.3}$$

where FRA_c is a dummy for cohorts affected by the increased retirement age. The β_1 coefficient identifies a causal effect if no other factors are influencing the size of the discontinuity other than the policy change. That is, any confounders changing discontinuously at the cutoff must be constant over time.

3.4.3 Bandwidth Selection

In my primary specifications I use a 30-day bandwidth. The main reason to use short bandwidths is to avoid modeling the non-linear seasonal effects shown in Figure 3.2. Additionally, Social Security and Medicare eligibility reset every first of the month, so keeping the window within a single month avoids introducing additional discontinuities. As a robustness check, I also consider optimal bandwidth selection tools. First, I use the technique developed by Calonico, Cattaneo and Titiunik (2014) which selects a bandwidth and estimates robust bias-corrected confidence intervals for a single discontinuity. Second, I implement the methods in Cattaneo, Titiunik and Vazquez-Bare (2020) which account for settings with multiple cutoffs. The technique estimates a pooled coefficient as a weighted average of the average treatment effects at each cutoff.

3.5 Results

3.5.1 Visual Evidence and Parametric Estimation

Figure 3.3 plots disability rates within the narrow 30-day bandwidth around the January 2 cutoff for the 6 treated and 12 untreated cohort discontinuities. To aggregate the data, I collapse disability rates by distance from the cutoff and weight by the number of observations in each date-of-birth cell. The fitted line and shaded confidence interval are computed on the collapsed data. Although there may be modest increases in disability rates around the cutoff, the discontinuity does not appear to be larger for treated cohorts in either sex.

Regression evidence from Table 3.3 also fails to find the policy changes had any effect. Columns (1,3,5) estimate the stacked regression using equation (2) and a sample only of affected cohorts. Columns (2,4,6) estimate the stacked difference-in-discontinuities using equation (3) and a sample of treated and control cohorts. Under the full sample difference-in-discontinuities specification, we can reject the policy reduced disability enrollment by less than -0.26 percentage points or increased enrollment by more than 0.27 percentage points.

3.5.2 Non-Parametric Estimation

Although the parametric estimation finds no effect, non-parametric estimation tools may uncover a different result. I first consider estimating an effect for each cohort separately. Specifically, I use the procedure developed by Calonico, Cattaneo and Titiunik (2014) to select bandwidths for each cohort and compute robust confidence intervals. Figure 3.4 shows the point estimate for each cohort. The mean treatment effect for the control cohorts is 0.15 percentage points and 0.20 percentage points for the treated cohorts. We cannot reject the difference in means for affected cohorts vs control cohorts is less than -0.20 percentage points or more than 0.30 percentage points.[7]

An alternative technique for aggregating the discontinuities is the Cattaneo, Titiunik and Vazquez-Bare (2020) method which computes robust weighted and pooled estimators. Tables 3.4 and 3.5 show the results or the treated and control cohorts. Again, the pooled estimate for both treated and control cohorts are no different from zero.

3.5.3 Validating the Regression Discontinuity Design

A key concern for the validity of the design is whether beneficiaries are able to manipulate the running variable. Because the baseline regressions are collapsed at the date of birth level, I can test for bunching by running the regressions without weights and replacing the count of observations within a cell as the dependent variable. Table 3.6 reports the results. There is evidence of a discontinuous increase in the number of beneficiaries born in January, but the size of the discontinuity appears stable across treated and untreated cohorts. Rather

[7]For a parametric estimation we could assume each coefficient is independent, aggregate them using standard formulas for the sum and variance of random variables, and compute a single confidence interval. However, this is not possible for the non-parametric estimation because it does not produce symmetric confidence intervals.

than manipulation of the running variable, this is likely caused by differential seasonal patterns in fertility (Buckles and Hungerman, 2013). Although this threatens identification for the undifferenced discontinuity specifications, the difference-in-discontinuities specifications remain unaffected.

An additional concern is whether the results are sensitive to the length of the bandwidth around the discontinuity. To test for bandwidth sensitivity, I rerun the preferred regression for various bandwidth lengths and plot the coefficients with their 95% confidence intervals in Figure 3.5. The coefficient is consistently a precise zero regardless of the size of the window.

3.5.4 Disability Beneficiary Characteristics

Although I do not find evidence the policy affected the total number of disability beneficiaries, it may have affected the characteristics of beneficiaries. In particular, marginal applicants would be more likely to have higher benefit amounts or more difficult to diagnose disabilities. To test for changes in the type of beneficiaries receiving benefits, I rely on the public use Disability Analysis File. I continue to use the stacked difference-in-discontinuities specification of equation (3). Because date of birth in the file is only reported at the month level, I use a wider 6 month bandwidth.[8]

Table 3.7 shows the results from the Disability Analysis File are consistent with the results from the Medicare file. The first column tests for bunching in the distribution of beneficiary birth dates among the 10% of SSDI beneficiaries. There is no evidence the treated cohorts have a larger spike in January birth dates relative to untreated cohorts.[9]

I also explore if there is a change in the type of primary disability diagnosis around the cutoff. I focus specifically on conditions classified as mental disorders or musculoskeletal

[8] Persons with January 1 birthdates should technically be classified on the left side of the discontinuity, but the public use data do not permit this.

[9] Figure 3.6 plots the density of births by month for treated and control cohorts and shows no evidence of bunching.

disabilities. Because these conditions may be hard to verify, some authors have suggested they are more likely among marginal applicants (Autor and Duggan, 2006). I do not find any change in the frequency of these conditions or other covariates such as the age of disability onset or the benefit amount. Overall, my results from the Disability Analysis File do not show evidence that less generous retirement benefits had any measurable effect on the disability program.

3.6 Comparison with Duggan, Singleton, and Song (2007)

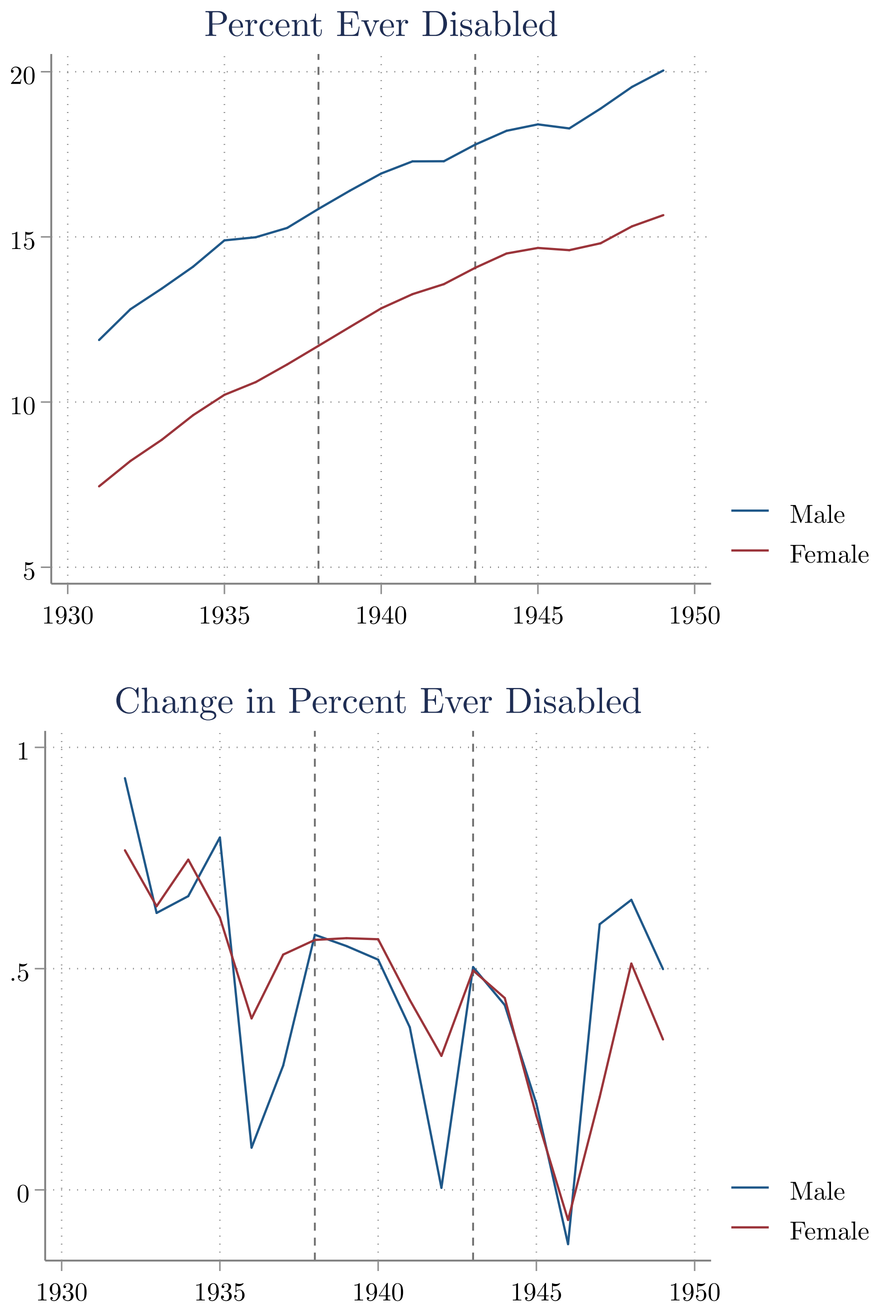
Why do DSS estimate a large effect while I estimate no effect? The different results arise from different identification strategies. I only exploit a single level of variation at the cohort level, but they exploit two levels of variation at the cohort-by-age level. Specifically, they compute the change in the present value of retirement benefits across cohorts over time. Their computation uses the mean benefit amount in 1999 and computes the discounted lifetime benefit using a 3% real rate as well as the mortality probability at each age. Thus, for a single cohort with no variation in the retirement age there is still variation across the lifecycle. Because the present value of benefits is lower at age 45 than at age 62, they predict the policy change will have a greater effect at older ages. Using data at the level of an age-by-time cell, they run the following specification:

$$\Delta SSDI_{A,t} = \alpha_t + \beta * \Delta PVR_{A,t} + \mu_A + e_{At} \tag{3.4}$$

where $\Delta SSDI_{A,t}$ is the change in per capita enrollment, $\Delta PVR_{A,t}$ is the present value of benefits,α_t is a time fixed effect and μ_A is an age fixed effect.

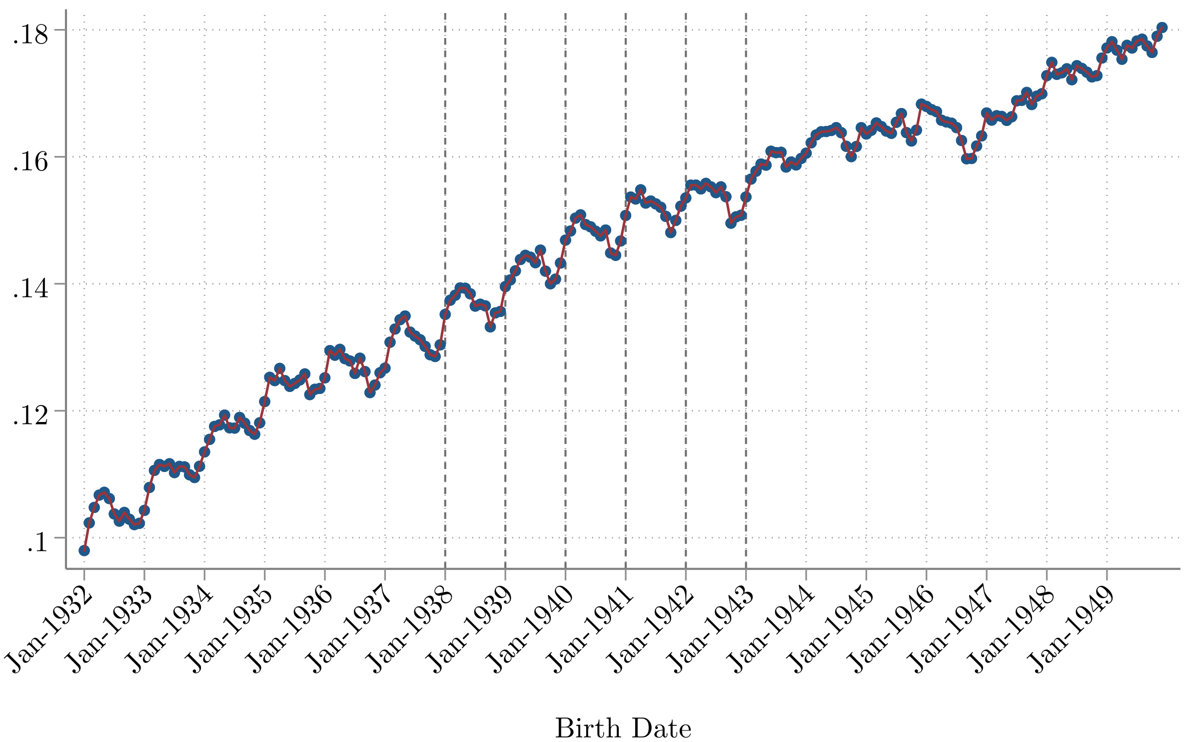
If within-cohort variation in the present value of benefits is driving greater enrollment, we would expect beneficiaries in affected cohorts to be claiming at older ages. To examine this I use the Disability Analysis File for all individuals who became entitled after age 40. Figure 3.7 shows a clear downward trend in the age at first entitlement that is unchanged by the policy phase-in. Similarly, Table 7 does not find evidence of discontinuities in the age of disability onset. Although the calculations necessary to compare the DSS coefficient to my coefficient are beyond the scope of this paper, my results suggest that within-cohort variation in the present value of benefits is not a primary driver of disability enrollment.

Figure 3.1: Disability Rates by Cohort and Sex



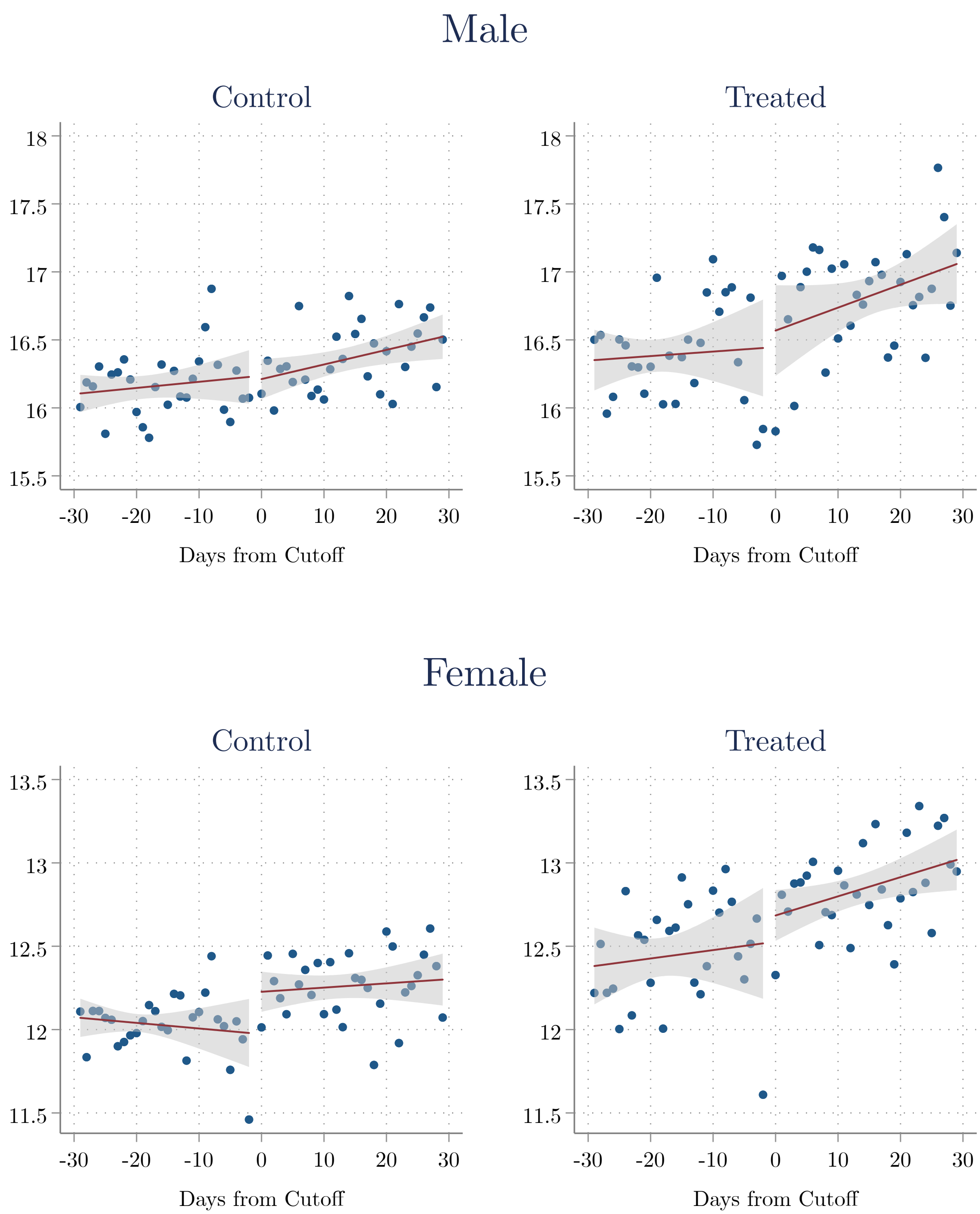
Notes: Percentage is defined as the share of Medicare beneficiaries who received entitlement through SSDI. Entitlement data is computed from 100% Master Beneficiary Summary File for selected years from 1999 to 2017. The left dashed line denotes when the policy change started and the right dashed lined denotes when the policy change ended.

Figure 3.2: Disability Rates by Birth Date



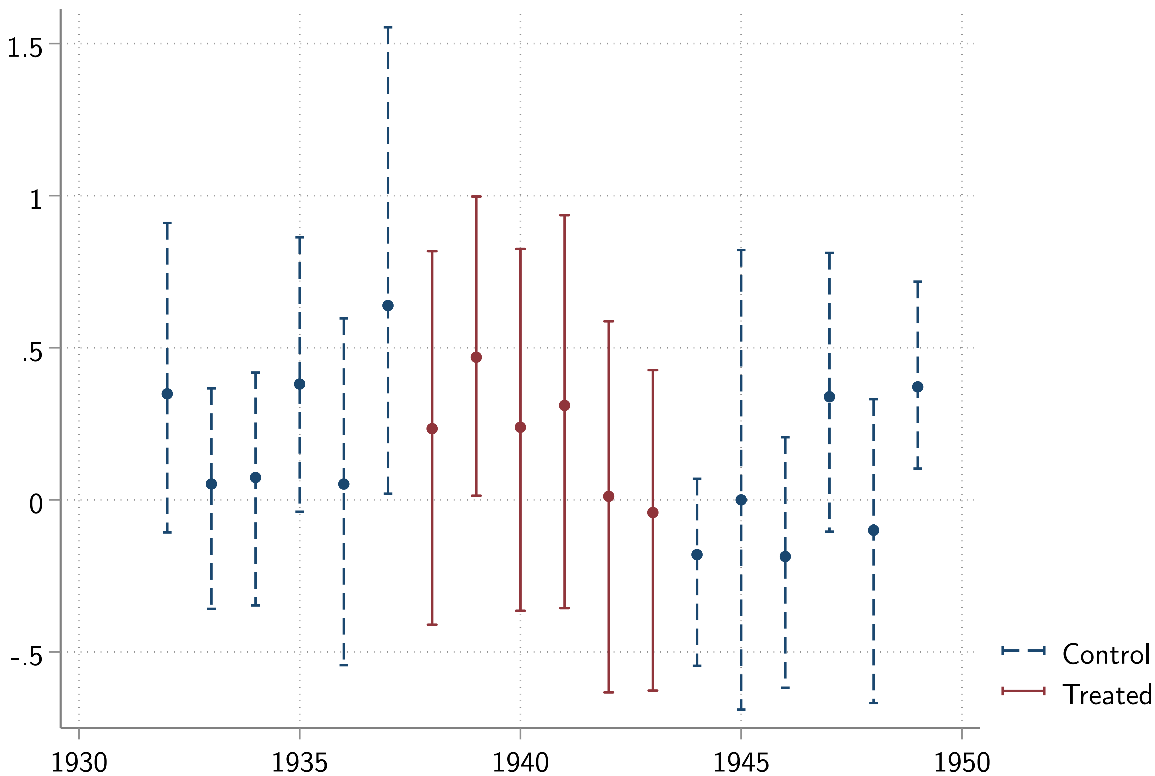
Notes: Data is collapsed at the month of birth level. Dashed lines denote discontinuities in the full retirement age.

Figure 3.3: Disability Rates For Treated and Control Cohorts



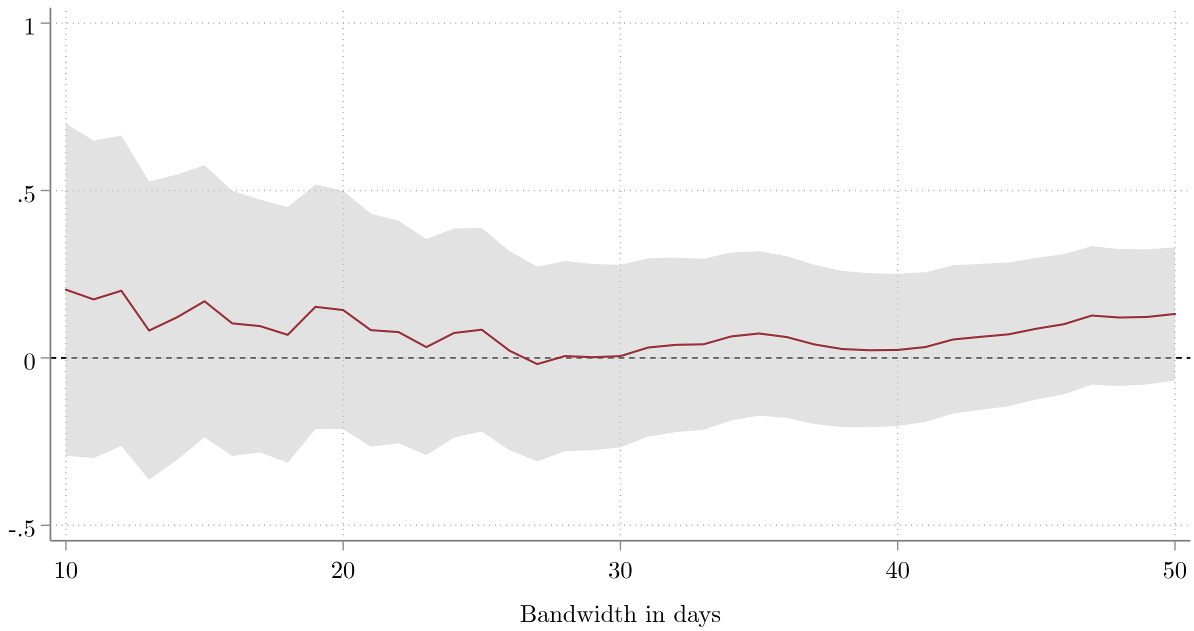
Notes: Data is collapsed by sex around the cutoff for 12 untreated cohorts and 6 treated cohorts. See text for details.

Figure 3.4: Coefficients from Nonparametric Estimators



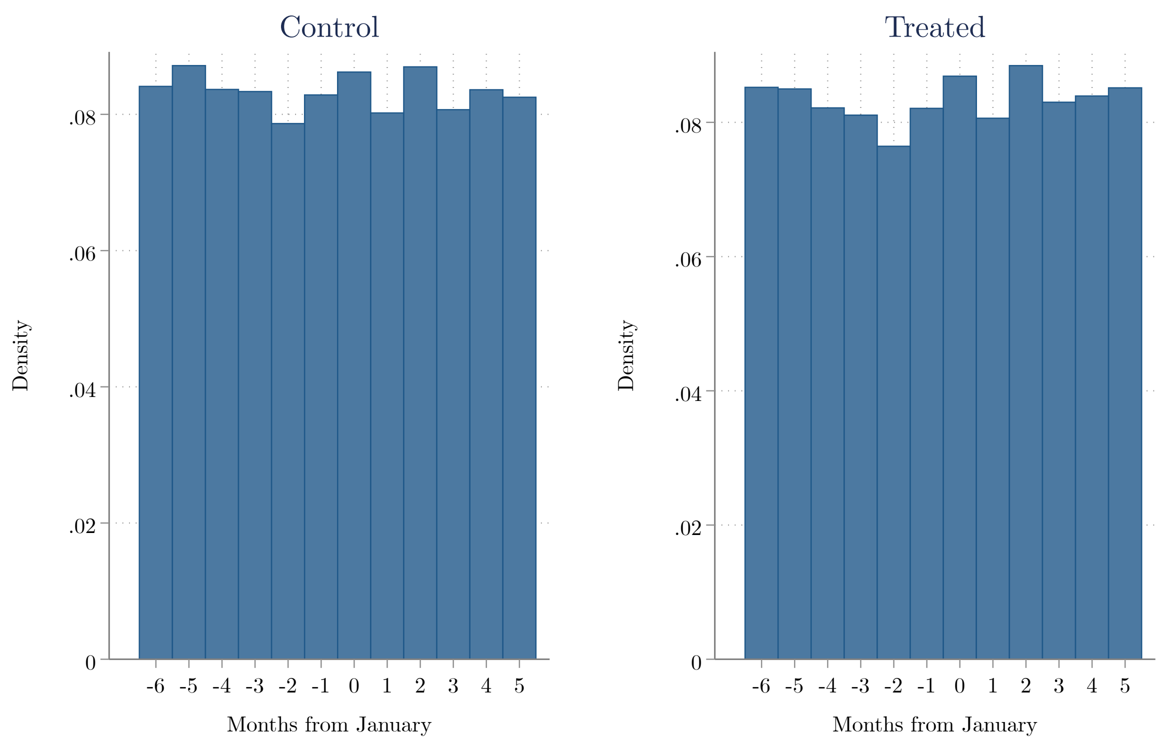
Notes: Each coefficients and its associated confidence intervals is estimated using the Calonico, Cattaneo and Titiunik (2014) method. The sample is the full population of males and females.

Figure 3.5: Bandwidth Sensitivity



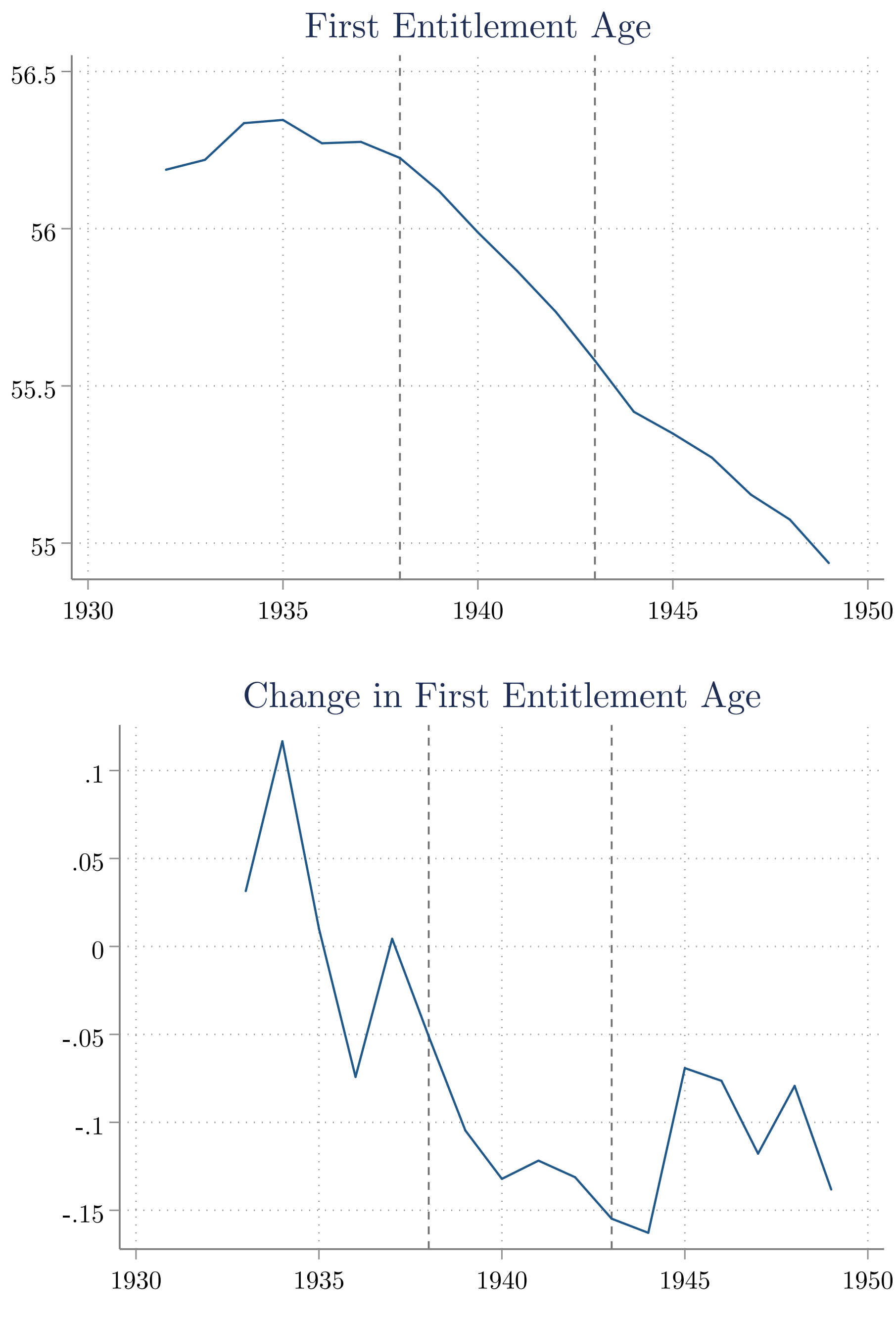
Notes: The treatment effect is estimated using the difference-in-discontinuities design and is not sensitive to bandwidth choice.

Figure 3.6: Histogram from Disability Analysis File



Notes: There is not evidence of differential bunching by date of birth in Disability Analysis File.

Figure 3.7: Trends in Entitlement Age by Cohort



Notes: Age at entitlement computed using public use Disability Analysis File. The left dashed line denotes when the policy change started and the right dashed lined denotes when the policy change ended. The policy changes does not appear to change the underlying downward trend.

Table 3.1: Retirement Benefit Parameters

Birth Year	FRA	% of PIA at claiming age			
		62	65	66	67
1931-32	65	80	100	105	110
1933-34	65	80	100	105.5	111
1935-36	65	80	100	106	112
1937	65	80	100	106.5	113
1938	65.2	79.2	98.9	105.4	111.9
1939	65.3	78.3	97.8	104.6	111.7
1940	65.5	77.5	96.6	103.5	110.5
1941	65.6	76.6	95.6	102.5	110
1942	65.8	75.8	94.4	101.3	108.8
1943-54	66	75	93.3	100	108
1955	66.2	74.2	92.2	98.9	106.7
1956	66.3	73.3	91.1	97.8	105.3
1957	66.5	72.5	90	96.7	104.0
1958	66.6	71.7	88.9	95.6	102.7
1959	66.8	70.8	87.6	94.4	101.3
1960	67	70	86.7	93.3	100

Notes: The table shows how benefits vary by birth cohort and claiming age.

Table 3.2: Estimation Sample Summary Statistics

	Mean	Standard Deviation
Restricted Use Medicare Data		
Female	0.54	
Non-White	0.12	
Ever Entitled to SSDI	0.15	
Public-Use DAF Benefits File		
Monthly Benefit Amount	1,346	541
Age at Onset	53.4	8.8
Musculoskeletal or Mental Disability	0.43	

Notes: Medicare data is computed from 100% Master Beneficiary Summary File for selected years from 1999 to 2017. Disability data is from 2018 version of the 10% public-use subsample of the Disability Analysis File. Both samples include beneficiaries born between 1932 and 1949.

Table 3.3: Fraction Ever Disabled

	Male		Female		All	
	(1)	(2)	(3)	(4)	(5)	(6)
Jan 2 Dummy	0.15	0.03	0.15	0.26***	0.17	0.16**
	(0.17)	(0.11)	(0.14)	(0.09)	(0.12)	(0.08)
FRA Interaction		0.12		-0.10		0.01
		(0.20)		(0.16)		(0.14)
Observations	360	1080	360	1080	360	1080
R^2	0.514	0.914	0.671	0.949	0.721	0.963
Demograpics Controls		X		X		X
Dep Var Mean	16.65	16.84	12.72	12.77	14.61	14.71

Notes: Columns (1,3,5) are OLS regressions from equation (2). Columns (2,4,6) are OLS regressions from equation (3). All specification use a 30 day bandwidth. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.4: Robust Multi Cutoff for Treated Cohorts

	Coefficient	Robust 95% Confidence Interval	P-value
1937/1938	0.20	[-0.43, 0.75]	0.60
1938/1939	0.46	[0.02, 0.98]	0.04**
1939/1940	0.23	[-0.36, 0.81]	0.45
1940/1941	0.31	[-0.34, 0.91]	0.37
1941/1942	0.02	[-0.63, 0.59]	0.95
1942/1943	-0.05	[-0.63, 0.39]	0.65
Weighted	0.18	[-0.07, 0.40]	0.16
Pooled	0.19	[-0.26, 0.61]	0.43

Notes: Pooled non-parametric estimates using Cattaneo, Titiunik and Vazquez-Bare (2020) method.

Table 3.5: Robust Multi Cutoff for Control Cohorts

	Coefficient	Robust 95% Confidence Interval	P-value
1932/1933	0.30	[-0.12, 0.81]	0.15
1933/1934	0.05	[-0.36, 0.37]	0.98
1934/1935	0.06	[-0.36, 0.41]	0.91
1935/1936	0.36	[-0.05, 0.82]	0.08*
1936/1937	0.05	[-0.51, 0.57]	0.91
1937/1938	0.57	[-0.03, 1.45]	0.06*
1943/1944	-0.19	[-0.56, 0.06]	0.11
1944/1945	-0.02	[-0.71, 0.79]	0.92
1945/1946	-0.19	[-0.63, 0.19]	0.30
1946/1947	0.33	[-0.11, 0.80]	0.14
1947/1948	-0.06	[-0.61, 0.34]	0.59
1948/1949	0.32	[0.05, 0.63]	0.02**
Weighted	0.12	[-0.01, 0.25]	0.07*
Pooled	0.10	[-0.75, 0.94]	0.83

Notes: Pooled non-parametric estimates using Cattaneo, Titiunik and Vazquez-Bare (2020) method.

Table 3.6: Number of Observations Within Date-of-Birth Cell

	Male		Female		All	
	(1)	(2)	(3)	(4)	(5)	(6)
Jan 2 Dummy	116.48***	130.42***	67.75**	90.91***	184.23***	221.33***
	(21.96)	(19.48)	(30.70)	(23.43)	(48.63)	(40.38)
FRA Interaction		-13.93		-23.16		-37.09
		(29.34)		(38.60)		(63.19)
Observations	360	1080	360	1080	360	1080
R^2	0.935	0.981	0.917	0.977	0.937	0.982
Demograpics Controls		X		X		X
Dep Var Mean	3098.38	3235.31	3369.24	3536.08	6467.62	6771.39

Notes: Columns (1,3,5) are OLS regressions from equation (2). Columns (2,4,6) are OLS regressions from equation (3). All specification use a 30 day bandwidth. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.7: Characteristics of SSDI Beneficiaries

	Cell Count	Log Count	Disability Type	Age at Onset	Benefit Amount
Jan Dummy	2304.39***	0.07***	0.55**	-0.03	1.59
	(673.07)	(0.02)	(0.26)	(0.05)	(2.93)
FRA Interaction	1397.28	0.02	0.28	0.02	1.28
	(1201.88)	(0.03)	(0.49)	(0.08)	(5.82)
Observations	216	216	216	216	216
R^2	0.990	0.990	0.964	0.969	0.982
Dep Var Mean	43849.54	10.63	41.27	53.52	1341.35

Notes: OLS regressions from equation (3). All specification use a 6 month bandwidth. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

APPENDIX A

Appendix

These appendices provide additional results and details. Appendix Section 1 presents robustness checks for the empirical strategy and provides econometric details. Appendix Section 2 shows descriptive results from other data sources on the correlation between income and mortality and income and health expenditures. Appendix Section 3 shows how to derive the income discontinuities from the Social Security benefit formula. Appendix Section 4 describes a model of health capital to provide intuition for the results.

A.1 Robustness and Additional Results

A.1.1 Stacked Regression Discontinuity in Difference

To test the sensitivity of the results, I also consider a specification that builds off the stacked RD design with an added feature from difference-in-discontinuities designs. Intuitively, a difference-in-discontinuities is estimating the change between two distinct regression discontinuity estimates over time. A standard difference-in-difference compares outcomes between a treatment and a control group before versus after a policy change. A difference-in-discontinuities makes the same comparison except treatment and control groups are defined within a narrow bandwidth on either side of a cutoff (Duggan, Gupta and Jackson, 2019; Persson, 2020).

In my setting, the discontinuities are differenced with respect to the discontinuities for the placebo years (1928 and 1934). In the context of a stacked RD, we can achieve this by interacting the January 2 dummy D_i with a vector of 0.01.

$$\log(Y_i) = \beta_0(D_i * 0.01) + \beta_1(D_i * S_c) + \sum_{c \in C} \beta_{2,c} DOB_i + \sum_{c \in C} \beta_{3,c}(DOB_i * D_i) + \beta_c + e_i \quad \text{(A.1)}$$

Now, β_1 measures the spending elasticity relative to the placebo cohorts. This is equivalent to plotting the coefficients against the income shock and allowing the intercept to vary. An intercept (β_0) statistically different from zero would provide evidence that the classic RD identification assumptions do not hold. Appendix Table A.6 shows for all major categories of spending the intercept is near zero, and elasticities are similar to the stacked RD design.

A.1.2 Stacked Regression Discontinuity by Sample Year

In the primary specification, the unit of observation is a date-of-birth cell and the dependent variable is the log sum of Medicare nominal spending over 6 years. Because Medicare prices are readjusted every year, there is some risk that the outcome is distorted by inflation in medical costs over time. To account for this, I consider an alternate specification where the unit of observation is a (date of birth) by (sample year) cell. This allows for the inclusion of sample year fixed effects which absorbs changes in prices over time.

$$\log(Y_{it}) = \beta_1(D_{it} * S_c) + \sum_{c \in C} \beta_{2,c} DOB_{it} + \sum_{c \in C} \beta_{3,c}(DOB_{it} * D_{it}) + \beta_c + \gamma_t + e_{it} \qquad \text{(A.2)}$$

Column (1) in Appendix Table A.9 presents estimates from this regression. Column (2) allows for cohort by sample fixed effects and column (3) interacts the slopes with the sample year to allow for different trends within cohort over time. In all cases, the elasticity is unchanged. This suggests that price changes within the sample period are not a concern.

A.1.3 Bandwidth Selection

In my preferred specification, I use a 30-day bandwidth for all cohorts. Using a narrow, consistent sample has the advantage of avoiding seasonality concerns, and allowing for clean decomposition of total spending by category. Nevertheless, it may be too narrow for certain cohorts. In particular, cohorts with fewer observations or higher variance of spending may require longer bandwidths. To account for this, I use the procedure developed by Calonico, Cattaneo and Titiunik (2014) to select bandwidths separately for each cohort, and then estimate equation (3) in the main text using these cohort-specific bandwidths.[1] Appendix Table A.10 presents the results. In general, bandwidths which are larger generates smaller standard errors and smaller point estimates, but the qualitative results are unchanged. We cannot reject these are different from the estimates in the main text.

A.1.4 Threats to Identification

Although there are several potential identification threats for regression discontinuity designs using date-of-birth variation, most are not relevant for my cohorts born in the 1920s and 1930s. For example, more recent cohorts can be affected by discontinuities in the tax code. Under the modern tax code, a mother giving birth in December can claim child-tax benefits when she files her tax return a few months later (Wingender and LaLumia, 2017). None of these benefits existed when my cohorts were born. Although parents might have been able to claim an additional exemption, very few households even filed taxes during this period.[2] A similar concern arises from the use of cesarean sections which may allow mothers to select their child's exact date of birth. Because C-section rates were under 5% during this period, selection by birth date is unlikely to confound the results. A final concern is whether school

[1] More recent work by Cattaneo et al. (2016) describes how to interpret regression discontinuity designs with multiple cumulative or non-cumulative cutoffs, but they do not consider a pooling approach when treatments have different signs.

[2] See https://www.irs.gov/pub/irs-prior/f1040–1931.pdf

entry cutoffs change around the calendar year. This is not likely to be an issue because cutoffs in most states were in August or September.

A.2 Income Correlations

A.2.1 Evidence on Expenditures from Survey Data

In order to give the causal estimates context, it is useful to explore the correlations of income on health expenditures and health outcomes. I rely on survey data because income is not directly observed in the administrative data. Although questions about health expenditures are included in many surveys, accurate measurement is challenging. Definitions of "health expenditures" are not always consistent, there are often multiple payers, and prices are rarely transparent. The Medical Expenditure Panel Survey (MEPS) attempts to minimize these sources of error by linking interviews from households to billing data from their medical providers. Merging the two data sources also allows MEPS to decompose expenditures made by payer.

I consider a sample of males receiving Social Security benefits between ages 65 and 84. Appendix Table A.12 presents OLS results with log annual payments for health care as the dependent variable. This excludes premium payments for insurance plans. I calculate spending elasticities with respect to total income for different payer sources. All models include survey year fixed effects, and a quadratic age term. Demographic controls include dummies for race, ethnicity, education, martial status, and census region. Columns (1-2) show the result across all payers, (3-4) show out-of-pocket payments, and (5-6) show payments made by Medicare on beneficiaries' behalf.

Two patterns emerge. First, demographic controls reduce the coefficient. This suggests that simple bivariate correlations between income and utilization have an upward bias. Second, price exposure matters. Elasticities are positive for out-of-pocket payments not covered by insurance, while elasticities are near zero for total payments across all payers. This makes sense under a model of full insurance. If consumers do not face prices at point of service, then income changes should not directly affect their willingness to use medical services. Expenditures covered by Medicare are almost fully insured because consumers have most of their cost-sharing covered by supplement plans. Given that a large share of Medicare payments are for hospital visits and other less "discretionary" services, the negative coefficient suggests improvements in health. Under this view, income buys more unobserved health investment which leads to better health. Better health reduces the need for Medicare services which are more curative than preventive. Other empirical research in a setting of near full insurance finds similar results. In the RAND Health Insurance Experiment, Phelps (1992) calculated income elasticities of 0.2 or less.

A.2.2 Evidence on Mortality from Administrative Data

Chetty et al. (2016) provides the best data on the correlation between total income and mortality. They use administrative tax records linked to Social Security mortality data for all individuals with a valid Social Security Number from 1999 to 2014. Their main results are for the full population, but in Online Table 15 they present summary data disaggregated by age. For the elderly, real income is measured at age 61 and mortality is measured as an average annual rate. Although Social Security benefits are computed over lifetime earnings, earnings at 61 provides a reasonable proxy.

Appendix Figure A.11 plots the relationship between log income and the log average mortality rate for adults aged over 65. The cross sectional elasticity of mortality with respect to total income (measured as the slope of the OLS fit line) is equal to -0.41. Because the Social Security formula has higher replacement rates for low-income workers, the elasticity with respect to Social Security income would be more negative (larger in absolute value).

A.3 More Detail on Social Security Rules and Data

The benefit discontinuities in this paper arise from the use of wage and price deflators in the Social Security benefit formula. Policymakers have struggled to implement these adjustments in a consistent way throughout the history of the program. The original 1935 Social Security Act had no wage or benefit indexation, and all cost-of-living adjustments required Congress to pass new legislation. Under this regime, the real value of benefits would erode over several years then abruptly increase when Congress intervened.[3] Policymakers attempted to automate this process in the 1972 Social Security Act, but made a technical error in the indexation formula. This led benefits to increase at nearly double the rate of inflation. In 1977, Congress created the modern benefit formula to correct this error, but did so in a way that abruptly cut benefits for individuals born on or after January 2, 1917 – a benefit discontinuity known as "the Notch."

Under the modern formula, workers and their spouse (either current or divorced) are eligible for Social Security benefits if the worker has at least 10 years of creditable labor market earnings. A creditable year is defined as earning above a certain inflation-indexed threshold ($4,480 for 2010).

A.3.1 Formula in Detail

Suppose individual i is born in calendar year b with taxable nominal earnings n_{it} in year t. SSA computes the average wage index (AWI_t) using IRS administrative data. This is used as an earnings deflator where indexed earnings y_{it} are defined as

$$y_{it} = \begin{cases} n_{it} \cdot \frac{AWI_{b+60}}{AWI_t} & t \leq b+60 \\ n_{it} & t > b+60 \end{cases}$$

[3]The lump-sum death benefit is the only feature of the original Social Security program which still exists today but has not been indexed to inflation. It was designed to help families pay for burial expenses after a worker's death. Congress set the maximum benefit at $255 in 1954 and has not updated it since.

so that earnings after the year an individual turns 60 are not adjusted. Indexed earnings are ordered such that $y_{i(1)} < y_{i(2)} \cdots < y_{i(n)}$ where $y_{i(1)}$ denotes the minimum indexed earnings in a person's wage history and $y_{i(n)}$ denotes the maximum indexed earnings. Average Indexed Monthly Earnings ($AIME_i$) is calculated as the average of the highest 35 years of indexed annual earnings divided by 12 months

$$AIME_i = \frac{1}{12}\left(\frac{1}{35}\sum_{j=0}^{34} y_{i(n-j)}\right)$$

Next, the Principal Insurance Amount (PIA_i) is computed as a concave function of $AIME_i$. The progressive formula has a marginal replacement rate that starts at 90%, declines to 32%, and reaches a minimum of 15% for the highest earners

$$PIA_i = \begin{cases} 0.9 \cdot AIME_i & AIME_i < k_{1,b} \\ 0.9 \cdot k_{1,b} + 0.32 \cdot (AIME_i - k_{1,b}) & k_{1,b} < AIME_i < k_{2,b} \\ 0.9 \cdot k_{1,b} + 0.32 \cdot (k_{2,b} - k_{1,b}) + 0.15 \cdot (AIME_i - k_{2,b}) & AIME_i > k_{2,b} \end{cases}$$

The kinks in the benefit schedule vary according to any individual's birth year where

$$k_{1,b} = 180 \cdot \frac{AWI_{b+60}}{AWI_{1977}} \qquad \text{and} \qquad k_{2,b} = 1085 \cdot \frac{AWI_{b+60}}{AWI_{1977}}$$

There is a cost-of-living adjustment ($COLA_t$) which increases benefits by a fixed percentage every December

$$COLA_t = \max\left(\frac{CPI_t}{\max\limits_{\tau<t} CPI_\tau}, 1\right)$$

where CPI_t denotes the mean of CPI-W in the third quarter of the year. This formula is designed to protect beneficiaries against deflation. When the level of the CPI index declines (as occurred in 2009 and 2010) nominal benefits levels stay constant. This protection, however, is "paid for" since later CPI increases are lower. Once the CPI starts to increase again, the base year for computing the percentage is the last highest year, not the previous year. Finally, the benefit is adjusted by the delayed retirement credit (DRC_b) which reduces the benefit for those who retire before the full retirement age (FRA_b), or increases it for those who claim after the FRA. These are indexed by b since they vary by year of birth as shown in Appendix Table A.13. The amount credited in year t can be summarized as

$$Benefit_{it} = PIA_i \cdot DRC_b \cdot \prod_{j=b+62}^{t} COLA_j$$

Consider two workers with identical earnings histories who claim benefits at their full retirement age of 65 ($DRC_b = 1$). One worker is born in December of year $b-1$ and the other is born one month later in January of year b. To simplify the calculation, we assume that the highest indexed earnings occur before age 60 and that the same years are included in the maximum 35 year average. Under these assumptions we can express

$$\frac{AIME_{Jan}}{AIME_{Dec}} = \frac{AWI_{b+60}}{AWI_{b+59}}$$

Because the kinks in the PIA formula also indexed by AWI we can similarly write

$$\frac{PIA_{Jan}}{PIA_{Dec}} = \frac{AWI_{b+60}}{AWI_{b+59}}$$

Regardless of when an individual claims, a cost-of-living adjustment (COLA) is applied in the calendar year after the year of first eligibility when they turn 62. Thus, benefits credited in the month after both individuals have claimed are

$$\frac{Benefit_{Jan}}{Benefit_{Dec}} = \frac{PIA_{Jan} \cdot COLA_{b+62} \cdot COLA_{b+63} \cdot COLA_{b+64}}{PIA_{Dec} \cdot COLA_{b+61} \cdot COLA_{b+62} \cdot COLA_{b+63} \cdot COLA_{b+64}}$$

Assuming that inflation is positive this simplifies to

$$\frac{Benefit_{Jan}}{Benefit_{Dec}} = \frac{AWI_{b+60} \cdot CPI_{b+60}}{AWI_{b+59} \cdot CPI_{b+61}}$$

Then, taking log differences we can express the percentage change in benefits as

$$\%\Delta Benefits \approx \%\Delta AWI_{b+60} - \%\Delta CPI_{b+61}$$

In practice, the impact of the change in base years will be somewhat different since earnings after 60 are not indexed, and indexation may change which the years which are included in the maximum 35 years.

Appendix Figure A.1 depicts how changes in parameters of the benefit formula contribute to the net change in benefits. Each calculation assumes nominal wage are identical on either side of birth date cutoff. Wage growth is positive in every year except for 2009 (birth cohort 1949) when it declined for the first time. Growth in benefits from the CPI is constrained by law to be positive. For beneficiaries born from 1938 to 1943 and 1954 to 1960, the net effect is 1.1 percentage points lower due to the rising retirement age.

A.3.2 Changes in the Delayed Retirement Credit

Appendix Table A.13 shows that half of the cohorts in the sample are affected by a 0.5% change in the delayed retirement credit (DRC). Although these changes violate the assumption that other policies change smoothly around the January 2 cutoff, they are small enough to be ignored. According to SSA Statistical Supplement 2007, Table 6.B5, only 3.6% of men are affected which means the change in real benefits across the whole cohort is less than 0.05%. I exclude disability conversions from the denominator. Figure 1.1 provides further evidence the DRC changes can be ignored. It shows there is not evidence of changes in claiming behavior.

A.3.3 Disability Benefits

Although beneficiaries who previously received Social Security disability payments use the same benefit formula, the details of its application differ. Both the number of years in the average and the base year are selected to maximize the PIA computation. The base year is either two years before the onset of their disability, or the normal base year at 60. A further complication is there is substantial bunching in the onset date of disability (29 months before the start of Medicare coverage). Appendix Figure A.10 uses the 2017 enrollment file to plot the number of Medicare beneficiaries who have ever been disabled by the date of disability onset. For clarity, the label next to each data point denotes the month of the onset. This bunching is due to an SSA policy that sometimes allows disability examiners to a select on an onset date that results in a more favorable benefit.[4] For this reason, the benefit discontinuities cannot be applied for disability beneficiaries.

[4]POMS DI 25501.300

A.4 A Model of Income and Health Spending

The standard framework for studying the demand for health and medical care is the Grossman (2000) model of health capital. In this model, individuals face a tradeoff between health and consumption of other commodities. Health is a durable stock variable that increases due to health investment or declines due to depreciation. The model makes a clear prediction that increased income will result in higher health investment and improved health outcomes.

A major challenge to testing this prediction in data is defining health investment. In the model, health investment is a single input of health goods and services sold in the market. This input includes everything from emergency department visits and vaccines to good nutrition and a low-stress lifestyle. As Kaestner (2013) notes, aggregating these inputs into a single index of health investment overlooks potential substitution possibilities between inputs. For example, preventive investments in healthy diet or exercise can substitute for medical investments in managing chronic disease.

A useful distinction to make is between ex-ante and ex-post health spending. Ex-ante spending can be broadly defined to include investments in health such as exercise, good nutrition, a low-pollution environment, or preventive care. Ex-post health spending can be narrowly defined to include acute medical care that mitigates the health loss due to a current illness. Because health insurance is more likely to cover ex-post spending than ex-ante spending, an income shock will affect these two categories differently.

To account for these features, I follow Grossman (2000) and Kaestner (2013) to develop a health capital model that distinguishes between ex-ante and ex-post health investment. Suppose that agents live for two-periods and have preferences over health and other consumption. In the first period, they inherit a health stock, h_1, and tradeoff between consumption, c_1, and how much to invest in future health, i_1. In the second period utility is discounted by β. Consumers are either sick or healthy and the probability of becoming sick, $\rho(i_1)$, declines with the level of investment such that $\rho'(i_1) < 0$. Consumers maximize expected utility

$$U(c_1, h_1) + \beta\rho(i_1)U(c_{2s}, h_{2s}) + \beta(1 - \rho(i_1))U(c_{2h}, h_{2h}) \tag{A.3}$$

If the consumer remains healthy, the stock of health evolves according to the standard law of motion.

$$h_{2h} = h_1(1 - \delta) + f(i_1) \tag{A.4}$$

If the consumer gets sick, they decide how much to spend on medical care, m_2. The additional loss to their health stock is denoted by $\lambda(m_2) > 0$ with $\lambda'(m_2) < 0$.

$$h_{2s} = h_1(1 - \delta - \lambda(m_2)) + f(i_1) \tag{A.5}$$

Income, y, is fixed and medical care is never consumed if a person is healthy. Prices for investment inputs and medical care are, p_i and, p_m, respectively. If there is no saving or borrowing, the budget constraints can be summarized as

$$y = c_1 + p_i i_1 \tag{A.6}$$

$$y = c_{2h} \tag{A.7}$$

$$y = c_{2s} + p_m m_2 \tag{A.8}$$

To summarize, agents have preferences over health and other consumption $U(c, h)$ with two periods and two states of health.

- At $t = 1$ everyone is healthy and inherits health stock h_1
  - They decide consumption and investment (c_1, i_1)
- At $t = 2$ the health shock is realized

– If they remain healthy, there is no need for medical care

  * $h_{2h} = h_1(1-\delta) + f(i_1)$
  * $c_{2h} = y$

– If they become sick, they consume medical care to mitigate the loss in health capital

  * Loss to health is $\lambda(m_2) > 0$ with $\lambda'(m_2) < 0$
  * $h_{2h} = h_1(1-\delta-\lambda(m_2)) + f(i_1)$
  * $c_{2h} = y - p_m m_2$

On average, health status is $\rho(i_1)h_{2s} + (1-\rho(i_1))h_{2h}$ and medical spending is $\rho(i_1)m_2$. Because consumption decisions in the second period are made after the shock is realized, we solve the problem by backward induction. We first maximize $U(c_{2s}, h_{2s})$ subject to the constraints in equations (9) and (12) to solve for the optimal consumption bundles as a function of first period investment $c^*_{2s}(i_1)$ and $h^*_{2s}(i_1)$. We then plug these expressions into equation (7) and maximize subject to the constraints in equations (8), (10), and (11).

The first order conditions are complicated and do not provide useful intuition, so instead I rely on simulations to study how income shocks affect medical spending and health. In particular, I follow Koka, Laporte and Ferguson (2014) and make the following assumptions regarding functional forms:

$$U(c,h) = c^{\alpha}h^{1-\alpha}$$

$$f(i_1) = i_1^{\gamma}$$

$$\lambda(m_2) = 1 - k_m m_2$$

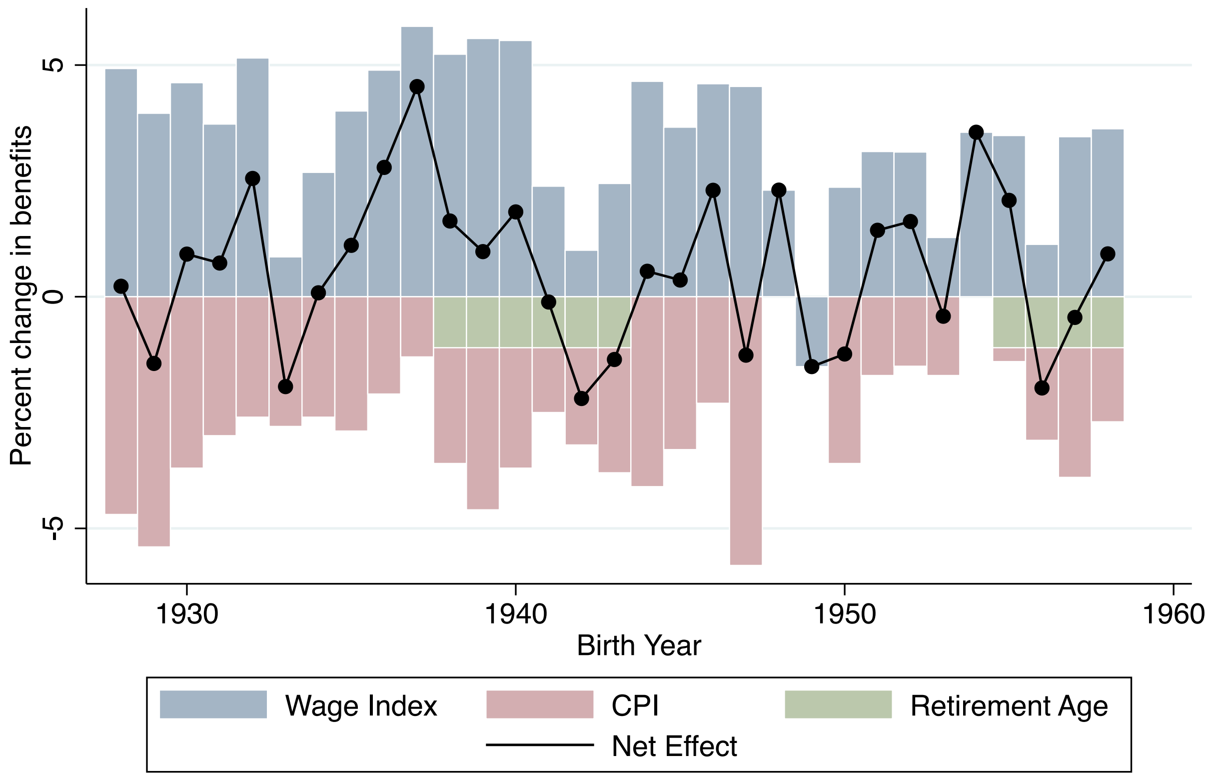
$$\rho(i) = exp(k_i i)$$

and the following assumptions for parameters:

α	γ	δ	p_i	p_m	h_0	k_m
0.6	0.8	0.1	1	1	1	1.5

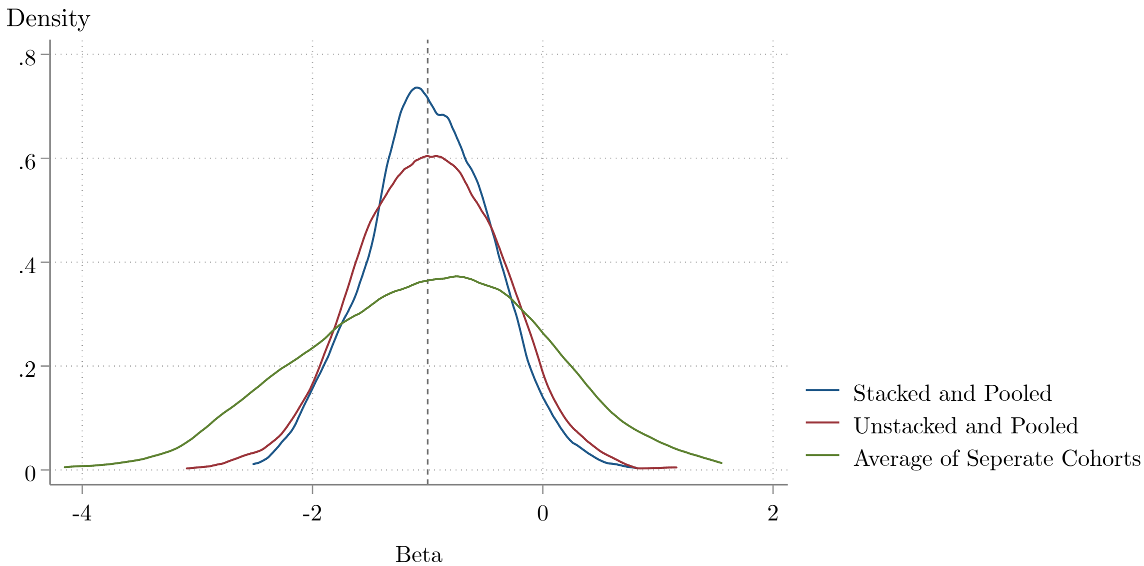
With most of the parameters fixed, I can solve the model numerically and test the relationship between expected medical spending, health investment, and income. In particular, I am interested in k_i, the parameter which affects how much ex-ante investment reduces the probability of illness. I consider two cases: high efficiency of health investment $k_i = -1$, and low efficiency of health investment $k_i = -4$. Appendix Figures 12 and 13 depict these relationships. As in the standard Grossman model, income and health investment are positively related because health is a normal good. In contrast, the relationship between income and medical spending is ambiguous. On one hand, income will increase ex-ante investment which will reduce the probability of illness. On the other hand, if the consumer does get sick, more income will reduce price sensitivity to acute care spending and so medical spending will increase. My empirical results find medical spending declines and health improves matching the simulation with a high efficiency of health investment.

Figure A.1: Parameter Changes for all Benefit Discontinuities



Notes: Benefit discontinuities occur because the Average Wage Index (AWI) and the Consumer Price Index (CPI) grow at different rates. The net effect in the figure is the change in benefits for a person born in January relative to a person born in December with the same nominal wage history. By law, CPI changes can only be positive. For beneficiaries born from 1938 to 1943, the net effect is lower due to the rising retirement age.

Figure A.2: Monte Carlo Comparison of Different Estimators

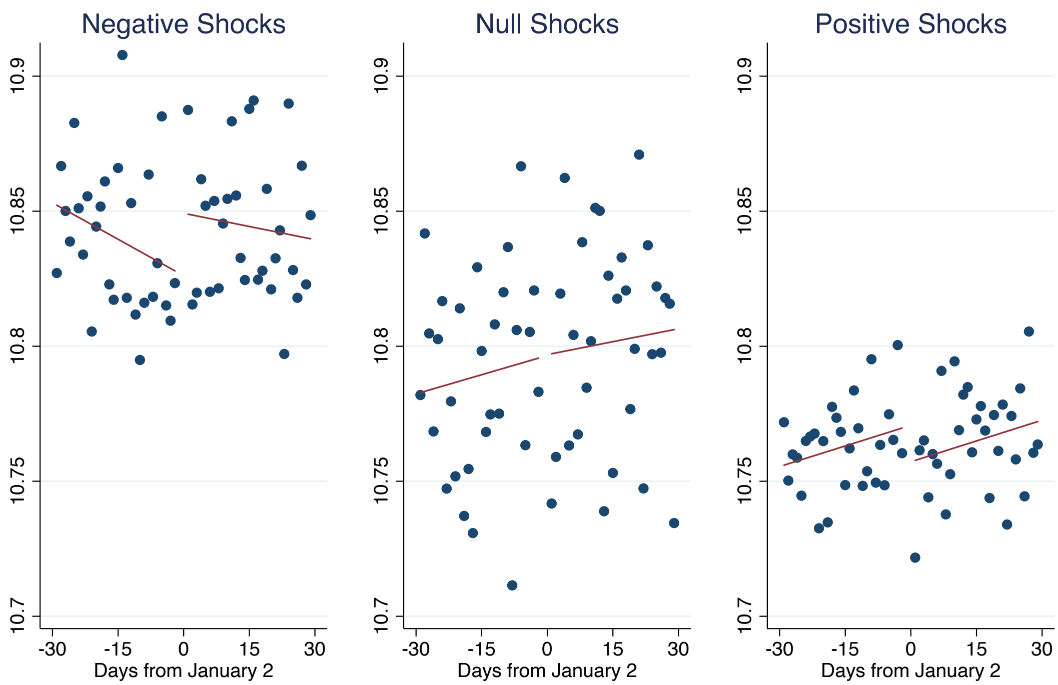


Notes: To construct a simple Monte Carlo simulation, I assume log spending is drawn from a normal distribution $Y \sim \mathcal{N}(10.8, 0.14)$, the treatment elasticity is $\beta = -1$, and there 10 income shocks equal to the predictions from equation (1). The "Average of Separate Cohorts" specification computes treatment effects by averaging together separate RD specifications in equation (2) and then rescaling the coefficient. The "Unstacked and Pooled" specification is the slope coefficient of equation (3). The "Stacked and Pooled" specification is the β_1 coefficient from equation (4). The figure plots the treatment estimates from 1,000 simulations. All estimates are unbiased, but the stacked and pooled specification is most efficient.

Figure A.3: Non-Parametric Spending Effects for Each Cohort

Notes: Each observation is the β_1 coefficient from equation (2) with its associated confidence interval. The x-axis denotes the predicted difference implied by equation (1).

Figure A.4: Log Total Medicare Spending by Treatment Type



Notes: The outcome is total log Medicare Part payments. See notes on Figure 3.

Figure A.5: Log Total Medicare Spending by Each Cohort

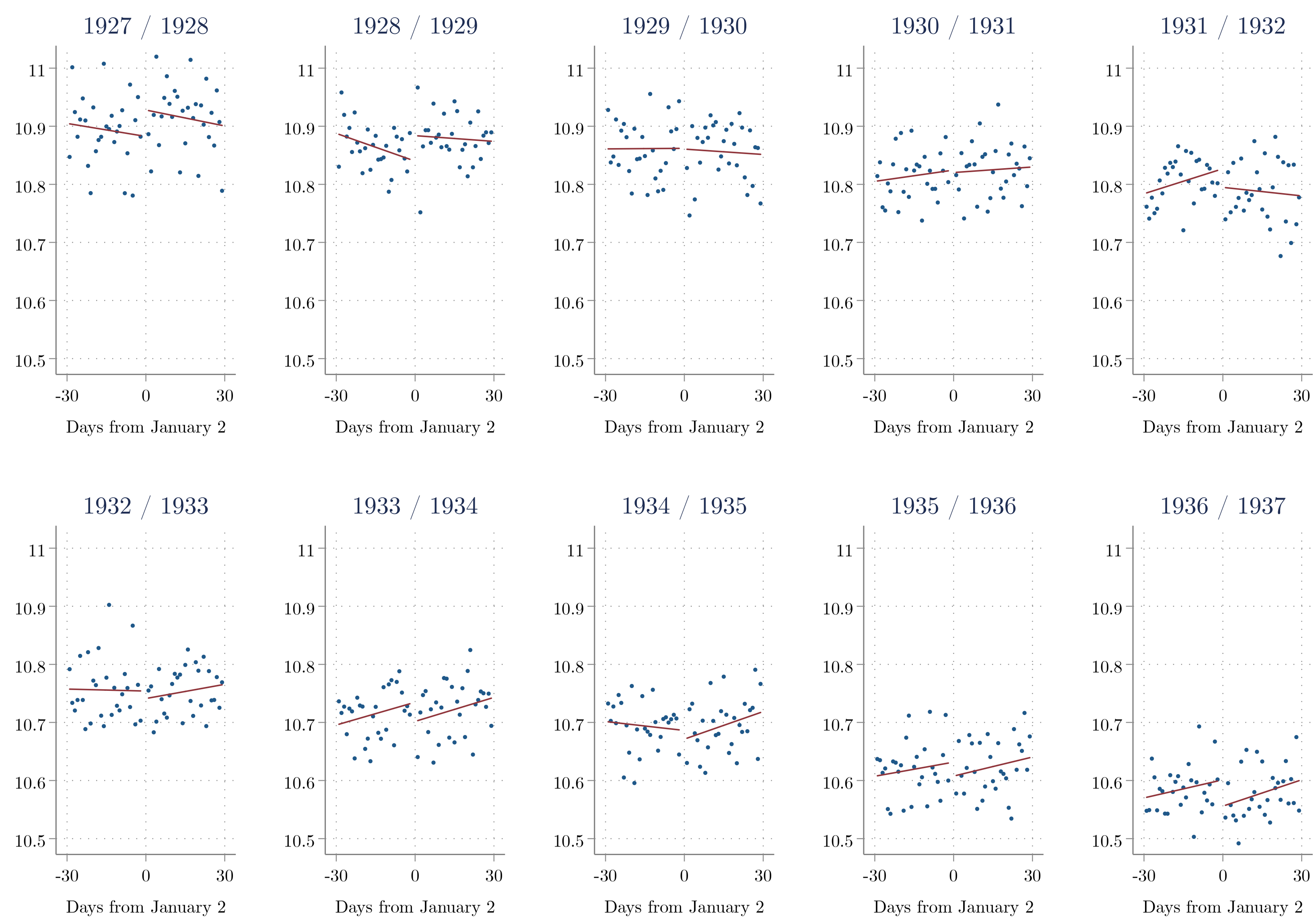
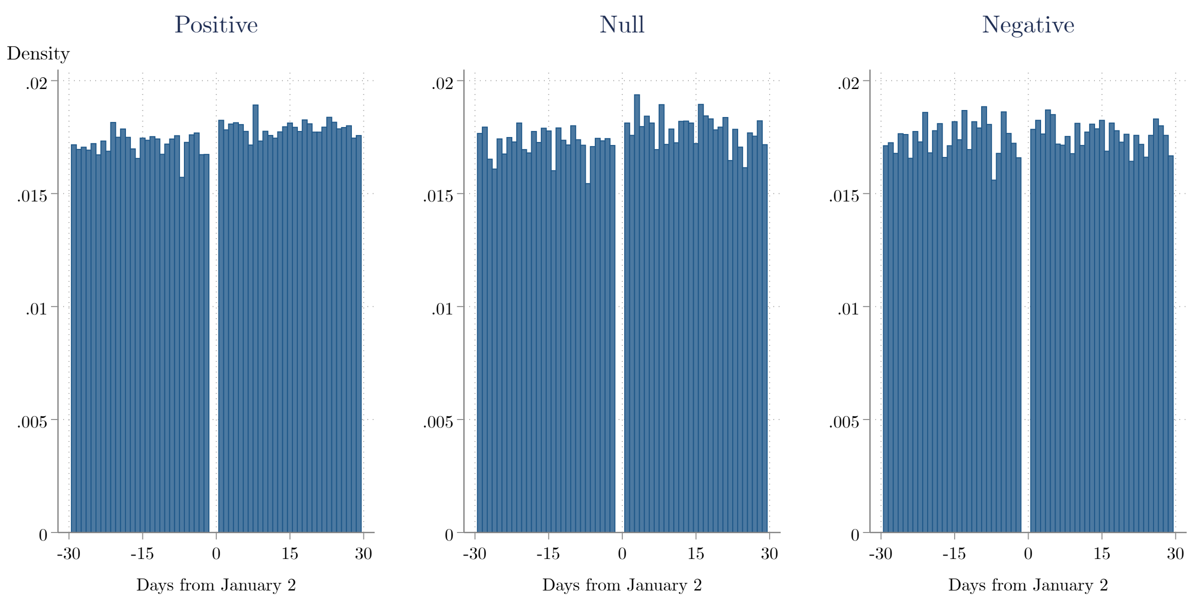
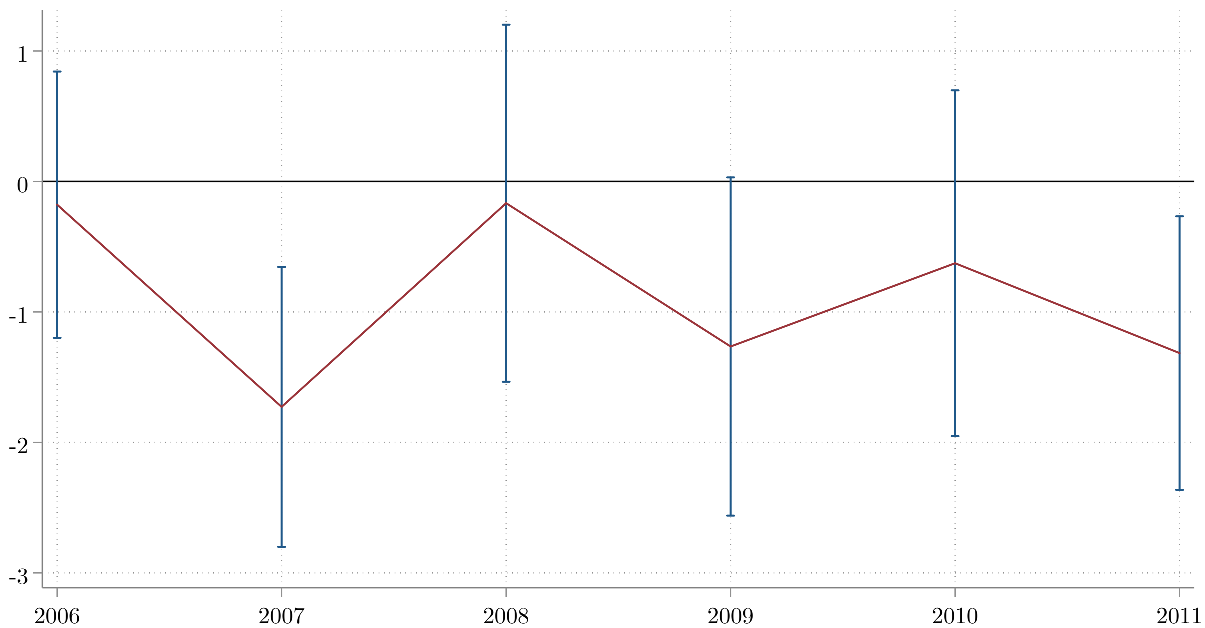


Figure A.6: Histogram by Treatment Type



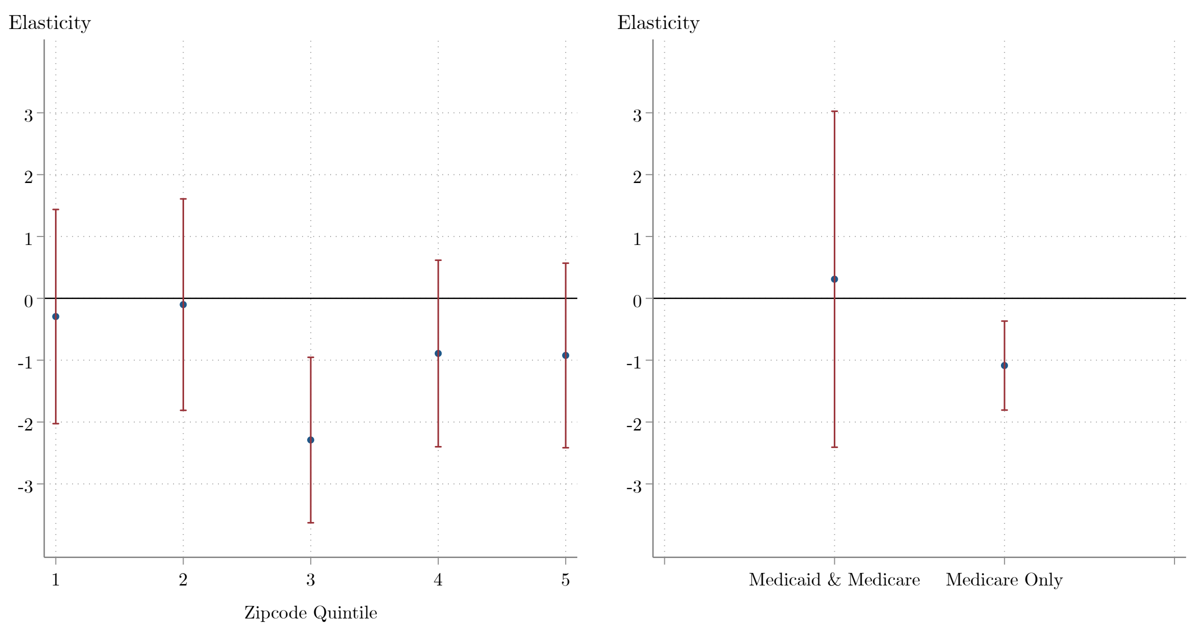
Notes: Density of observations around 30 days from the cutoff in the estimation sample, with January 1 and January 2 birth dates dropped. For all cohorts, there is a decline in reported births on December 26, the day following Christmas.

Figure A.7: Spending Elasticities over Time



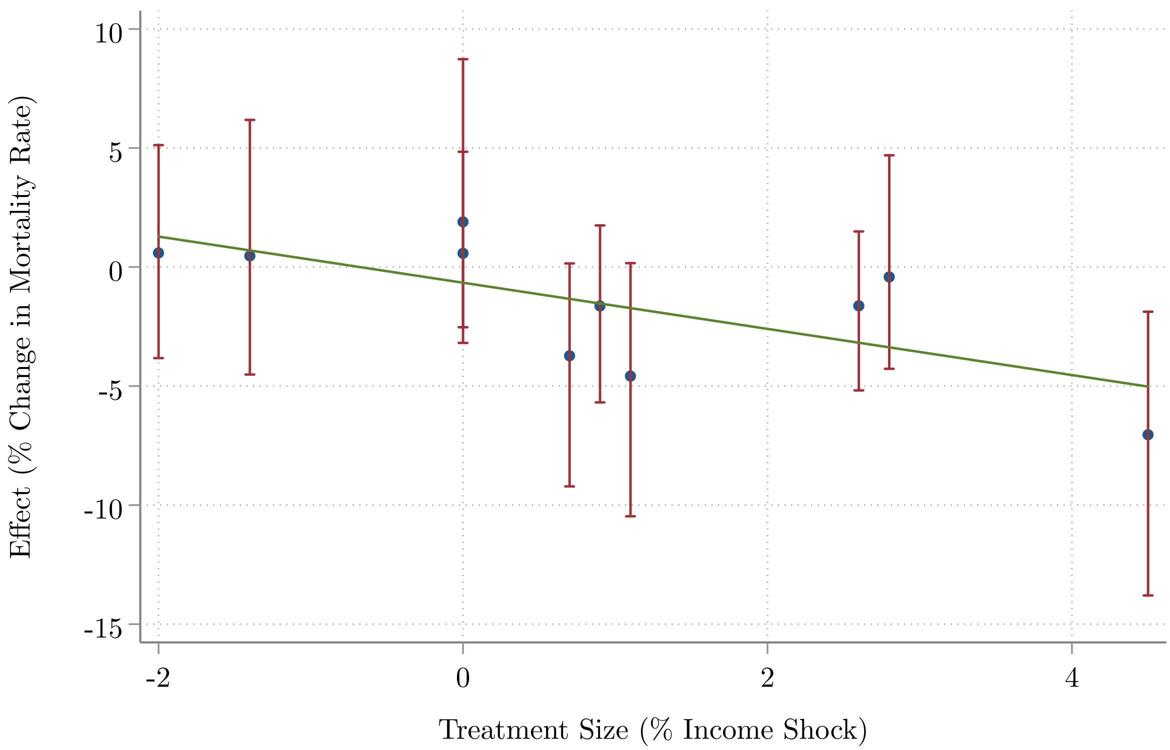
Notes: Elasticities of spending are computed by estimating equation (4). The dependent variable is log total spending for Medicare Part A and B for a given year. Although there is a slight downward slope, there is no clear pattern.

Figure A.8: Spending Elasticities by Zipcode Quintile



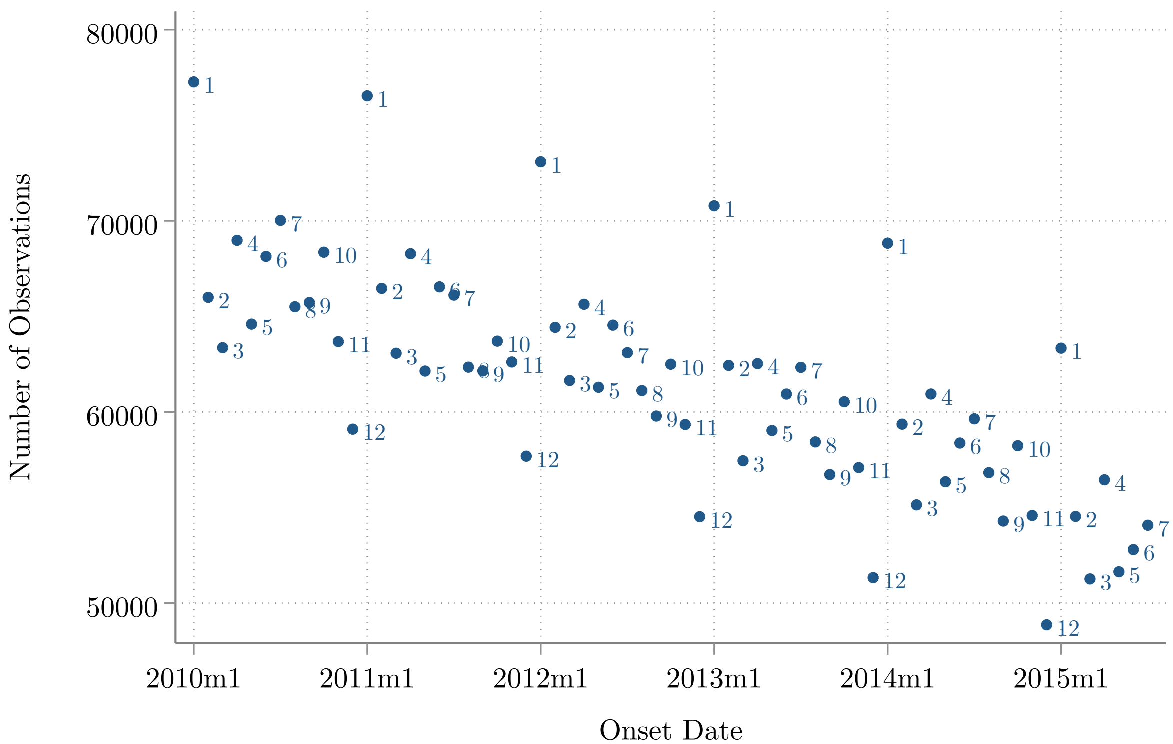
Notes: Elasticities of spending are computed by estimating equation (4) where the sample is divided into 5 zipcode quintiles. The dependent variable is log total spending within a zipcode quintile.

Figure A.9: Non-Parametric Mortality Effects for Each Cohort



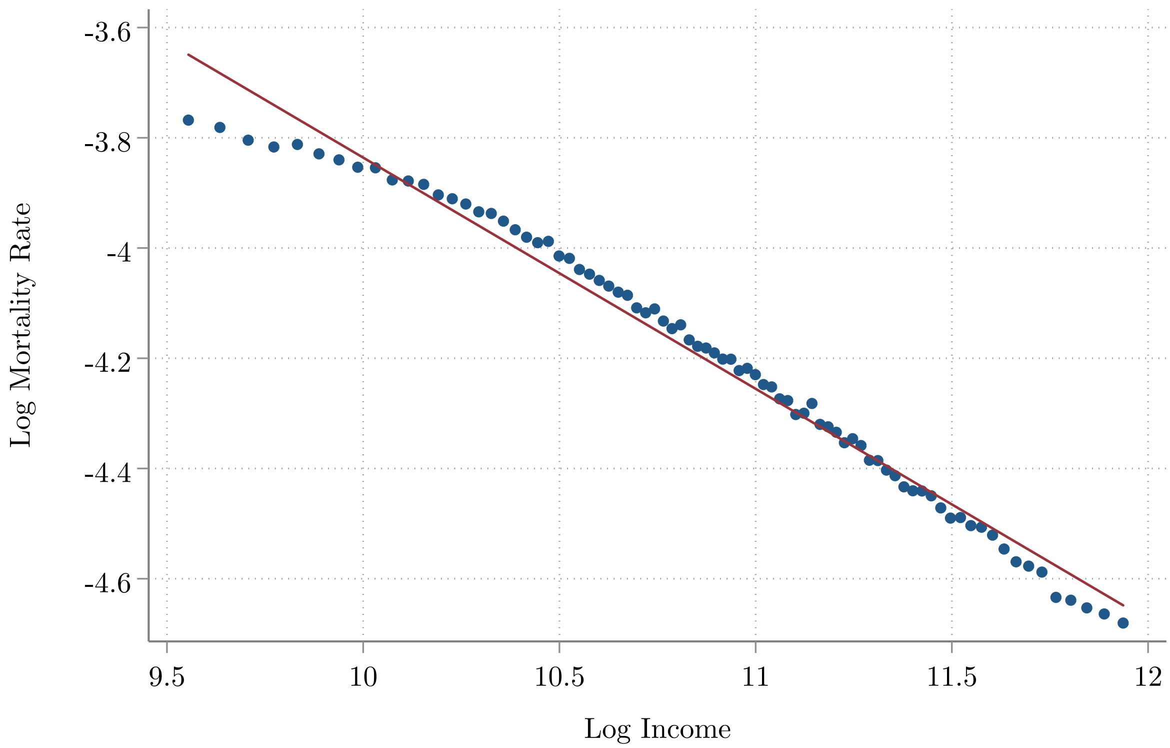
Notes: Each observation is the β_1 coefficient from equation (2) for a given cohort where the dependent variable is the log fraction of the baseline sample that has died by the end of 2017. Bandwidth selection and confidence intervals are computed using CCT procedure.

Figure A.10: Evidence of Bunching Among Disability Beneficiaries



Notes: The figure depicts counts of observations by disability onset date using the 2017 enrollment file. Labels denote month of onset. The bunching is due to an SSA rule which allows flexibility in the date of disability onset if it is advantageous to the beneficiary. Because of this, the benefit formula discontinuities cannot be used for disability beneficiaries.

Figure A.11: Elasticities of Mortality with Respect to Total Income



Notes: Summary data provided by Chetty et al. (2016) for adults over 65. Income measured at age 61 and mortality measured from 1999 to 2014. The slope of the OLS fit line is equal to -0.41.

Figure A.12: High Efficiency: $k_i = -1$

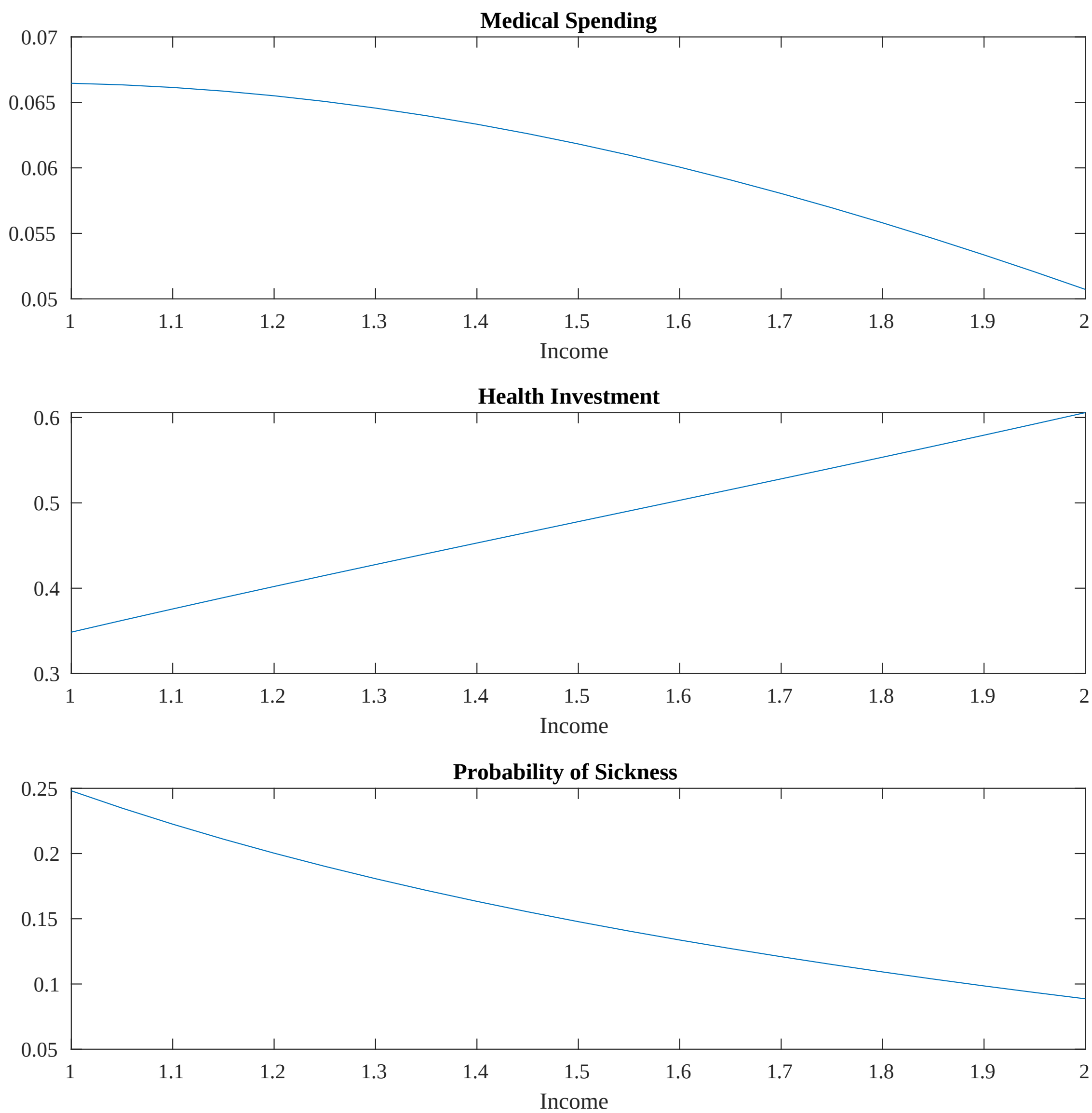


Figure A.13: Low Efficiency: $k_i = -4$

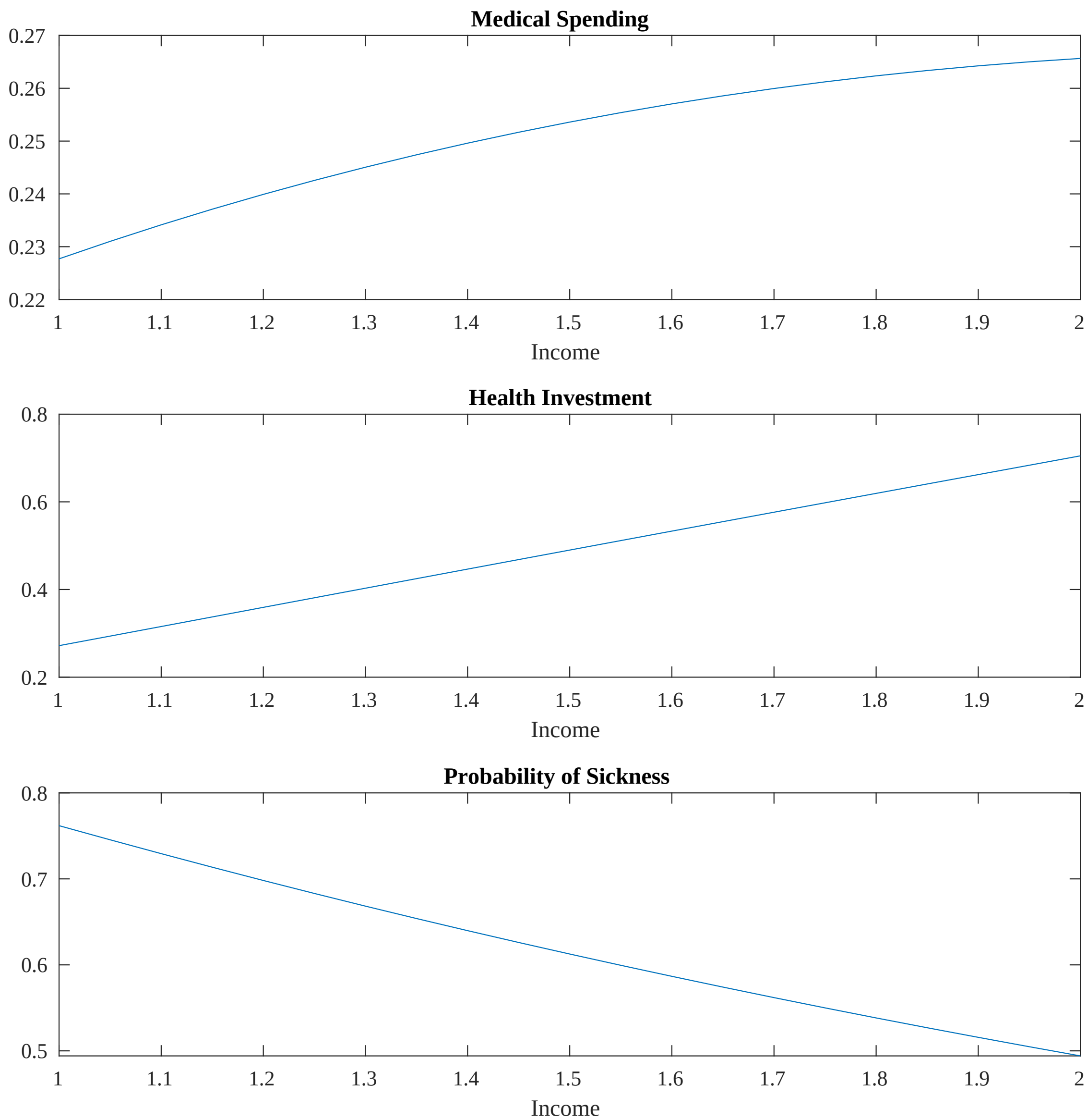


Table A.1: Average Demographics of FFS Only and Full Medicare Sample

	FFS Only	Full Sample
Age	76.0	75.9
White	0.88	0.83
Dual Coverage	0.08	0.09
Part D Coverage	0.42	0.57
Dead by 2017	0.34	0.34
Total Observations	3,287,465	5,435,185

Notes: FFS sample excludes beneficiaries without continuous Part A or Part B coverage as well as those who ever enrolled in Medicare Advantage.

Table A.2: Specification Tests

	(1) Baseline	(2) Baseline w/ Controls	(3) Global Quadratic	(4) Cohort Quadratic
Elasticity	-0.928**	-0.903**	-1.112***	-0.971*
	(0.376)	(0.377)	(0.402)	(0.581)
Observations	570	570	570	570
R^2	0.841	0.842	0.842	0.848
AIC	-1860.1	-1856.3	-1858.2	-1844.5
BIC	-1725.3	-1704.2	-1714.8	-1622.9

Notes: This table compares the fit of various specifications. The global quadratic specification is equation (4) with distance from cutoff squared and its interaction terms included. Cohort-specific quadratic allow these two terms to be estimated separately for each cohort.

Table A.3: Log Observation Count Within Date-of-Birth Cell

	Enrollment Sample		Estimation Sample	
	(1)	(2)	(3)	(4)
Elasticity	1.39***	0.40	1.09***	0.21
	(4.28)	(1.05)	(2.97)	(0.49)
Constant		4.66***		4.15***
		(5.33)		(4.19)
Observations	570	570	570	570
R^2	0.954	0.957	0.920	0.923

Notes: This table tests for possible manipulation of the running variable (date-of-birth). Columns (1) and (3) use equation (4) from the main text, and columns (2) and (4) use a regression discontinuity-in-difference specification from equation (5) in the Appendix. The enrollment sample is similar to the estimation sample, except without restrictions on enrollment in Part A, Part B, or Medicare Advantage.

Table A.4: Log Share of Population

	Estimination Sample	Part A/B	FFS	Black	Hispanic
Elasticity	-0.30	-0.00	-0.25	-0.02	0.92
	(0.19)	(0.06)	(0.16)	(0.99)	(0.99)
Observations	570	570	570	570	570
R^2	0.688	0.927	0.214	0.416	0.226
Outcome Population Share	60.91	93.09	66.61	6.40	7.14

Notes: This table estimates the elasticity of enrollment outcomes with respect to benefit shocks using equation (4). The elasticity can be interpreted as the effect of 1% change in benefits on the % share of the population with an income. The sample is similar to the estimation sample, except without restrictions on enrollment in Part A, Part B, or Medicare Advantage. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.5: Log Share of Population

	Dual Eligible	Part D Enrolled	Any Subsidy
Elasticity	0.336	-0.318	0.428
	(1.077)	(0.322)	(1.073)
Observations	570	570	570
R^2	0.127	0.360	0.125
Outcome Population Share	7.7	42.3	7.8

Notes: This table estimates the elasticity of enrollment outcomes with respect to benefit shocks using equation (4). The sample is the baseline estimation sample.

Table A.6: Log Expenditures by Payer using Difference in Discontinuity Design

	Total	Medicare	Beneficary	Part A	Part B
Intercept	-0.197	-0.212	-0.136	-0.512	0.010
	(0.941)	(0.981)	(0.793)	(1.393)	(0.869)
Elasticity	-0.884**	-0.896*	-0.816**	-0.584	-1.077***
	(0.436)	(0.457)	(0.359)	(0.661)	(0.383)
Observations	570	570	570	570	570
R^2	0.841	0.839	0.831	0.822	0.761

Notes: The table estimates the elasticity of spending using Appendix equation (5). The intercept term denotes the percentage change in spending for placebo cohorts.

Table A.7: Individual and Aggregate Regressions on Levels

	(1)	(2)	(3)	(4)	(5)
Scaled Dummy	-399.0**	-399.0**	-399.0**	-399.0**	-399.0**
	(161.1)	(192.0)	(155.7)	(197.2)	(160.0)
Observations	455968	455968	455968	570	570
Cluster	Robust	Cutoff distance	Date of birth	Cutoff distance	Robust

Notes: OLS regressions from equation (4) with spending level as the dependent variable. The estimate is the effect of a 1% change in benefits on total Medicare outlays over 6 years. Columns (1-3) are regressions from individual level micro data. Columns (4-5) are collapsed data at the date-of-birth level.

Table A.8: Log Medicare Expenditures with Heterogeneity by Age

	(1)	(2)	(3)	(4)	(5)	(6)
Elasticity	-0.880**	-0.884**	-0.968**	-0.874**	-1.022	-0.846**
	(0.355)	(0.367)	(0.386)	(0.358)	(0.708)	(0.400)
Age Interaction		-0.003				
		(0.048)				
Cohort Interaction			0.067			
			(0.145)			
Sample Interaction				0.012		
				(0.051)		
Observations	3420	3420	3420	3420	1881	1539
R^2	0.905	0.905	0.905	0.905	0.846	0.846
Sample	Baseline	Baseline	Baseline	Baseline	Above 75	75 and below

Notes: The table estimates the elasticity of spending using variations on the specification described in Appendix equation (6). The unit of observation is a (date of birth) by (sample year) cell. Column (2-4) include an income elasticity by time interaction. Columns (5) and (6) consider different samples after and below 75.

Table A.9: Log Medicare Expenditures with Flexible Slopes

	(1)	(2)	(3)
Elasticity	-0.909***	-0.909***	-0.909**
	(0.257)	(0.256)	(0.360)
Observations	3420	3420	3420
R^2	0.903	0.905	0.908
Fixed Effect Level	Cohort and Sample	Cohort by Sample	Cohort and Sample
Slope Level	Cohort	Cohort	Cohort by Sample

Notes: The table estimates the elasticity of spending using variations on the specification described in Appendix equation (5). The unit of observation is a (date of birth) by (sample year) cell. Column (1) constrains the slope for each cohort to be equal across sample years. Column (2) keeps the slopes constraint but allows for individual cohort-sample fixed effects. Column (3) allows different slopes and fixed effects for each of the 60 cohort-sample year combinations. Errors are clustered at the date of birth level.

Table A.10: Log Expenditures by Payer using Separate Cohort Bandwidths

	Total	Medicare	Beneficary	Part A	Part B
Elasticity	-0.579**	-0.643**	-0.501**	-0.674	-0.604**
	(0.280)	(0.290)	(0.253)	(0.417)	(0.271)
Observations	1132	1150	1004	1214	996
R^2	0.831	0.831	0.806	0.827	0.732
Average CCT Bandwidth	57.47	58.48	51.54	61.69	50.96

Notes: OLS estimates from equation (4) where bandwidths are selected using the CCT procedure.

Table A.11: Log Total Expenditures with Dates of Birth Dropped

	(1)	(2)	(3)	(4)
Elasticity	-0.928**	-0.841**	-0.783**	-0.982**
	(0.376)	(0.366)	(0.338)	(0.390)
Observations	570	580	590	560
R^2	0.841	0.842	0.843	0.840
Dropped Days	Jan 1, Jan 2	Jan 1	None	Dec 26, Jan 1, Jan 2

Notes: The table shows the results are not sensitive to dropping or including particular birth dates.

Table A.12: Elasticities of Health Expenditures by Payer

	All Payers		Out-of-Pocket		Medicare	
	(1)	(2)	(3)	(4)	(5)	(6)
Log Income	0.03**	-0.04***	0.36***	0.20***	-0.13***	-0.17***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
Observations	18144	18144	18144	18144	18144	18144
R^2	0.020	0.039	0.059	0.115	0.062	0.072
Demograpics Controls		X		X		X

Notes: The table shows results from OLS regressions of male Social Security beneficiaries aged 65-84 in the Medical Expenditure Panel Survey (2000-2017). All models include survey year fixed effects, and a quadratic age term. Demographic controls include dummies for race, ethnicity, education, martial status, and census region. Robust standard errors in parentheses.

Table A.13: Delayed Retirement Credits and Full Retirement Age

Birth Year	FRA	Benefit (% of PIA) claiming at age			
		62	65	66	70
1924	65	80	100	103	115
1925-26	65	80	100	103.5	117.5
1927-28	65	80	100	104	120
1929-30	65	80	100	104.5	122.5
1931-32	65	80	100	105	125
1933-34	65	80	100	105.5	127.5
1935-36	65	80	100	106	130
1937	65	80	100	106.5	132.5
1938	65.17	79.17	98.89	105.42	131.42
1939	65.33	78.33	97.78	104.67	132.67
1940	65.50	77.5	96.67	103.5	131.5
1941	65.67	76.67	95.56	102.5	132.5
1942	65.83	75.83	94.44	101.25	131.25
1943-54	66	75	93.33	100	132

Notes: The table shows how benefits vary by birth cohort and claiming age.

Bibliography

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. "Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program." *Journal of the American statistical Association*, 105(490): 493–505.

Alexander, Diane, and Molly Schnell. 2019. "The impacts of physician payments on patient access, use, and health." National Bureau of Economic Research.

Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics. An Empiricist's Companion.*

Apouey, Benedicte, and Andrew E. Clark. 2015. "Winning big but feeling no better? the effect of lottery prizes on physical and mental health." *Health Economics (United Kingdom).*

Aron-Dine, Aviva, Liran Einav, and Amy Finkelstein. 2013. "The RAND health insurance experiment, three decades later." *Journal of Economic Perspectives*, 27(1): 197–222.

Autor, David H, and Mark Duggan. 2010. "Supporting work: A proposal for modernizing the US disability insurance system." *Washington, DC: Center for American Progress and the Hamilton Project.*

Autor, David H., and Mark G. Duggan. 2006. "The Growth in the Social Security Disability Rolls: A Fiscal Crisis Unfolding." *Journal of Economic Perspectives*, 20(3): 71–96.

Ayyagari, Padmaja, and David Frisvold. 2016. "The Impact of Social Security Income on Cognitive Function at Older Ages." *American Journal of Health Economics*, 2(4): 463–488.

Barham, Tania, and Jacob Rowberry. 2013. "Living longer: The effect of the Mexican conditional cash transfer program on elderly mortality." *Journal of Development Economics*, 105: 226 – 236.

Barreca, Alan I., Jason M. Lindo, and Glen R. Waddell. 2016. "Heaping-Induced Bias in Regression-Discontinuity Designs." *Economic Inquiry*, 54(1): 268–293.

Bee, Adam, and Joshua Mitchell. 2017. "Do Older Americans Have More Income Than We Think?" *Proceedings. Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association*, 110: 1–85.

Bishop, Kelly C, Jonathan D Ketcham, and Nicolai V Kuminoff. 2018. "Hazed and Confused: The Effect of Air Pollution on Dementia." National Bureau of Economic Research Working Paper 24970.

Borghans, Lex, Anne C. Gielen, and Erzo F. P. Luttmer. 2014. "Social Support Substitution and the Earnings Rebound: Evidence from a Regression Discontinuity in Disability Insurance Reform." *American Economic Journal: Economic Policy*, 6(4): 34–70.

Brot-Goldberg, Zarek C., Amitabh Chandra, Benjamin R. Handel, and Jonathan T. Kolstad. 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.

Brown, Jeffrey R., Arie Kapteyn, and Olivia S. Mitchell. 2016. "Framing And Claiming: How Information-framing Affects Expected Social Security Claiming Behavior." *Journal of Risk and Insurance*, 83(1): 139–162.

Brown, Jeffrey R., Arie Kapteyn, Erzo F.P. Luttmer, Olivia S. Mitchell, and Anya Samek. 2019. "Behavioral Impediments to Valuing Annuities: Complexity and Choice Bracketing." *The Review of Economics and Statistics*, 1–45.

Buckles, Kasey S., and Daniel M. Hungerman. 2013. "Season of Birth and Later Outcomes: Old Questions, New Answers." *The Review of Economics and Statistics*, 95(3): 711–724.

Burns, Marguerite, and Laura Dague. 2017. “The effect of expanding Medicaid eligibility on Supplemental Security Income program participation.” *Journal of Public Economics*, 149: 20 – 34.

Calonico, Sebastian, Matias Cattaneo, and Rocio Titiunik. 2014. “Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs.” *Econometrica*, 82(6): 2295–2326.

Card, David, David S. Lee, Zhuan Pei, and Andrea Weber. 2015. “Inference on Causal Effects in a Generalized Regression Kink Design.” *Econometrica*, 83(6): 2453–2483.

Carey, Colleen M, Sarah Miller, and Laura R Wherry. 2018. “The Impact of Insurance Expansions on the Already Insured: The Affordable Care Act and Medicare.” National Bureau of Economic Research Working Paper 25153.

Caswell, Kyle J, and Timothy A Waidmann. 2017. “Medicare Savings Program Enrollees and Eligible Non-Enrollees.” *Report for Medicaid and CHIP Payment and Access Commission (MACPAC).*

Cattaneo, Matias D., Luke Keele, Rocío Titiunik, and Gonzalo Vazquez-Bare. 2016. “Interpreting Regression Discontinuity Designs with Multiple Cutoffs.” *The Journal of Politics*, 78(4): 1229–1248.

Cattaneo, Matias D, Rocio Titiunik, and Gonzalo Vazquez-Bare. 2020. “Analysis of regression-discontinuity designs with multiple cutoffs or multiple scores.” *The Stata Journal*, 20(4): 866–891.

Cawley, John, John Moran, and Kosali Simon. 2010. “The impact of income on the weight of elderly Americans.” *Health Economics.*

Cesarini, David, Erik Lindqvist, Robert Östling, and Björn Wallace. 2016. “Wealth, Health, and Child Development: Evidence from Administrative Data on Swedish Lottery Players." *The Quarterly Journal of Economics*, 131(2): 687–738.

Chandra, Amitabh, and Craig Garthwaite. 2019. "Economic Principles for Medicare Reform." *The ANNALS of the American Academy of Political and Social Science*, 686(1): 63–92.

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

Chen, Alice. 2014. "Do the poor benefit from more generous Medicaid policies?" *Available at SSRN 2444286*.

Chetty, Raj, Michael Stepner, Sarah Abraham, Shelby Lin, Benjamin Scuderi, Nicholas Turner, and Augustin Bergeron. 2016. "The Association Between Income and Life Expectancy in the United States, 2001-2014." *Journal of the American Medical Association*, 315(16): 1750–1766.

Clemens, Jeffrey, and Joshua D Gottlieb. 2014. "Do physicians' financial incentives affect medical treatment and patient health?" *American Economic Review*, 104(4): 1320–49.

Clemens, Jeffrey, and Joshua D. Gottlieb. 2017. "In the Shadow of a Giant: Medicare's Influence on Private Physician Payments." *Journal of Political Economy*, 125(1): 1–39.

CMS. 2015. "Access to Care Issues Among Qualified Medicare Beneficiaries (QMB)." Centers for Medicare and Medicaid Services.

Conley, Timothy G, and Christopher R Taber. 2011. "Inference with "difference in differences" with a small number of policy changes." *The Review of Economics and Statistics*, 93(1): 113–125.

Cutler, David, Angus Deaton, and Adriana Lleras-Muney. 2006. "The Determinants of Mortality." *Journal of Economic Perspectives*, 20(3): 97–120.

Deryugina, Tatyana, Garth Heutel, Nolan H. Miller, David Molitor, and Julian Reif. 2019. "The Mortality and Medical Costs of Air Pollution: Evidence from Changes in Wind Direction." *American Economic Review*, 109(12): 4178–4219.

Deschênes, Olivier, and Enrico Moretti. 2009. "Extreme Weather Events, Mortality, and Migration." *The Review of Economics and Statistics*, 91(4): 659–681.

Deshpande, Manasi, Itzik Fadlon, and Colin Gray. 2020. "How Sticky is Retirement Behavior in the U.S.? Responses to Changes in the Full Retirement Age." National Bureau of Economic Research Working Paper 27190.

Dobkin, Carlos, and Steven L. Puller. 2007. "The effects of government transfers on monthly cycles in drug abuse, hospitalization and mortality." *Journal of Public Economics*, 91(11-12): 2137–2157.

Doyle, Joseph J., Steven M. Ewer, and Todd H. Wagner. 2010. "Returns to physician human capital: Evidence from patients randomized to physician teams." *Journal of Health Economics*, 29(6): 866 – 882.

Doyle, Joseph, John Graves, and Jonathan Gruber. 2018. "Evaluating Measures of Hospital Quality: Evidence from Ambulance Referral Patterns." *The Review of Economics and Statistics*, 101(5): 841–852.

Duggan, Mark, and Scott A Imberman. 2009. "Why are the disability rolls skyrocketing? The contribution of population characteristics, economic conditions, and program generosity." In *Health at older ages: The causes and consequences of declining disability among the elderly.* 337–379. University of Chicago Press.

Duggan, Mark, Atul Gupta, and Emilie Jackson. 2019. "The Impact of the Affordable Care Act: Evidence from California's Hospital Sector." National Bureau of Economic Research Working Paper 25488.

Duggan, Mark, Perry Singleton, and Jae Song. 2007. "Aching to retire? The rise in the full retirement age and its impact on the social security disability rolls." *Journal of Public Economics*, 91(7): 1327 – 1350.

Engelhardt, Gary V., Jonathan Gruber, and Cynthia D. Perry. 2005. "Social Security and Elderly Living Arrangements: Evidence from the Social Security Notch." *Journal of Human Resources*, XL(2): 354–372.

Evans, William, and Timothy Moore. 2012. "Liquidity, Economic Activity, and Mortality." *The Review of Economics and Statistics*, 94(2): 400–418.

Fahle, Sean, Kathleen McGarry, and Jonathan Skinner. 2016. "Out-of-Pocket Medical Expenditures in the United States: Evidence from the Health and Retirement Study." *Fiscal Studies*, 37(3-4): 785–819.

Ferman, Bruno, and Cristine Pinto. 2019. "Inference in Differences-in-Differences with Few Treated Groups and Heteroskedasticity." *The Review of Economics and Statistics*, 101(3): 452–467.

Fitzpatrick, Maria D., and Timothy J. Moore. 2018. "The mortality effects of retirement: Evidence from Social Security eligibility at age 62." *Journal of Public Economics*.

Frakt, Austin B. 2011. "How much do hospitals cost shift? A review of the evidence." *The Milbank Quarterly*, 89(1): 90–130.

Gardner, Jonathan, and Andrew J. Oswald. 2007. "Money and mental wellbeing: A longitudinal study of medium-sized lottery wins." *Journal of Health Economics*.

Gelber, Alexander M, Adam Isen, and Jae Song. 2016. "The effect of pension income on elderly earnings: Evidence from social security and full population data."

Gelber, Alexander, Timothy Moore, and Alexander Strand. 2018. "Disability insurance income saves lives." *SIEPR Working Papers*.

Gelman, Andrew, and Guido Imbens. 2019. "Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs." *Journal of Business & Economic Statistics*, 37(3): 447–456.

Goda, Gopi Shah, Ezra Golberstein, and David C. Grabowski. 2011. "Income and the utilization of long-term care services: Evidence from the Social Security benefit notch." *Journal of Health Economics*, 30(4): 719–729.

Golberstein, Ezra. 2015. "The Effects of Income on Mental Health: Evidence from the Social Security Notch." *The journal of mental health policy and economics*, 18(1): 27–37.

Grabowski, David C. 2007. "Medicare and Medicaid: Conflicting Incentives for Long-Term Care." *The Milbank Quarterly*, 85(4): 579–610.

Gross, Tal, and Jeremy Tobacman. 2014. "Dangerous Liquidity and the Demand for Health Care." *Journal of Human Resources*, 49(2): 424–445.

Gross, Tal, Timothy Layton, and Daniel Prinz. 2020. "The Liquidity Sensitivity of Healthcare Consumption: Evidence from Social Security Payments." National Bureau of Economic Research Working Paper 27977.

Grossman, Michael. 2000. "The Human Capital Model." In *Handbook of Health Economics.* Vol. 1 of *Handbook of Health Economics*, , ed. Anthony J. Culyer and Joseph P. Newhouse, 347 – 408. Elsevier.

Hackmann, Martin B. 2019. "Incentivizing better quality of care: The role of Medicaid and competition in the nursing home industry." *American Economic Review*, 109(5): 1684–1716.

Hall, Robert E., and Charles Jones. 2007. "The Value of Life and the Rise in Health Spending." *The Quarterly Journal of Economics*, 122(1): 39–72.

Handwerker, Elizabeth Weber. 2011. "What can the Social Security Notch tell us about the impact of additional income in retirement?" *Journal of Economic & Social Measurement*, 36(1/2): 71–92.

Jacobson, Mireille, Maria Kogelnik, and Heather Royer. 2020. "Holiday, Just One Day Out of Life: Birth Timing and Post-natal Outcomes." *National Bureau of Economic Research Working Paper Series*, No. 27326.

Jensen, Robert T, and Kaspar Richter. 2004. "The health implications of social security failure: evidence from the Russian pension crisis." *Journal of Public Economics*, 88(1): 209 – 236.

Kaestner, Robert. 2013. "The Grossman model after 40 years: a reply to Peter Zweifel." *The European Journal of Health Economics*, 14(2): 357–360.

Kaestner, Robert, and Anthony T. Lo Sasso. 2015. "Does seeing the doctor more often keep you out of the hospital?" *Journal of Health Economics*, 39: 259 – 272.

Kim, Beomsoo, and Christopher J. Ruhm. 2012. "Inheritances, health and death." *Health Economics*.

Koka, Katerina, Audrey Laporte, and Brian Ferguson. 2014. "Theoretical Simulation in Health Economics: An Application to Grossman's Model of Investment in Health Capital." Canadian Centre for Health Economics Working Papers 140010.

Kopczuk, Wojciech, and Jae Song. 2008. "Stylized Facts and Incentive Effects Related to Claiming of Retirement Benefits Based on Social Security Administration Data." University of Michigan, Michigan Retirement Research Center Working Papers wp200.

Kuhn, Andreas, Stefan Staubli, Jean-Philippe Wuellrich, and Josef Zweimüller. 2019. "Fatal attraction? Extended unemployment benefits, labor force exits, and mortality." *Journal of Public Economics*, 104087.

Li, Xiaoyan, and Nicole Maestas. 2008. "Does the rise in the full retirement age encourage disability benefits applications? Evidence from the health and retirement study." *Michigan Retirement Research Center Research Paper*, , (2008-198).

Liebman, Jeffrey B. 2015. "Understanding the Increase in Disability Insurance Benefit Receipt in the United States." *Journal of Economic Perspectives*, 29(2): 123–50.

Liebman, Jeffrey B., and Erzo F. P. Luttmer. 2015. "Would People Behave Differently If They Better Understood Social Security? Evidence from a Field Experiment." *American Economic Journal: Economic Policy*, 7(1): 275–99.

Lindahl, Mikael. 2005. "Estimating the Effect of Income on Health and Mortality Using Lottery Prizes as an Exogenous Source of Variation in Income." *Journal of Human Resources*.

Lindner, Stephan. 2016. "How Do Unemployment Insurance Benefits Affect the Decision to Apply for Social Security Disability Insurance?" *Journal of Human Resources*, 51(1): 62–94.

MacKinnon, James G., and Matthew D. Webb. 2017. "Wild Bootstrap Inference for Wildly Different Cluster Sizes." *Journal of Applied Econometrics*, 32(2): 233–254.

Maestas, Nicole. 2019. "Identifying Work Capacity and Promoting Work: A Strategy for Modernizing the SSDI Program." *The ANNALS of the American Academy of Political and Social Science*, 686(1): 93–120.

McInerney, Melissa, Jennifer M. Mellor, and Lauren Hersch Nicholas. 2013. "Recession depression: Mental health effects of the 2008 stock market crash." *Journal of Health Economics*.

MedPAC. 2018. "Data Book: Medicare beneficiary and other payer financial liability."

Meer, Jonathan, Douglas L Miller, and Harvey S Rosen. 2003. "Exploring the health–wealth nexus." *Journal of Health Economics*, 22(5): 713 – 730.

Moran, John R, and Kosali Ilayperuma Simon. 2006. "Income and the Use of Prescription Drugs by the Elderly: Evidence from the Notch Cohorts." *Journal of Human Resources*, XLI(2): 411–432.

Mueller, Andreas I., Jesse Rothstein, and Till M. von Wachter. 2016. "Unemployment Insurance and Disability Insurance in the Great Recession." *Journal of Labor Economics*, 34(S1): S445–S475.

Persson, Petra. 2020. "Social Insurance and the Marriage Market." *Journal of Political Economy*, 128(1): 252–300.

Philipson, Tomas J., and Gary S. Becker. 1998. "OldAge Longevity and MortalityContingent Claims." *Journal of Political Economy*, 106(3): 551–573.

Reid, Rachel, Cheryl Damberg, and Mark W. Friedberg. 2019. "Primary Care Spending in the Fee-for-Service Medicare Population." *JAMA Internal Medicine*, 179(7): 977–980.

Ridley, Matthew W, Gautam Rao, Frank Schilbach, and Vikram H Patel. 2020. "Poverty, Depression, and Anxiety: Causal Evidence and Mechanisms." National Bureau of Economic Research Working Paper 27157.

Roberts, Eric T., Alexandra Glynn, Noelle Cornelio, Julie M. Donohue, Walid F. Gellad, J. Michael McWilliams, and Lindsay M. Sabik. 2021. "Medicaid Coverage 'Cliff' Increases Expenses And Decreases Care For Near-Poor Medicare Beneficiaries." *Health Affairs*, 40(4): 552–561. PMID: 33819086.

Robst, John. 2006. "Estimation of a Hedonic Pricing Model for Medigap Insurance." *Health services research*, 41(6): 2097–2113.

Salm, Martin. 2011. "The Effect of Pensions on Longevity: Evidence From Union Army Veterans." *The Economic Journal*, 121(552): 595–619.

Schmidt, Lucie, Lara D. Shore-Sheppard, and Tara Watson. 2020. "The Impact of the ACA Medicaid Expansion on Disability Program Applications." *American Journal of Health Economics*, 6(4): 444–476.

Schwandt, Hannes. 2018. "Wealth shocks and health outcomes: Evidence from stock market fluctuations." *American Economic Journal: Applied Economics*.

Shepard, Mark, Katherine Baicker, and Jonathan Skinner. 2020. "Does One Medicare Fit All? The Economics of Uniform Health Insurance Benefits." *Tax Policy and the Economy*, 34(1): 1–41.

Skinner, Jonathan, and Douglas Staiger. 2015. "Technology Diffusion and Productivity Growth in Health Care." *The Review of Economics and Statistics*, 97(5): 951–964.

Snyder, Stephen E., and William N. Evans. 2006. "The effect of income on mortality: Evidence from the social security Notch." *Review of Economics and Statistics*, 88(3): 482–495.

SSA. 2015. "Population Profile: Never Beneficiaries, Aged 60–89."

Stepner, Michael, and Allan Garland. 2017. "CALIPMATCH: Stata module for caliper matching without replacement." *Statistical Software Components, Boston College Department of Economics*.

Tsai, Yuping. 2015. "Social security income and the utilization of home care: Evidence from the social security notch." *Journal of Health Economics*, 43: 45 – 55.

Van Kippersluis, Hans, and Titus J. Galama. 2014. "Wealth and health behavior: Testing the concept of a health cost." *European Economic Review*.

Wingender, Philippe, and Sara LaLumia. 2017. "Income effects on maternal labor supply: Evidence from child-related tax benefits." *National Tax Journal*, 70(1): 11–52.

Zhao, Kai. 2014. "Social security and the rise in health spending." *Journal of Monetary Economics*, 64: 21 – 37.

Zheng, Nan Tracy, Sonja Hoover, and Zhanlian Feng. 2014. "Effect of State Medicaid Payment Policies for Medicare Cost Sharing on Access to Care for Dual Eligibles."