Essays on Applied Public Finance

Jeremy Everett McCauley

Thesis submitted for a Ph.D. in Economics.

Department of Economics
UCL
May 2018

Declaration

I, Jeremy Everett McCauley, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Jeremy McCauley, May 2018

Statement of Conjoint Work

Two out of my three primary chapters that form this thesis involve conjoint work, as specified below.

Chapter II "Medical Spending of the US Elderly" is conjoint work with Mariacristina De Nardi, Eric French, and John Bailey Jones. Overall, my contribution amounts to two thirds of the total paper. A version of this essay has been published as:

> De Nardi, M., French, E., Jones, J.B. and McCauley, J., 2016. Medical spending of the US elderly. *Fiscal Studies*, 37(3-4), pp.717-747.

Chapter III "The Effect of Disability Insurance Receipt on Mortality" is conjoint work with Bernard Black, Eric French, and Jae Song. Overall, my contribution amounts to two thirds of the total paper.

Acknowledgements

Throughout my PhD I have been helped, encouraged and supported by countless individuals, and I apologize for not being able to name and thank each person individually.

First and foremost, I would like to thank my principal supervisor Eric French, who has provided me with invaluable guidance and insights, given me many opportunities, and has taught me a lot about research. I would also like to thank my other supervisors, Fabien Postel-Vinay and Magne Mogstad, for their guidance and advice.

I have been fortunate to work with many co-authors during my PhD, each of whom have taught me something valuable and are owed my thanks. To name a few: Mariacristina De Nardi, John Bailey Jones, Bernard Black, Jae Song, Manudeep Bhuller, Tarjei Havnes, and Elaine Kelly.

The Economics Department at UCL has been a great environment to learn and conduct research. I have greatly benefited from the comments and questions of many faculty members, as well as my fellow PhD students. A special thanks goes to Jeff Rowley for being a good friend throughout the PhD.

Finally, I would like to thank my family and friends for all their care and encouragement. In particular, my wife, Nadine, who has been my main support and the key motivation to finish this PhD.

Funding Acknowledgements

I gratefully acknowledge the funding that made my PhD possible. I was funded by the ESRC through their ‘1+3’ studentship.

Abstract

This thesis is made up of three main essays, each utilizing applied micro-econometric techniques to develop a deeper understanding of issues involving healthcare and benefit receipt.

The first essay (Chapter 2) documents the medical spending of the US population aged 65 and older. It establishes some important facts, including that the government provides over 65% of the elderly's medical expenses. Despite this, the expenses that remain after government transfers are even more concentrated among a small group of people. Thus, government health insurance, while valuable, is far from complete.

The second essay (Chapter 3) estimates the effect of Disability Insurance benefit receipt on mortality. Those receiving benefits receive large cash transfers, and health insurance, but also face work disincentives. Each of these factors could affect mortality. Identifying the overall mortality effect is difficult, however, because those allowed benefits may be unobservably less healthy than those denied. I exploit the random assignment of judges to disability insurance cases to create instrumental variables that address this selection effect, and find considerable heterogeneity in the mortality response.

The final essay (Chapter 4) assesses whether the low observed rate of welfare migrants is due to individuals not knowing the quality of welfare programs in their area. I focus on the elderly in England and use a policy introduced in 2002, where the national government gave a publicly-released rating of the quality of each area's social services (which includes social care). I treat this public release of the ratings as an "information shock" and analyze the distribution of the elderly population across areas before and after the star ratings became public. I use the facts from my empirical analysis to motivate a search model with nested learning, where individuals search for the areas with the best social services and gradually learn about their unobserved quality. Estimates suggest that there is a lot of noise in the learning process, but overall the information release led to increases in utility by affecting migration decisions.

Impact Statement

This thesis has direct policy relevance for agencies providing healthcare and social insurance in both the US and UK. Each chapter presents new empirical evidence that can inform the wider research community and help shape future policy.

Chapter 2 has policy relevance for healthcare provision in the US. I show that despite the government covering 67 per cent of the elderly's total medical expenses, what remains after government transfers is even more concentrated among a small group of people. Therefore, health insurance is still important for this age group.

Chapter 3 has direct policy implications for the Social Security Administration in the US, who provide disability insurance (DI). Given the large and increasing cost of the program, many reform proposals have been put forward to make it more sustainable, including making the disability criteria more stringent. My results suggest that for maximizing the longevity of current DI applicants, the current disability thresholds are at about the right level.

Chapter 4 makes several novel academic contributions. It shows that information can offer a partial explanation for why there is a lack of internal welfare-migration. The chapter presents a tractable model that can provide insights about how beliefs, the speed of learning, and information shocks can affect migration. This model can be applied to a wide range of applications where individuals do not have perfect information about areas. This impacts the current literature on migration (both within Economics and other academic fields), which generally assumes individuals have perfect information about the quality of area-based amenities. If individuals instead have imperfect beliefs, this could affect the estimation of other parameters, such as moving costs. The findings also have implications for policymakers regarding information releases, especially those relating to the quality of local services or benefits. I show that the form that the information release takes, and whether it is personalized or not, can lead to very different outcomes.

I have disseminated this research to academic audiences in the UK. I will further disseminate this research through scholarly publications.

Contents

1 Introduction **13**

Bibliography . . . 15

2 Medical Spending of the US Elderly **16**

2.1 Introduction . . . 16

2.2 Health Care for the Population Aged 65 and Over in the US . . . 18

2.3 The MCBS Data Set . . . 22

2.4 Exp. in the Cross-Section, Over Time and Across Incomes . . . 26

2.5 Medical Spending of the Disabled . . . 35

2.6 Average Medical Spending Over the Life Cycle . . . 36

2.7 Medical Spending Before Death . . . 38

2.8 Conclusion . . . 41

Bibliography . . . 43

2A Appendix . . . 45

3 The Effect of Disability Insurance Receipt on Mortality **50**

3.1 Introduction 50
3.2 Literature Review 52
3.3 The Disability Insurance System 54
3.4 Estimating Equations 57
3.5 Data 61
3.6 Results 66
3.7 The Channels by which DI Affects Mortality 78
3.8 Robustness 84
3.9 Conclusion and Policy Implications 86
Bibliography 88
3A Appendix 92

4 Information and Welfare-Induced Migration **114**

4.1 Introduction 114
4.2 Related Literature 117
4.3 Context, Data and Descriptive Statistics 119
4.4 Identification Strategy 124
4.5 Main Empirical Results 129
4.6 Search Model with Nested Learning 136
4.7 Conclusion 148
Bibliography 149
4A Appendix 153

List of Figures

2.1 Personal Health Care Expenditures for Whole Population 22

2.2 CDFs of Medical Expenditures, Averaged Over 1, 2, and 3 Years. 34

2.3 Average Total Medical Expenditures 37

2.4 Spending in the Last 12 Months of Life, by Expenditure and Payer Type 42

3.1 Allowance at Different Stages of the Applications and Appeals Process 55

3.2 Allowance Rate of ALJs, De-meaned (Left Panel), and De-meaned by Hearing Office and Day (Right panel). 62

3.3 Cumulative Mortality Rates, Allowed versus Denied 66

3.4 Marginal Treatment Effects: Mortality Response by De-meaned Allowance 74

3.5 Estimated 5 Year and 10 Year Mortality Effect of Allowance by the Medical Expenditure for Each Health Condition 82

3.6 Estimated 5 Year and 10 Year Mortality Effect of Allowance by the Mortality Rate for Each Health Condition 83

3.7 Estimated 5 Year and 10 Year Mortality Effect of Allowance by the Labor Supply Response for Each Health Condition 84

3A.1 Mortality of Those Allowed and Denied, by Age at Application 101

3A.2 Mortality of Those Allowed and Denied 101

3A.3 Marginal Treatment Effects: Mortality Response by De-meaned Allowance, Without Covariates . . . 112

3A.4 Marginal Treatment Effects: Mortality Response by De-meaned Allowance, Using Local Polynomial Smoothed Estimates . . . 113

4.1 Revenues and Expenditures of All Local Authorities in England and Wales, 2001 120

4.2 Spending on Personal Social Services for the Elderly by Local Authorities in England, 2001 . . . 121

4.3 Geographic Variations in 2001 . . . 126

4.4 Response to Information Shock in Model . . . 146

4.5 Response to Changes in Size and Form of Information Shock . . . 147

4A.1 Geographic Variation: 2002 Actual and Predicted SSPR Ratings . . . 155

4A.2 Mean Trends, by 2002 SSPR rating . . . 157

4A.3 Placebo Test . . . 159

List of Tables

2.1 Funding Sources of the Elderly's Personal Health Care Expenditures, 2010 20

2.2 Percentage of Personal Health Care Expenditures, by Payer and Expenditure Type: National Data 21

2.3 Medicare and Medicaid Enrolment and Expenditures for the Population Aged 65 and Over: Comparisons 25

2.4 Percentage of Total Expenditures, by Payer, Expenditure Type and Gender: MCBS Data 27

2.5 Medical Spending Percentiles: MCBS 29

2.6 Mean Medical Expenditures, by Spending Quintile and Gender 29

2.7 Income and Medical Expenditures, by Income Quintile and Gender 30

2.8 Mean Medical Expenditure, by Income Quintile and Payer / Expenditure Type . 31

2.9 Correlation of Medical Spending in Year t with Spending in years $t+1$ and $t+2$ 32

2.10 Transition Matrices for Total Medical Expenditure 33

2.11 Measures of the Concentration of Medical Spending Over One, Two and Three years 35

2.12 Mean Monthly Spending of Those with Disabilities, by Age 36

2.13 Medical Spending in the Last Years of Life 39

2A.1 Medicaid Enrolment and Expenditures by Enrollee Spending Percentile: MSIS versus MCBS 45

2A.3 Percentage of Total Expenditures, by Payer and Expenditure Type: MCBS data 46

2A.4 Correlation of Medical Spending in Year t with Spending in Years t+1 and t+2, by Gender . . . 46

2A.5 Transition Matrices for Total Medical Expenditure Excluding Nursing Home Costs 47

2A.6 Transition Matrices for Hospital Expenditure . . . 48

2A.8 Out-of-Pocket Insurance Premia Spending, by Year . . . 48

2A.9 Last Year of Life Spending, From Data . . . 49

3.1 Estimated Percentage of U.S. Deaths Included in the SSA Death Data and Underreporting Correction, by Age Group . . . 64

3.2 Predictors of Allowance and Judge Leniency, Aged 55-64 . . . 68

3.3 First Stage Estimates: Regression of Allowance Rates on Judge Leniency Variable, by Demographics, Aged 55-64 . . . 69

3.4 Estimated Effect of DI Recipiency on Mortality . . . 71

3.5 Estimated Effect of DI Recipiency on Mortality, Aged 55-64, Disaggregated . . . 77

3.6 Key Outcome Differences Between Those Allowed versus Denied . . . 79

3.7 Robustness Checks . . . 85

3A.1 Sample Selection . . . 94

3A.2 Estimated Percentage of US Deaths Included in the SSA Death Data, 1990-2014, by Age Group . . . 95

3A.3 Predictors of Allowance and Judge Leniency, All Ages . . . 96

3A.4 First Stage Estimates: Regression of Allowance Rates on Judge Leniency, By Demographics, All Ages . . . 97

3A.5 Estimated Effect of DI Recipiency on Mortality, Disaggregated cont'd, Aged 55-64 98

3A.6 Estimated Effect of DI Recipiency on Mortality, Disaggregated, All Ages . . . 99

3A.7 Estimated Effect of DI Recipiency on Mortality, Disaggregated cont'd, All Ages . 100

4.1 2001 Mean Local Authority Level Characteristics by 2002 SSPR Rating 125

4.2 Main Results: Distribution of Pensioners Across Areas 129

4.3 Further Robustness and Alternative Methodological Approaches 130

4.4 Distribution of Benefit Claimants Across Areas 132

4.5 Effects on Migration 134

4.6 Model Parameter Values, Pre-set (a) and Estimates (b) 142

4.7 Data and Model Simulated Moments 143

4A.1 Predicting SSPR for Wales 154

4A.2 2001 Mean Local Authority-Level Characteristics by Predicted 2002 SSPR Rating, England and Wales 156

4A.3 P-values related to the null hypothesis $H_0 : \theta_{99} = \theta_{00} = \theta_{01} = 0$ 158

4A.4 Alternative specification 160

Chapter 1

Introduction

In this thesis, I use applied micro-econometric techniques to address a range of questions pertaining to health care and benefit receipt, drawing evidence from the US and UK. There are three primary chapters, all of which address topics that should be of interest to policy makers. Each chapter can be read as a free-standing essay, however there are important connections between each and all highlight the importance of government provision for the health care needs of the elderly and disabled.

The first two chapters focus on health spending and outcomes. Chapter 2 (based on co-authored work with Mariacristina De Nardi, Eric French and John Bailey Jones) documents the medical spending of the US population aged 65 and older. We provide, to the best of our knowledge, the most thorough and accurate description of spending for this age group. We show where the spending goes and who paid for it. We also present facts about spending for the disabled US population, a lot of whom are eligible for government health insurance (Medicare and Medicaid) through the Social Security Disability Insurance (SSDI) and Supplemental Security Income (SSI) programs. Chapter 3 (co-authored with Bernard Black, Eric French and Jae Song) considers whether receipt of government benefits is associated with better health outcomes. Specifically, we estimate the effect of SSDI and SSI benefit receipt on mortality.

In Chapter 2 we find that medical expenses more than double between the ages of 70 and 90 and that they are very concentrated: The top 10% of all spenders are responsible for 52% of medical spending in a given year. In addition, those currently experiencing either very low or very high medical expenses are likely to find themselves in the same position in the future. We also find that the poor consume more medical goods and services than the rich and have a much larger share of their expenses covered by the government. Overall, the government pays for 65% of the elderly's medical expenses. Despite this, the expenses that remain after government transfers are even more concentrated among a small group of people. Thus, government health insurance, while potentially very valuable, is far from complete. Finally, while medical expenses before death can be large, on average they constitute only a small fraction of total spending, both in the aggregate and over the life cycle. Medical expenses before death do not appear to be an important driver of the high and increasing medical spending found in the US.

The statistics presented in Chapter 2 were carefully chosen and constructed to be part of a series of studies examining the properties of individual-level medical spending both across several data sets for a given country and across countries.[1] While the US spends more as a share of GDP on health care than other countries, it is not an outlier in patterns of individual-level medical spending or spending at the end of life.

In Chapter 3 we find considerable heterogeneity in the mortality response to being allowed SSDI and SSI benefits. For marginal recipients, who receive benefits if seen by lenient judges, but would be denied by stricter judges, we find no detrimental effects of being denied on mortality. Instead, we find that for these individuals benefit receipt slightly increases mortality within the first 10 years of benefit receipt, consistent with the view that reduced labor supply from benefit receipt increases mortality. However, Marginal Treatment Effects estimates suggest that benefit receipt reduces mortality for inframarginal benefit recipients, who would receive benefits even if seen by a relatively strict judge. The findings suggest that for maximizing the longevity of current SSDI and SSI applicants, the current disability thresholds are at about the right level.

In the final chapter of the thesis, Chapter 4, I assess a slightly different issue – whether the low observed rate of internal welfare migrants is due to individuals not knowing the quality of welfare programs in their area. Despite much research on the topic, there is little evidence to suggest that geographic differences in the generosity of welfare programs drives migration decisions. The chapter seeks to establish the role of information as part of the explanation for the lack of welfare-induced migration. I assess whether the low observed rate of welfare migrants is due to individuals having little knowledge of the quality of welfare programs in their area compared to other areas. I focus on the elderly in England, where social services provision (which includes social care and can be viewed as a welfare benefit) is decentralized to local authorities. I use a policy introduced in 2002, where the national government gave a publicly-released rating of the quality of each area's social services on a scale from zero to three stars. I treat this public release of the ratings as an "information shock" and analyze the distribution of the elderly population across areas before and after the star ratings became public. My findings show that a one increase in publicly-released star rating is associated with a 0.01 percentage point increase in the percentage of the elderly population living in that area relative to others. This corresponds

[1] To compare the statistics presented in Chapter 2 with different countries and age groups see: Hirth et al. (2016), Fahle et al. (2016) and Pashchenko and Porapakkarm (2016) for other US data sets; Christensen et al. (2016) for Denmark; Gastaldi-Ménager et al. (2016) for France; Karlsson et al. (2016) for Germany; Ibuka et al. (2016) for Japan; Bakx et al. (2016) for the Netherlands; Aragon et al. (2016) and Kelly et al. (2016) for England; Côté-Sergent et al. (2016) for the province of Quebec in Canada; and Chen and Chuang (2016) for Taiwan.

to a 1.3 percent increase in the number of elderly people in the area. I use these facts to motivate a search model with nested learning, where elderly individuals search for the areas with the best social services and gradually learn the quality of these services. The model shows that to generate the observed migration response to the public release of star ratings, the information shock would have to bring individuals' beliefs about the quality of their areas closer to the true mean quality of those areas by 58%. The form that the information shock takes, and whether it is personalized or not, is shown to be important and can lead to different utility and migratory responses.

Bibliography

Aragon, M. J., Chalkley, M., and Rice, N. (2016). Medical spending and hospital inpatient care in england: an analysis over time. *Fiscal Studies*, 37(3-4):405–432.

Bakx, P., O'Donnell, O., and Doorslaer, E. (2016). Spending on health care in the netherlands: not going so dutch. *Fiscal Studies*, 37(3-4):593–625.

Chen, S. H. and Chuang, H. (2016). Recent trends in taiwanese medical spending. *Fiscal Studies*, 37(3-4):653–688.

Christensen, B. J., Gørtz, M., and Kallestrup-Lamb, M. (2016). Medical spending in denmark. *Fiscal Studies*, 37(3-4):461–497.

Côté-Sergent, A., Échevin, D., and Michaud, P.-C. (2016). The concentration of hospital-based medical spending: Evidence from canada. *Fiscal Studies*, 37(3-4):627–651.

Fahle, S., McGarry, K., and Skinner, J. (2016). Out-of-pocket medical expenditures in the united states: Evidence from the health and retirement study. *Fiscal Studies*, 37(3-4):785–819.

Gastaldi-Ménager, C., Geoffard, P.-Y., and De Lagasnerie, G. (2016). Medical spending in france: concentration, persistence and evolution before death. *Fiscal Studies*, 37(3-4):499–526.

Hirth, R. A., Calónico, S., Gibson, T. B., Levy, H., Smith, J., and Das, A. (2016). Long-term health spending persistence among the privately insured in the us. *Fiscal Studies*, 37(3-4):749–783.

Ibuka, Y., Chen, S. H., Ohtsu, Y., and Izumida, N. (2016). Medical spending in japan: an analysis using administrative data from a citizen's health insurance plan. *Fiscal Studies*, 37(3-4):561–592.

Karlsson, M., Klein, T. J., and Ziebarth, N. R. (2016). Skewed, persistent and high before death: Medical spending in germany. *Fiscal Studies*, 37(3-4):527–559.

Kelly, E., Stoye, G., and Vera-Hernández, M. (2016). Public hospital spending in england: evidence from national health service administrative records. *Fiscal Studies*, 37(3-4):433–459.

Pashchenko, S. and Porapakkarm, P. (2016). Medical spending in the us: facts from the medical expenditure panel survey data set. *Fiscal Studies*, 37(3-4):689–716.

Chapter 2

Medical Spending of the US Elderly

2.1 Introduction

We use data from the Medicare Current Beneficiary Survey (MCBS) to document the medical spending of people aged 65 and over in the United States. The medical spending of the US population aged 65 and over is notable for several reasons.

First, the typical elderly American receives far more medical services than those of younger ages. In 2010, average medical expenditures for an American aged 65 or older were 2.6 times the national average.[1] In the same year, people of 65 and older accounted for over one-third of US medical spending. As the population continues to age, this fraction will likely grow. Given that much of the elderly's medical expenditures are financed by the government, their spending is of increasing fiscal importance. A particularly contentious issue is spending at the end of life (Scitovsky, 1984, 1994). Even though studies, such as Hoover et al. (2002), have found that over a quarter of Medicare spending on the elderly is for end-of-life care, proposals to reform this spending have generated scepticism ((Emanuel and Emanuel, 1994) and sometimes strident resistance(Daly, 2009). As our results suggest, end-of-life spending in the US is not unusually high relative to that in other countries.

A second notable feature of this population is that virtually every American aged 65 or older is eligible for Medicare, a government-provided health insurance program. Medicare pays much of

[1]Centers for Medicare and Medicaid Services (2015).

the cost of short hospital stays, doctor visits and, since 2006, pharmaceuticals. This is in sharp contrast to the younger population. The majority of Americans younger than 65 are covered through employer-provided health insurance, but many others are covered by privately-purchased health insurance or government-provided insurance. Moreover, because privately-purchased insurance can be expensive, and because the eligibility criteria for government insurance are strict for the non-elderly, many people younger than 65 are uninsured. A number of studies, such as Wagstaff and Van Doorslaer (2000), suggest that access to health care in the US is unequal across the income distribution.[2] This inequality is likely more pronounced among the younger population than among the elderly, where Medicare mitigates disparities in health care access. In fact, the US health care system for those aged 65 and over looks much more similar to health care systems elsewhere in the OECD. Perhaps unsurprisingly, many of the patterns of medical spending we document in this paper are similar to the patterns documented for other countries. More surprisingly, however, many of the patterns we document in this essay are similar to the patterns documented in Pashchenko and Porapakkarm (2016) for the under-65 US population. For both the over- and under-65 populations in the US and elsewhere, there is a modest negative correlation between income and total medical spending.

A third reason to study medical spending among retirees is that medical expenses provide an important motive for retirement saving (De Nardi et al., 2010). This saving not only affects wages and economic growth, but is an important policy concern in its own right. Given the policy debates surrounding the financing of medical spending in the US, better knowledge of the patterns of existing medical spending risk faced by the elderly is of value. Despite near-universal health insurance for the population aged 65 and over, we show that many elderly people in the US still face the risk of catastrophic medical spending.

The MCBS links the administrative Medicare records to survey information from households. In addition to high-quality data on Medicare payments, the MCBS contains spending data for other payers from its survey component.

We find that medical expenses more than double between ages 70 and 90, with most of the increase coming from nursing home spending. Medical expenses are very concentrated: The top 10 per cent of all spenders are responsible for 52 per cent of medical spending in a given year. We also find that those currently experiencing either very low or very high medical expenses are likely to find themselves in the same position in the future. These features of the data are consistent with individuals or households facing a small risk of large medical expenses, which, once incurred, tend to be persistent over time. Because it is hard to self-insure against such risks by saving, they may be quite costly for consumers, especially if there are frictions in private health insurance markets. Government insurance mitigating these risks may thus be very valuable to consumers. This notwithstanding, and despite the fact the government pays for 67 per cent

[2] More precisely, the authors review the literature on inequalities in the delivery of health care.

of the elderly's medical expenses, the expenses that remain after government transfers are even more concentrated among a small group of people. Hence, government health insurance, while potentially very valuable, is far from complete. This is in part because the government's Medicaid program is the payer of last resort, contributing only after private funding has been exhausted. As a result, even though the poor on average consume more medical goods and services than the rich, they are responsible for a much smaller share of their costs. Finally, while medical expenses before death can be large, on average they constitute only a small fraction of total spending, both in the aggregate and over the life cycle. Therefore, medical expenses before death do not appear to be an important driver of the high and increasing medical spending found in the US.

The rest of this essay is organised as follows. Section 2.2 briefly describes the health care system for older Americans. Section 2.3 describes the MCBS data and compares them with administrative data. Section 2.4 documents the concentration of medical expenditures, both within a single year and across multiple years, and the concentration of medical spending across the income distribution. Section 2.5 considers the medical expenditures of the disabled. Section 2.6 shows the evolution of medical expenses and their payers during the retirement period. Section 2.7 presents new estimates of medical spending in the last three years of life and Section 2.8 concludes.

2.2 Health Care for the Population Aged 65 and Over in the US

2.2.1 Institutional Background

With some exceptions, US health care is privately provided. Most US hospitals are run either by non-profit institutions, such as universities or religious organisations, or by private for-profit companies. The employees of those hospitals, including doctors and nurses, are then paid by the hospitals. Hospitals, doctors and other health care providers are largely free to charge what they wish for their services. However, health care insurers (public and private) usually negotiate prices for their insurees.

We define a payer of health care as the final payer of billed medical goods and services. Thus, we count a payment by a private insurer as a private insurance payment, even though an individual paid insurance premiums to obtain these insurance services. So if an individual paid \$1 for insurance that paid \$1 to a hospital, we will count 'out-of-pocket spending' as 0 and payments by private insurance as \$1.

The main payer of health care amongst the elderly is Medicare, a federal program that provides health insurance to almost every person aged 65 or over. Individuals covered by Medicare have the option of traditional Medicare, where Medicare pays the providers, or Medicare Advantage, where Medicare provides payments to health maintenance organisations, which then provide care. Under traditional Medicare, the government sets a schedule of payments for most services. In order to discourage the over-provision of health care services, many health care treatments performed by hospitals are paid on the basis of the diagnosis rather than the treatment. Traditional Medicare pays for the great majority of the cost of short-term hospital stays, 80 per cent of the cost of doctor visits and, since 2006, most of the costs associated with pharmaceuticals. Medicare Advantage pays for close to 100 per cent of the cost of hospital stays, doctor visits and pharmaceuticals.

Many older individuals have private insurance plans that cover medical expenses not covered by Medicare, such as the residual share of the costs of doctor visits. However, some forms of care are largely uninsured by either Medicare or private health insurance, with the most important category being nursing home spending. A large share of nursing home costs is paid out-of-pocket. Because nursing home stays are expensive – of the order of $77,000–88,000 a year in 2014 – most individuals will be impoverished by a long nursing home stay. Those made financially destitute will be covered by Medicaid, a means-tested program that is run jointly by the federal and state governments.[3] In 2013, around 29 per cent of nursing home costs were paid out-of-pocket, while around 30 per cent were covered by Medicaid. Medicaid covers almost all the nursing home costs of poor, old recipients. More generally, Medicaid ends up financing 63 per cent of nursing home residents.[4] In 2009, 74 per cent of Medicaid's transfers to the elderly were for long-term care.[5] In large part because of its role in funding nursing home care, Medicaid is the second most important public health insurance program for the elderly in the US. Nonetheless, Medicaid is the payer of last resort, contributing only after private funding and Medicare support have been (nearly) exhausted.

The National Health Expenditure Accounts (NHEA), maintained by the Centers for Medicare and Medicaid Services (CMS), document how much is being spent on each type of health care service, as well as the payers of those services.[6] Tables 2.1 and 2.2 use these data to summarise the sources and uses of personal health care spending. Personal health care spending measures the total amount spent on all treatments for all individuals. It excludes government administration, government public health activities, and investment. We focus on personal health care expenditure since it is the concept that the MCBS data are designed to measure. Moreover, the bulk – in 2013, 85 per cent – of total national health care expenditures go to personal health care.

Table 2.1 shows how the personal health care expenditures of the elderly were funded in 2010, the most recent year the age-specific data are available in the CMS data set. Each column of the table corresponds to a particular type of service: hospital care; professional services such as doctor and dental visits; nursing home care; drugs; and other.[7] Each row corresponds to a payer:

[3] Gardner and Gilleskie (2006) and De Nardi et al. (2016) document many important aspects of Medicaid insurance in old age.
[4] Kaiser Family Foundation (2013), Figure 4.
[5] Kaiser Family Foundation (2013), Figure 12.
[6] Centers for Medicare and Medicaid Services (2015).
[7] 'Other' means 'Other payers and programs', which includes Department of Defense, Department of Veterans

	Type of Expenditure					
Payer	Hospitals	Professional Services	Nursing Care	Retail Drugs	Other	All
Out-of-pocket	1.1%	9.4%	28.2%	18.6%	27.9%	13.2%
Private Insurance	13.4%	18.6%	7.8%	23.4%	3.8%	13.3%
Medicaid	6.8%	2.1%	29.7%	1.3%	21.9%	11.1%
Medicare	69.7%	64.3%	24.3%	52.8%	36.5%	54.4%
Other	9.0%	5.6%	10.0%	4.0%	10.0%	8.0%

Source: National Health Expenditure Accounts.

Table 2.1. Funding Sources of the Elderly's Personal Health Care Expenditures, 2010

out-of-pocket; private health insurance; Medicaid; Medicare; and other. The table shows the fraction of each expenditure subtotal paid by each payer. For example, the first column shows that only 1 per cent of the costs of hospital care are paid out-of-pocket, while almost 70 per cent of the costs are covered by Medicare. In fact, Medicare is the largest payer for every type of expenditure with the exception of nursing home care. The final column of Table 2.1 shows that Medicare covers well over half of the elderly's medical expenditures. Private health insurance, Medicaid and out-of-pocket expenditures each cover between 11 and 13 per cent of the total.

2.2.2 Trends in Health Care Expenditures

Table 2.2 compares the spending of the elderly with that of the general population. The top panel shows the shares of medical spending covered by different payers. The first column in this panel repeats the final column of Table 2.1. The second column of Table 2.2 shows the equivalent to the first column for the under-65 population and the remaining four columns show results for the entire US population for 1970, 1990, 2010 and 2013. While Medicare pays a much bigger share of health care expenditures for the population aged 65 and over than for the population as a whole, in 2010 the share spent out-of-pocket barely falls after age 64. Instead, the rise in Medicare expenditures after age 64 mostly displaces private insurance expenditures. The second panel of Table 2.2 shows the shares of total medical spending across service categories. The biggest changes in expenditure shares for those aged 65 and over are a rise in nursing home care and a fall in professional services such as doctor visits.

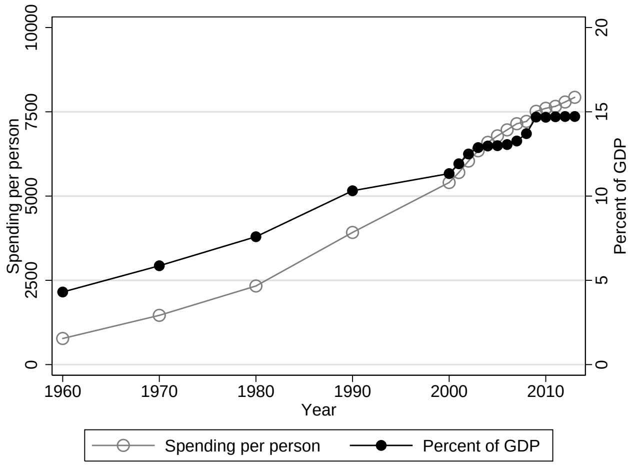
Affairs, worksite health care, other private revenues, Indian Health Service, workers' compensation, general assistance, maternal and child health, vocational rehabilitation, other federal programs, Substance Abuse and Mental Health Services Administration, and other state and local programs.

	65+ population	Under 65s	Whole population			
	2010	2010	1970	1990	2010	2013
Fraction by Payer						
Out-of-pocket	13.2%	14.3%	39.6%	22.5%	13.9%	13.7%
Private Insurance	13.3%	45.2%	22.2%	33.3%	34.4%	34.3%
Medicaid	11.1%	19.5%	7.9%	11.3%	16.7%	16.6%
Medicare	54.4%	5.9%	11.6%	17.4%	22.3%	22.3%
Other	8.0%	15.1%	18.7%	15.6%	12.7%	13.0%
Fraction by Type of Expenditure						
Nursing Care	16.2%	1.5%	6.3%	7.3%	6.5%	6.3%
Hospitals	35.3%	38.0%	43.1%	40.6%	37.1%	38.0%
Professional Services	23.2%	35.9%	31.4%	33.7%	31.6%	31.5%
Retail Drugs	10.3%	12.4%	8.7%	6.5%	11.7%	11.0%
Other	15.0%	12.1%	10.5%	11.9%	13.1%	13.2%
Total Personal Health Care Expenditures ($ billions)	800	1,550	310	990	2,350	2,500

Note: Dollar values are adjusted to 2014 dollars.
Source: National Health Expenditure Accounts.

Table 2.2. Percentage of Personal Health Care Expenditures, by Payer and Expenditure Type: National Data

As is well known, the US spends large and increasing amounts on medical care. The bottom panel of Table 2.2 shows that in 2013 personal health care expenditures amounted to $2.5 trillion in 2014 dollars, representing 14.7 per cent of GDP. This translates to $7,930 per person. Figure 2.1 shows personal health care spending in the US, both per person and as a percentage of GDP, from 1960 to 2013. By either measure, health care spending has risen dramatically. Table 2.2 reveals that while the shares of spending going to each category have been fairly stable over time, the share of spending covered out-of-pocket has fallen by nearly two-fifths. For most of this period, per-capita expenditures on the elderly have grown more rapidly than expenditures on the young. Meara et al. (2004) calculate that in 1963, average expenditures in the population aged 65 and over were 2.4 times the expenditures of those under 65. In 2000, the ratio had risen to 4.4. The authors also find, however, that this trend has reversed in recent decades, and per-capita expenditures on the elderly are now growing more slowly than those on the young. The spending ratio calculated with the National Health Expenditure Accounts has fallen from 3.7 in 2002 to 3.4 in 2010.



Note: Personal Health Care Expenditures are displayed per person (2014 Dollars, left scale), and as a Percentage of GDP (right scale).

Figure 2.1. Personal Health Care Expenditures for Whole Population

2.3 The MCBS Data Set

2.3.1 Description

Our principal data source is the 1996 to 2010 waves of the Medicare Current Beneficiary Survey. The MCBS is a nationally representative survey of disabled and elderly (aged 65 and over) Medicare beneficiaries.[8] Although the sample misses elderly individuals who are not Medicare beneficiaries, virtually everyone aged 65 and over is a beneficiary. The survey contains an over-sample of beneficiaries older than 80 and disabled individuals younger than 65. We exclude disabled individuals younger than 65 (apart from in section 2.5), and use population weights throughout, unless specified.

MCBS respondents are interviewed up to 12 times over a four-year period and are asked about (and matched to administrative data on) health care utilisation over three of the four years, forming panels on medical spending for up to three years. We aggregate the data to an annual level. These sample selection procedures leave us 66,790 different individuals who contribute 152,193 person-year observations.

The MCBS's unit of analysis is an individual. Respondents are asked about health status, income, health insurance, and health care expenditures paid out of pocket, by Medicaid, by Medicare, by

[8] Adler and Phil (1998) describes the MCBS in some detail. The MCBS sourcebook series (Centers for Medicare and Medicaid Services, multiple years) provides annual data summaries.

private insurance and by other sources. The MCBS survey data are then matched to Medicare records. For this reason, the medical spending data are of particularly high quality.

The key variable of interest is medical spending. This includes the cost of hospital stays, doctor visits, pharmaceuticals, nursing home care and other long-term care. The MCBS's medical expenditure measures are created through a reconciliation process that combines survey information with Medicare administrative files. As a result, the MCBS contains accurate data on Medicare payments and fairly accurate data on out-of-pocket, Medicaid, and other insurance payments. Out-of-pocket expenses include hospital, doctor and other bills paid out-of-pocket, but do not include insurance premiums paid out-of-pocket. Because the MCBS includes information on people who enter a nursing home or die, its medical spending data are very comprehensive.

In the MCBS, individuals are asked to report '... your and your spouse's total income before taxes during the past 12 months'. Respondents are asked to provide an income interval, rather than an exact dollar amount. The MCBS income measure appears to include household income, including transfer and asset income. In contrast, medical spending and most other variables in the MCBS are measured at the individual level. To make the income data compatible with the other variables, we rescale household income by standardised household size: [9]

$$Standardised\ household\ income = \frac{Total\ household\ income}{(Number\ of\ adults)^{0.7}}.$$

When taking logs, we bottom-code income and medical spending.[10] We adjust all dollar amounts to 2014 dollars using the personal consumption expenditure index.

De Nardi et al. (2016) benchmark the MCBS data to survey data from the Assets and Health Dynamics of the Oldest Old (AHEAD) data set and find that the MCBS and AHEAD match up well against each other, with the MCBS possibly being more accurate.[11]

[9] Michael and Citro (1995).

[10] Some people have zero medical spending, and so the log of their medical spending is undefined. To address this problem, we bottom-code the medical spending data whenever we take logs. We treat all values of medical spending that are less than 10 per cent of the mean of medical spending as equal to 10 per cent of the mean. So, if someone has medical spending equal to 5 per cent of the mean, we recode their medical spending as 10 per cent of the mean. We bottom-code income in the same way as medical spending.

[11] The authors show that, conditional on income quintile, average total income (including asset and other non-annuitised income), out-of-pocket medical spending and Medicaid recipiency rates in the AHEAD data are slightly lower than their counterparts in the MCBS data. The MCBS uses administrative data to identify Medicaid recipiency, which greatly reduces under-reporting problems. In addition, the MCBS imputes forgotten out-of-pocket expenses if Medicare had to pay a share of the total cost. In contrast, AHEAD uses a more detailed set of questions to measure out-of-pocket medical spending, including 'unfolding brackets', where respondents can give ranges for their spending instead of a point estimate or 'don't know' as in the MCBS.

2.3.2 Comparisons with Administrative Data

Although there is no high-quality administrative information for out-of-pocket and private insurance payments for the population aged 65 and over, we can compare the MCBS data with administrative data from the Medicare and Medicaid programs.

The first set of columns in Table 2.3 compares Medicare enrolment and average Medicare expenditures in the MCBS with the corresponding values in the aggregate data from the Census Bureau. It shows that, when using population weights, the number of Medicare beneficiaries and expenditures per beneficiary line up closely with the aggregate statistics. Over the 1996–2010 period, MCBS Medicare enrolment for the population aged 65 and over averages 36.7 million, only 3 per cent more than the average of 35.8 million from aggregate data. Over the same period, expenditures per beneficiary in the MCBS average $7,670, 14 per cent smaller than the value of $8,970 in the official statistics.[12] The expenditure match weakens over time, as mean expenditures in the MCBS go from 92 per cent of the data in 1996 to 81 per cent of the data in 2010. We are not sure of the source of the decline in the quality of the match.

The MCBS uses administrative data to determine whether an individual is receiving Medicaid benefits, but it does not have administrative data on the value of those payments. In order to assess the quality of the Medicaid expenditure data in the MCBS, we benchmark them against administrative data from the Medicaid Statistical Information System (MSIS). Table 2.3 shows that the MCBS also accurately measures the share of the Medicare population aged 65 and over receiving Medicaid payments, after adjusting the MSIS estimates to include only those receiving Medicare, and not the full population, a group sometimes known as 'dual-eligibles'.[13] According to MCBS data, there were on average 5.26 million aged Medicaid beneficiaries over the 1999–2010 period, versus 5.31 million aged Medicaid beneficiaries in the MSIS data, an underestimate of 1 per cent. However, for the same period, MCBS Medicaid payments for the population aged 65 and over are on average 29 per cent smaller than the MSIS data suggest. Part of this difference is explained by the MCBS payment data not including Medicaid payments to Medicare. After adjusting the MCBS estimates to also include estimated Medicaid contributions to Medicare, the MCBS captures 79 per cent of all Medicaid spending. As with the Medicare data, the discrepancy

[12] Medicare statistics come from the United States Census Bureau, The 2012 Statistical Abstract, Health & Nutrition. Medicare Part D payments are not disaggregated by age, so we assume 84 per cent of all Part D payments are for the 65-and-over age group, the same percentage as for Parts A and B.

[13] In order to construct Table 2.3, we made a number of adjustments to the raw counts in both the MSIS and the MCBS. Most importantly, we adjusted the MSIS to account for the fact that, being a sample of Medicare beneficiaries, the MCBS does not include those not receiving Medicare. About 98 per cent of Americans aged 65 and over receive Medicare. However, based on our analysis of 2008 MSIS data, of those on Medicaid and aged 65 and over, only 92 per cent are also receiving Medicare, and they make up 93 per cent of total Medicaid spending for the population aged 65 and over. These estimates are similar to those in Young et al. (2012). Thus we multiplied the Medicaid population and payments in the MSIS by 0.92 and 0.93 respectively. Medicaid MSIS statistics are located at https://www.cms.gov/Research-Statistics-Data-and-Systems/Computer-Data-and-Systems/MedicaidDataSourcesGenInfo/MSIS-Tables.html.

	Medicare				Medicaid (Medicaid and Medicare Dual Eligibles Only)				
	MCBS		U.S. Census Bureau		MCBS			MSIS	
Year	Population (millions)	Mean Exp. ($)	Population (millions)	Mean Exp. ($)	Population (millions)	Mean Exp. ($)	Adj. Mean Exp.[a] ($)	Population (millions)	Mean Exp. ($)
1996	34.8	6,430	33.4	6,970	4.71	9,800	10,510	–	–
1997	34.8	6,480	33.7	7,380	4.68	9,830	10,550	–	–
1998	34.9	6,170	33.8	7,380	4.64	9,590	10,300	–	–
1999	35	6,450	33.9	7,160	4.65	9,380	10,110	4.48	12,490
2000	35.1	6,650	34.3	7,120	4.80	9,830	10,540	4.60	13,270
2001	35.5	7,030	–	–	4.90	9,990	10,760	4.76	13,730
2002	35.9	7,490	–	–	5.09	10,100	10,940	5.12	13,740
2003	36.2	7,510	35	8,240	5.16	9,810	10,690	5.43	13,530
2004	36.3	7,690	35.4	8,590	5.39	9,560	10,530	5.47	14,040
2005	36.6	7,880	35.8	9,210	5.51	9,940	11,050	5.59	14,120
2006	36.9	8,640	36.3	9,910	5.38	8,760	9,980	5.66	12,340
2007	37.8	8,990	37	10,890	5.38	8,940	10,200	5.49	12,220
2008	38.7	9,110	37.9	10,750	5.46	8,760	10,010	5.58	12,410
2009	39.6	9,210	38.8	11,460	5.68	7,980	9,240	5.64	12,240
2010	40.6	9,340	39.6	11,530	5.73	8,820	10,240	5.85	12,560

[a] Adjusted mean expenditure is expenditure plus estimated Medicaid payments to Medicare Part B.
Note: '-' denotes that the data are unavailable. MSIS is the Medicaid Statistical Information System. See footnotes 14 and 15 for details on construction of the aggregate Medicare and Medicaid statistics. Adjusted to 2014 dollars.

Table 2.3. Medicare and Medicaid Enrolment and Expenditures for the Population Aged 65 and Over: Comparisons

between the MCBS data and the administrative data is growing over time.[14]

2.4 Expenditures in the Cross-Section, Over Time and Across Incomes

2.4.1 Cross-Sectional Distribution

The top panel of Table 2.4 shows a breakdown of medical spending in the MCBS among payers: out-of-pocket; private insurance; uncollected liabilities for treatments that have not been paid for; and government. The bottom panel shows a breakdown of spending among expenditure categories: nursing home care; hospital spending, by inpatients and outpatients; professional services; pharmaceutical costs; and home help and hospice care. Both panels use data from all waves.

The percentages shown in Table 2.4 are constructed in the same way as those in Table 2.2. Mean spending in each category is divided by the mean of total medical spending, so that the percentages represent the distribution of aggregate medical spending.[15] The percentages calculated for the MCBS are fairly similar to those for the aggregate data for the elderly in 2010 shown in Table 2.2. In both tables, the government covers over 65 per cent of the elderly's medical expenditures. The fraction of costs paid out-of-pocket is higher in the MCBS (19.4 per cent) than in the aggregate statistics (13.2 per cent), while the fraction covered by Medicaid is lower. Drug expenditures are relatively higher in the MCBS. These differences may in part reflect the lack of Medicare drug coverage in the years preceding 2006.

The two most notable differences between men and women in Table 2.4 involve Medicaid and nursing home care. The fraction of medical expenditures covered by Medicaid is nearly twice as large for women as it is for men. Similarly, women spend nearly twice as much on nursing home care as men. This is consistent with Table 2.1, which shows that Medicaid plays a particularly large role in funding nursing home care. Table 2.4 also shows that, in the aggregate, men rely more on Medicare (57.5 per cent) and spend relatively more on hospital care (40.0 per cent) than

[14] In Table 2A.1 in the appendix, we compare the distribution of Medicaid spending in the MCBS with the distribution of Medicaid spending in the MSIS administrative payment data reported by Young et al. (2012).

[15] An alternative approach is to construct spending ratios for each individual and calculate the means of these ratios. Table 2A.3 in the appendix displays these ratios.

	All	Men	Women
Fraction by Payer			
Out-of-Pocket[a]	19.4%	17.2%	21.0%
Private Insurance	12.5%	14.3%	11.3%
Uncollected liabilities	1.5%	1.7%	1.4%
Government	66.5%	66.9%	66.3%
Medicaid	9.4%	6.0%	11.6%
Medicare	54.7%	57.5%	52.8%
Other government	2.5%	3.4%	1.9%
Fraction by Type of Expenditure			
Nursing Home Care	20.6%	14.4%	24.8%
Hospitals	34.7%	40.0%	31.1%
Inpatients	25.8%	29.8%	23.0%
Outpatients	8.9%	10.1%	8.0%
Professional Services	27.1%	28.9%	25.9%
Drugs	13.1%	13.1%	13.2%
Home Health and Hospice	4.5%	3.7%	5.0%
Premium to Total Expenditure Ratio[b]	0.13	0.14	0.13

[a]Includes all medical bills paid out-of-pocket, but does not include insurance premiums.
[b]Total insurance premiums paid by individuals divided by total billed medical expenses.
Note: This table reports total spending in each category divided by total overall medical spending.

Table 2.4. Percentage of Total Expenditures, by Payer, Expenditure Type and Gender: MCBS Data

women (52.8 per cent and 31.1 per cent, respectively). This too is consistent with Table 2.1, which shows that Medicare reimburses nearly 70 per cent of hospital costs.

The last row of Table 2.4 presents the 'premiums to total expenditure ratio', which is calculated by dividing total private insurance premiums by total medical spending. Many elderly individuals have 'Medigap' health insurance plans that pay for items such as Medicare co-payments for doctor visits. As it turns out, this ratio is 13 per cent (for all), which is very close to the 12.5 per cent share of aggregate costs paid for by private insurers, shown in the top panel of the table.

Table 2.5 shows the cross-sectional distribution of medical spending by expenditure type and for the most important payer types, with the results for each spending type sorted by that type's spending. The top panel shows the distributions of total medical spending, total spending excluding nursing home care, and spending on hospitals. Individuals in the top 5 per cent of the total expenditure distribution spend $97,880 apiece, nearly seven times the overall average of $14,120, and constitute nearly 35 per cent of all medical spending. For hospitals, 50 per cent of individuals have almost zero spending and those in the top 5 per cent of the distribution account for over 52 per cent of the spending. The bottom panel of Table 2.5 shows results for out-of-pocket expenditures, Medicare and Medicaid. Although out-of-pocket expenditures are on average much lower than total expenditures, the distribution of out-of-pocket expenditures is more concentrated. Almost half of the out-of-pocket expenditure is made by the top 5 per cent. Even with public and private insurance, out-of-pocket medical expenditure risk is significant.

To examine how the cross-sectional distribution of medical spending differs by gender, we sort medical spending for men and women into quintiles, calculating the quintiles separately for each gender. Table 2.6 shows mean medical spending within each spending quintile. Total expenditures are higher for women than for men at every spending quintile. This difference is largely due to expenditures on nursing home care. Once we exclude nursing home care, men have higher expenditures on average ($11,540 versus $10,970) and in the top two spending quintiles. Men in particular incur higher hospital costs ($5,390 versus $4,530), consistent with Table 2.4. However, the overall shapes of the medical spending distributions are similar across genders.

2.4.2 Distribution by Income

To document how medical spending is distributed by income, Table 2.7 displays mean income and medical expenditures by gender in the MCBS, broken down by income quintile. Low-income

By Expenditure Type

	All		All (excl. nursing homes)		Hospitals	
Spending Percentile	Average Spending	Perc. of total	Average Spending	Perc. of total	Average Spending	Perc. of total
All	14,120	100.0%	11,210	100.0%	4,890	100.0%
95-100%	97,880	34.6%	76,860	34.3%	51,400	52.5%
90-95%	48,890	17.3%	34,360	15.3%	18,880	19.3%
70-90%	20,540	29.1%	16,080	28.7%	6,030	24.6%
50-70%	7,750	11.0%	6,980	12.4%	760	3.1%
0-50%	2,250	8.0%	2,080	9.3%	50	0.1%

By Payer

	Out-of-Pocket		Medicare		Medicaid	
Spending Percentile	Average Spending	Perc. of total	Average Spending	Perc. of total	Average Spending	Perc. of total
All	2,740	100.0%	7,720	100.0%	1,320	100.0%
95-100%	26,930	49.1%	67,560	43.7%	24,980	94.7%
90-95%	6,700	12.2%	28,370	18.4%	1,360	5.2%
70-90%	2,920	21.3%	10,280	26.6%	10	0.1%
50-70%	1,360	9.9%	2,980	7.7%	0	0.0%
0-50%	420	7.6%	550	3.5%	0	0.0%

Note: The results for each expenditure type or payer are sorted by that expenditure type's or payer's spending. Adjusted to 2014 dollars.

Table 2.5. Medical Spending Percentiles: MCBS

	Total Expenditure			Total Expenditure (excl. nursing homes)			Hospitals		
	All	Men	Women	All	Men	Women	All	Men	Women
All	14,120	13,480	14,600	11,210	11,540	10,970	4,900	5,390	4,530
Bottom	740	600	860	670	560	760	0	0	0
Fourth	2,640	2,390	2,840	2,450	2,270	2,580	30	20	40
Third	5,430	5,100	5,670	4,980	4,820	5,090	310	270	330
Second	11,690	11,090	12,170	10,090	10,100	10,090	2,110	2,230	2,030
Top	50,110	48,250	51,440	37,870	39,970	36,330	22,030	24,410	20,260

Note: Adjusted to 2014 dollars.

Table 2.6. Mean Medical Expenditures, by Spending Quintile and Gender

Income	Mean Income			Mean Expenditure		
Quintile	All	Men	Women	All	Men	Women
All	28,280	31,920	25,600	14,120	13,480	14,590
Bottom	8,000	8,700	7,630	17,410	16,180	18,020
Fourth	14,260	16,060	13,250	14,940	14,050	15,890
Third	20,620	23,150	18,890	13,180	12,720	13,380
Second	30,080	33,410	27,650	12,650	12,120	13,050
Top	68,930	79,080	60,910	12,430	12,360	12,620

Income	Mean Expenditure (excl. nursing homes)			Mean Hospitals		
Quintile	All	Men	Women	All	Men	Women
All	11,210	11,540	10,970	4,890	5,390	4,530
Bottom	11,890	12,190	11,650	5,660	6,280	5,300
Fourth	11,490	11,990	11,420	5,370	6,080	5,070
Third	10,990	11,240	10,680	4,840	5,170	4,430
Second	10,900	11,020	10,730	4,430	4,720	4,190
Top	10,800	11,280	10,370	4,180	4,680	3,670

Note: Adjusted to 2014 dollars.

Table 2.7. Income and Medical Expenditures, by Income Quintile and Gender

people consume more medical resources per year. Of course, this higher spending of those with low incomes would be at least partly offset if we accounted for the fact that those at the top of the income distribution live longer than those at the bottom.[16] The observation also does not take into account the fact that those at the top of the income distribution tend to be healthy and have less medical need than those at the bottom of the distribution. What the table shows, however, is that society does spend a fairly large amount of health care resources on low-income people in the US.

The higher spending on the poor consists mostly of greater expenditure on nursing homes. When nursing home care is excluded, the income gradient is much less pronounced. Excluding nursing home expenditures, men consume more medical resources than women at each income quintile. But because women use more nursing home care than men, they have higher total medical spending at every income quintile.

The top panel of Table 2.8 shows how these expenditures are funded. Medicare is an important payer at every income quintile, spending an average of $9,490 on individuals in the lowest income quintile and $6,270 on those in the top one. Out-of-pocket spending is almost constant across

[16] Rettenmaier, 2012.

	All	Bottom	Fourth	Third	Second	Top
Income	28,280	8,000	14,260	20,620	30,080	68,930
By Payer						
All Payers	14,120	17,410	14,940	13,180	12,650	12,430
Out-of-Pocket	2,740	2,480	2,780	2,700	2,750	3,000
Medicare	7,720	9,490	8,430	7,460	6,950	6,270
Medicaid	1,320	3,900	1,590	570	260	270
Government Other	360	510	460	320	270	230
Private Insurance	1,760	860	1,450	1,920	2,170	2,420
Uncollected liability	220	170	230	210	230	240
By Expenditure						
All	14,120	17,410	14,940	13,180	12,650	12,430
Nursing Home Care	2,910	5,520	3,450	2,190	1,750	1,630
All (excl. nursing homes)	11,210	11,890	11,490	10,990	10,900	10,800
Professional Services	3,830	3,510	3,580	3,750	4,030	4,270
Drugs	1,860	1,780	1,810	1,860	1,940	1,900
Home Health and Hospice	630	930	740	550	490	450
Hospitals	4,900	5,660	5,370	4,840	4,430	4,180
Inpatient	3,640	4,420	4,020	3,610	3,240	2,920
Outpatient	1,250	1,250	1,350	1,220	1,190	1,250

Note: Adjusted to 2014 dollars.

Table 2.8. Mean Medical Expenditure, by Income Quintile and Payer / Expenditure Type

the income distribution. De Nardi et al. (2016) find that high-income people spend significantly more out-of-pocket than low-income people: singles at the top of the income distribution spend almost twice as much out-of-pocket as those at the bottom. The difference in results comes from the measure of out-of-pocket spending. De Nardi et al. include insurance payments, which is the relevant measure for measuring the spending risk paid by a household, whereas here we include only out-of-pocket payers to providers, since we measure everything from the standpoint of the payer of a particular medical bill. The out-of-pocket payments in this paper are close to those in Fahle et al. (2016). Medicaid pays an average of $3,900 to those in the bottom quintile and only $270 to those in the top one, while private insurance pays an average of $2,420 a year to those in the top quintile and only $860 to those in the bottom one.

The bottom panel of Table 2.8 shows a breakdown of expenditures by service item for each income quintile. Those at the bottom of the income distribution receive more medical services ($17,410) than those at the top ($12,430). Interestingly, this difference seems to be mainly driven by

Total spending in levels	*t+1*	*t+2*	Total spending in logs	*t+1*	*t+2*
All	0.57	0.40	All	0.61	0.53
All (excl. nursing homes)	0.45	0.28	All (excl. nursing homes)	0.56	0.48
Hospitals	0.27	0.19	Hospitals	0.30	0.25

Table 2.9. Correlation of Medical Spending in Year t with Spending in years $t+1$ and $t+2$

nursing home care expenditures. Once nursing home care is excluded, the difference in spending between those at the bottom ($11,890) and those at the top ($10,800) almost disappears.

2.4.3 Correlation Over Time

The distribution of cumulative medical spending depends not only on the distribution of spending at each age but also on its persistence: If an individual has high medical spending this year, how likely are they to have high medical spending next year as well? Relative to the concentration of medical spending over a single year, there has been much less work on the concentration of medical spending over multiple years. Spillman and Lubitz (2000), Lubitz et al. (2003) and Alemayehu and Warner (2004) describe how lifetime expenditures vary by health and time of death, but they do not describe the expenditures' concentration. For the US, most of the research has focused on the persistence of medical spending across multiple years.[17]

Feenberg et al. (1994) and French and Jones (2004) analyse the persistence of *out-of-pocket* medical spending. Table 2.9 shows correlations, both in levels and in logs, of all medical spending, all spending excluding nursing home care, and hospital spending, one and two years apart, i.e. it shows the correlation of medical spending in year t with medical spending in years t+1 and t+2. In our analysis, we include everyone who was alive one year (respectively, two years) after the initial period and we exclude those who died during that time. The correlation of total medical spending between adjacent years is 0.57 in levels and 0.61 in logs. The correlation of total medical spending between years two years apart is 0.40 in levels and 0.53 in logs. Although medical spending is not perfectly correlated over time, its serial correlation is still relatively high two years later. Thus, even on a lifetime basis, there is likely to be a large amount of concentration of medical spending. The correlation drops slightly when nursing home care is excluded, and it drops considerably when we only consider hospital spending. Table 2A.4 in the appendix shows the results disaggregated by gender.

[17] For example, Feenberg and Skinner (1994) and French and Jones (2004).

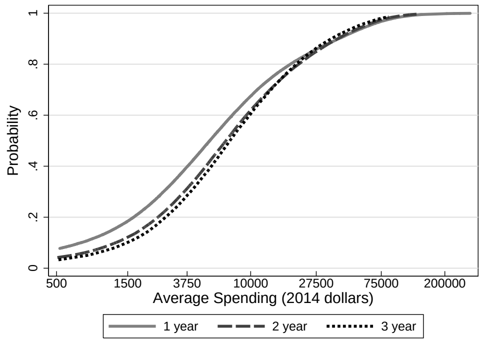
Quintile current year	*Quintile next year*				
	Bottom	Second	Third	Fourth	Top
Bottom	61.9	17.8	8.9	6.5	5.0
Second	24.1	36.6	19.4	12.1	7.8
Third	9.8	25.4	32.3	21.0	11.5
Fourth	6.0	13.6	25.9	34.2	20.3
Top	3.5	6.6	11.9	24.3	53.8

Quintile current year	*Quintile two years ahead*				
	Bottom	Second	Third	Fourth	Top
Bottom	58.3	17.6	10.3	7.5	6.3
Second	26.0	32.2	19.0	12.7	10.2
Third	11.9	25.6	28.3	20.5	13.8
Fourth	7.3	15.3	25.7	31.0	20.6
Top	4.7	8.5	13.5	25.1	48.2

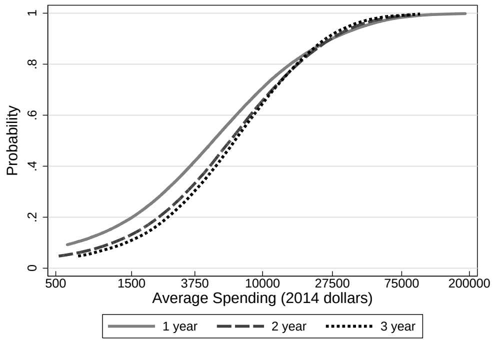
Table 2.10. Transition Matrices for Total Medical Expenditure

Correlation coefficients provide a single linear measure of co-movement. Table 2.10 presents transition matrices, which allow for more flexible relationships across time periods and spending bins. The top panel displays one-year transition probabilities and the bottom panel displays two-year probabilities for movements between the total medical spending quintiles shown in Table 2.6. The row j, column k element of a transition matrix gives the probability that an individual is in spending quintile k in year $t+1$ or $t+2$, given that the individual was in spending quintile j in year t. The tables show that medical spending is concentrated in the top and bottom tails of the distribution. Conditional on being in the top quintile of the medical spending distribution in a given year, there is a 53.8 per cent chance of being in the top quintile in the following year and a 48.2 per cent chance of being in the top quintile in two years' time. Tables 2A.5 and 2A.6 in the appendix report the transition matrices for total expenditures net of nursing home costs and for hospital expenditures, respectively.

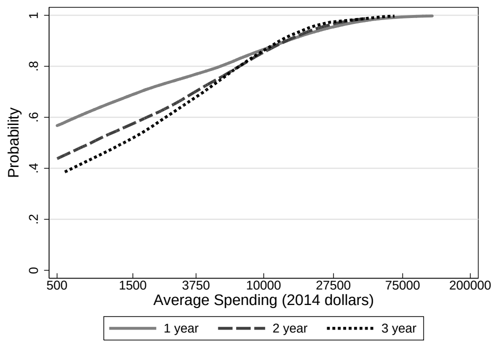
Figure 2.2 displays a more direct measure of how accumulated medical spending is concentrated, by displaying the cumulative distribution function (CDF) for medical spending averaged over one-, two- and three-year periods. Medical spending is highly concentrated even when the data



(a) Total Expenditure



(b) Total Expenditure (excl. nursing)



(c) Hospitals

Figure 2.2. CDFs of Medical Expenditures, Averaged Over 1, 2, and 3 Years.

are averaged across three years. For this to be the case, medical spending must be persistent across time, consistent with the preceding results.

Table 2.11 displays more measures of the concentration of medical spending over different durations – namely, the Gini coefficient[18] and the shares of total medical spending, total spending excluding nursing home costs, and hospital spending for the top 1 per cent and top 10 per cent of spenders. Again, results are shown for one-, two- and three-year periods. Although medical spending becomes less concentrated as the averages cover more years, medical spending remains very concentrated even at three years.

[18] The Gini coefficient is a measure of inequality. It is generally bounded between 0 and 1, where 0 corresponds to perfect equality and 1 corresponds to maximum inequality.

	Medical spending averaged over:		
	1 year	2 years	3 years
	All		
Gini coefficient on medical spending	0.67	0.61	0.58
Perc. spent by top 1% of spenders	11.9%	9.4%	8.7%
Perc. spent by top 10% of spenders	52.0%	45.5%	42.9%
	All (excluding nursing homes)		
Gini coefficient on medical spending	0.64	0.57	0.54
Perc. spent by top 1% of spenders	12.9%	10.0%	8.9%
Perc. spent by top 10% of spenders	49.6%	42.1%	38.7%
	Hospitals		
Gini coefficient on medical spending	0.84	0.77	0.72
Perc. spent by top 1% of spenders	21.4%	16.0%	14.0%
Perc. spent by top 10% of spenders	71.8%	59.1%	53.3%

Table 2.11. Measures of the Concentration of Medical Spending Over One, Two and Three years

2.5 Medical Spending of the Disabled

Table 2.12 presents mean monthly medical spending in the MCBS for those who began receiving Medicare benefits before age 65 using data from the full sample. Those receiving Medicare benefits before age 65 are largely those who qualified because they were eligible for Disability Insurance benefits. In this sense, the data provide us with a portrait of medical spending for the same disabled population at different ages. Table 2.12 provides average total spending and spending paid for by Medicare and Medicaid. It also shows total spending by health condition. To focus on the importance of these health conditions, and not the demographic characteristics of those with those health conditions, we regression adjust the data. Specifically, we regress monthly medical spending on gender, age group, education, race, and whether the individual also receives Medicaid. We then predict medical spending using the regression coefficient on that health condition and also the mean demographics for that sample. For most age/health groups, these demographic adjustments have only a modest effect of the estimates. Table 2.12 reveals that those with endocrine conditions (e.g. diabetes) have the highest medical spending and those with mental retardation have the lowest medical spending. Groups with high medical spending before age 65 also have high medical spending after 65.

Disability*	Mean Monthly Spending ($) Age 25-64	Age 55-64	Age 65+	Age 75+
Neoplasms (e.g., cancer)	1,815	1,915	2,342	1,181
Mental Disorders	1,080	1,258	1,370	1,023
Mental Retardation	769	1,090	542	226
Nervous System	1,109	1,370	1,525	1,631
Circulatory system (e.g., heart disease)	1,750	1,779	1,729	2,189
Musculoskeletal disorders (e.g., back pain)	1,033	1,182	1,416	2,321
Respiratory system	1,679	1,719	2,243	3,995
Injuries	1,243	1,235	1,210	1,546
Endocrine system (e.g., diabetes)	2,309	2,908	2,224	1,103
All other	1,683	1,715	1,869	2,594
Overall:				
All spending	1,325	1,451	1,678	2,208
Medicaid and Medicare spending	952	1,001	1,240	1,770

Note: These estimates are unweighted and control for gender, age group, education, race, and whether the individual also receives Medicaid.

*Reason for 1st eligibility of Medicare before age 65.

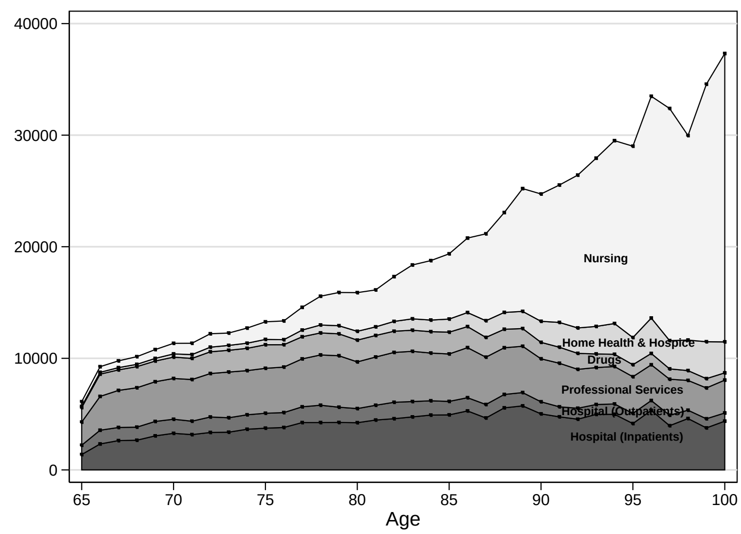
Table 2.12. Mean Monthly Spending of Those with Disabilities, by Age

2.6 Average Medical Spending Over the Life Cycle

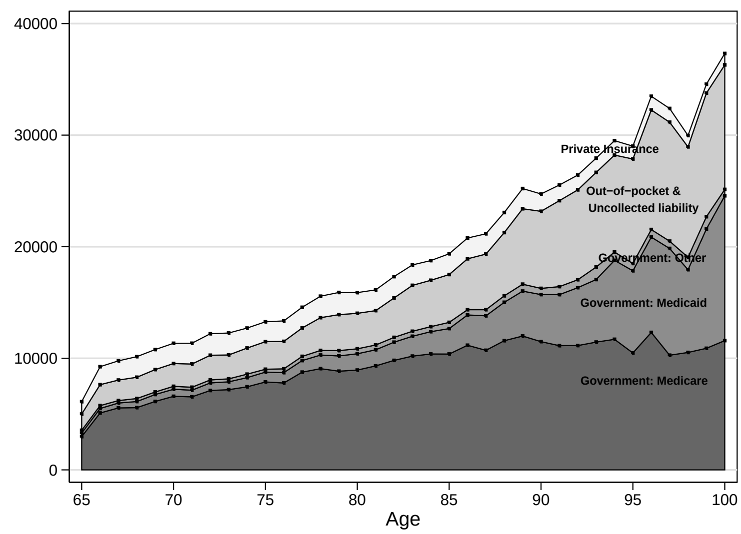
Figure 2.3 shows life-cycle profiles of mean total medical spending. The two graphs in this figure plot spending profiles, first by expenditure type and then by payer.[19] The estimates show that average medical spending exceeds $25,000 per year for those in their 90s. The top panel shows this is almost entirely due to nursing home expenditure. In fact, most other forms of expenditure fall with age after age 90. The bottom panel shows medical spending by payer. Given that nursing home care is mostly paid either out-of-pocket or by Medicaid and that nursing home spending rises quickly with age, it should come as no surprise that most of the increase in spending with age is paid either out-of-pocket or by Medicaid.

An interesting question is to what extent the rise in medical expenses with age is due to the fact that people require more expensive medical services at older ages and to what extent it is due to large medical expenditures right before death. Yang et al. (2003) argue that medical spending in the US increases with age primarily because mortality rates increase with age and

[19] We estimate total medical spending on a full set of age dummies, with age top-coded at 100, without adjusting for cohort effects.



(a) By Expenditure



(b) By Payer Type

Figure 2.3. Average Total Medical Expenditures

end-of-life expenditures are high. Other papers reach similar conclusions using data from different countries. For instance, Zweifel et al. (1999) use Swiss data, Seshamani and Gray (2004) use data from England and Polder, Barendregt and van Oers (2006) use data from the Netherlands. Interestingly, de Meijer et al. (2011) use Dutch data to find that time-to-death predicts long-term care expenditures primarily by capturing the effects of disability. Yang et al. (2003) find that inpatient expenditures incurred near the end of life are higher at younger ages, while long-term care expenditures rise with age. Braun et al. (2017) find that total end-of-life costs rise with age. Scitovsky (1994), Spillman and Lubitz (2000) and Levinsky et al. (2001) have also studied this question.

2.7 Medical Spending Before Death

It is often argued that people in the US spend too much on health care at the end of their lives. A number of studies have shown that end-of-life spending is significant. For example, Hoover et al. (2002) find that 22 per cent of all medical spending in the MCBS is for those in the last 12 months of life.[20] Here we revisit and update their estimates. We estimate medical spending in the calendar year of death and in the two years before death. We also compare medical spending before death with total aggregate medical spending.

Table 2.13 presents key facts on medical spending in the final three years of life, relative to the medical spending of the whole population. The top panel displays aggregate statistics on medical spending and mortality for the US in 2008 that are useful for making these calculations. National statistics for spending come from the aggregate NHEA data. The rightmost column displays corresponding statistics from the MCBS. Data on mortality come from the National Vital Statistics Reports.[21] The top panel of the table shows that the MCBS matches the aggregate spending statistics reasonably well and that it matches mortality statistics very well, giving us additional confidence in the data.

The bottom panel of Table 2.13 displays medical spending in the last years of life. The leftmost column refers to mean spending in the last one, two and three calendar years before death. If an individual dies in March, medical spending in the year of death will refer only to medical spending between January and March. All the data in Table 2.13 are for 2008, so spending in the 'next-to-last' and 'second-to-last' years is by people who go on to die in 2009 and 2010, respectively. Spending in the last calendar year of life averages $43,030, or about six times average spending for the entire population and over twice the average medical spending of the population aged 65 and over. Average medical spending in the previous year is $42,810, again

[20] Other studies include Lubitz and Riley (1993), Scitovsky (1994), Levinsky et al. (2001), Riley and Lubitz (2010) and Marshall, McGarry and Skinner (2011).

[21] Miniño et al., 2011.

Aggregate Medical Spending and Mortality			
	Total Population	*Population Aged 65 and Over*	
	National Stats	National Stats	MCBS
Personal health care expenditure			
Mean spending per person ($)	7,220	19,110	15,570
Aggregate spending ($ billion)	2,190	740	600
Mortality			
Deaths (million)	2.47	1.80	1.71

Medical Spending in Last Years of Life				
		As a Percentage of Aggregate Spending		
		Total Population	Age-65+ Population	
	Mean Spending	(National Stats)	(National Stats)	(MCBS)
Last years of life from data				
Year of death	**43,030**	**4.9%**	**10.5%**	**12.2%**
Hospitals	21,650	2.4%	5.3%	6.1%
Nursing Home Care	9,150	1.0%	2.2%	2.6%
Next to last year	**42,810**	**4.8%**	**10.4%**	**12.2%**
Hospitals	13,790	1.6%	3.4%	3.9%
Nursing Home Care	14,490	1.6%	3.5%	4.1%
Second to last	**32,860**	**3.7%**	**8.0%**	**9.3%**
Hospitals	8,560	1.0%	2.1%	2.4%
Nursing Home Care	12,290	1.4%	3.0%	3.5%
Sum of last 3 years	**118,690**	**13.4%**	**28.9%**	**33.7%**
Hospitals	44,000	5.0%	10.7%	12.5%
Nursing Home Care	35,920	4.0%	8.7%	10.2%
Hoover et al. method				
Final 12 months	**59,100**	**6.7%**	**14.4%**	**16.8%**
Hospitals	26,870	3.0%	6.5%	7.6%
Nursing Home Care	14,990	1.7%	3.6%	4.3%

Note: All data are for 2008, adjusted to 2014 dollars.
Source: Last years of life spending data from MCBS. Aggregate medical spending data from NHEA. Aggregated death data from National Vital Statistics Reports.

Table 2.13. Medical Spending in the Last Years of Life

about six times average medical spending per person, and spending in the second-to-last year is \$32,860. Of the \$43,030 spending in the last year of life, \$21,650 is on hospital care and \$9,150 is on nursing home care.

The right-hand block of the lower panel in Table 2.13 presents medical spending in the last years of life as a percentage of medical spending at all ages and as a percentage of medical spending for the population aged 65 and over. We calculate these percentages by multiplying the mean spending values in this panel by the number of deaths in the top panel and dividing the resulting product by the aggregate spending values reported in the top panel. By way of example, data from the National Vital Statistics Reports indicate that 2.47 million individuals died in 2008, of whom 73 per cent were aged 65 or older. Assuming that medical spending on people who die aged 65 or over is the same as medical spending for those who die younger than 65, we can infer that aggregate medical spending on all those who died in 2008 was \$43,030 $\times$ 2.47 = \$106.3 billion, which constitutes 4.9 per cent of aggregate medical spending.

Medical spending for the 'year of death' mixes together those who died in January (and so had only one month of spending in the 'year of death') and those who died in December (and so had 12 months of spending), along with those dying in other months. To estimate total medical spending in the last 12 months of life, we apply the approach taken in Hoover et al. (2002) and estimate the following regression:

$$E_i=\beta_0+\beta_1\sqrt{m_i}+\beta_2 m_i+\beta_3 m_i^2+\epsilon_i \tag{2.1}$$

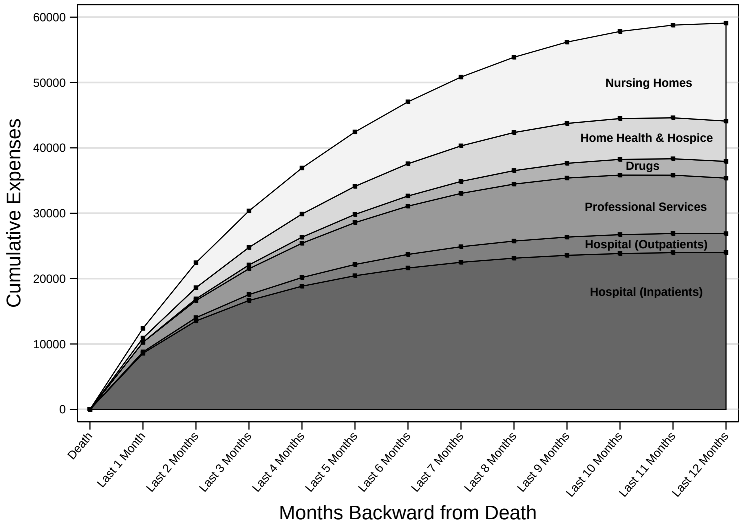
where E_i is total medical spending in the calendar year for individual i and m_i is individual i's exact month of death, where $m_i = 1$ if the month of death is January, $m_i = 2$ if the month of death is February, and so on. The last three rows of Table 2.13 present our results. Using MCBS data from 2008, we find that 16.8 per cent of all medical spending for the population aged 65 and over occurs in the last 12 months of life. Using MCBS data for 1992 to 1996, Hoover et al. (2002) found that 22 per cent of all medical spending for the population aged 65 and over occurs in the last 12 months of life. Our lower estimate appears to be the result of using more recent data. For example, if we use data from just 1996, the estimate becomes 20.9 per cent, much closer to Hoover et al.'s estimate.

Because those aged 65 and over are more likely to die, end-of-life spending is far more important for that age group than for the population as a whole. The population aged 65 and over accounts for only 34 per cent of all medical spending but for 73 per cent of all deaths. The percentage of medical spending at all ages going towards individuals in the last 12 months of life is only 6.7 per cent. Medical spending in the last three years of life represents 13.4 per cent of aggregate medical spending. Thus, while end-of-life spending is high in the US, it hardly explains why total per-capita medical spending is so much higher in the US than in other countries. For example, Polder et al. (2006) find that 10 per cent of all medical expenditures in the Netherlands are made in the last year of life, a higher percentage than (our estimates) for the US.

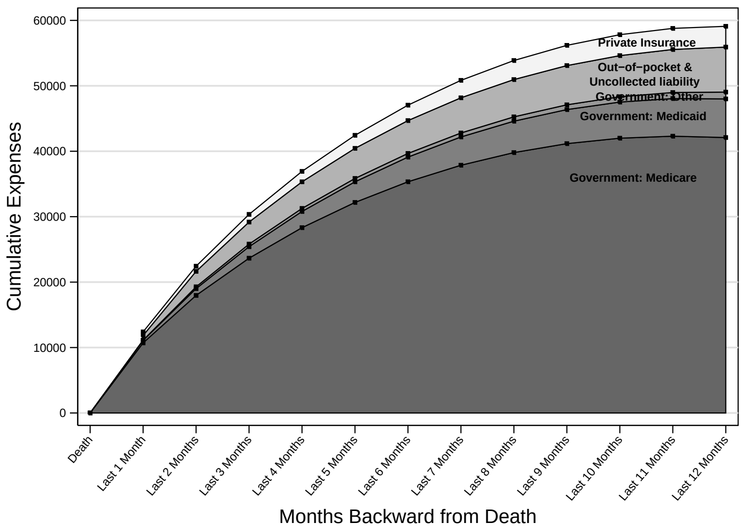
Figure 2.4 shows mean cumulative medical spending over the last 12 months of life as a function of the number of months from death. It decomposes medical spending into spending by payer and expenditure types. Total medical spending in the last month of life averages $12,400, the great majority of which is paid by the government, through Medicare, Medicaid and veterans' programs. Over the final year, total medical spending is $59,100. Of this total, $42,100, or 71 per cent, is covered by Medicare, while $5,900, or 10 per cent, is covered by Medicaid and $1,040 is covered by other government programs. Relative to medical spending for all the elderly (see Table 2.4), the government picks up a larger share of medical spending amongst those near death, most notably through Medicare. Out-of-pocket expenses in the last year of life are $6,500, somewhat lower than found by French et al. (2006) or Marshall, McGarry and Skinner (2011). Uncollected liabilities are $380, while $3,180 is covered by private insurance. The greatest expenditure type is hospital inpatients at $24,000, or 41 per cent, followed by nursing homes at $14,990, or 25 per cent. Expenditures are $8,500 on professional services, $6,170 on home help and hospice, $2,870 on hospital outpatients and $2,560 on drugs.

2.8 Conclusion

We find that medical expenses in the US more than double between ages 70 and 90 and that they are very concentrated: The top 10 per cent of all spenders are responsible for 52 per cent of medical spending in a given year. In addition, those currently experiencing either very low or very high medical expenses are likely to find themselves in the same position in the future. We also find that the poor consume more medical goods and services than the rich and have a much larger share of their expenses covered by the government. Overall, the government covers 67 per cent of the elderly's total medical expenses. Despite this, the expenses that remain after government transfers are even more concentrated among a small group of people. Thus, government health insurance, while potentially very valuable, is far from being complete. Finally, while medical expenses before death can be large, on average they constitute only a small fraction of total spending, both in the aggregate and over the life cycle. Hence, medical expenses before death do not appear to be an important driver of the high and increasing medical spending found in the US.



(a) By Expenditure Type



(b) By Payer Type

Figure 2.4. Spending in the Last 12 Months of Life, by Expenditure and Payer Type

Bibliography

Adler, G. and Phil, M. (1998). Concept and development of the medicare current beneficiary survey. In *Proceedings of the Survey Research Methods Section, American Statistical Association*, volume 153, page 155.

Alemayehu, B. and Warner, K. E. (2004). The lifetime distribution of health care costs. *Health services research*, 39(3):627–642.

Braun, R. A., Kopecky, K. A., and Koreshkova, T. (2017). Old, sick, alone, and poor: A welfare analysis of old-age social insurance programmes. *The Review of Economic Studies*, 84(2):580–612.

Centers for Medicare and Medicaid Services et al. (2015). National health expenditure data. *historical tables available at http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NationalHealthAccountsHistorical.html.*

Daly, M. (2009). Palin stands by "death panel" claim on health bill. *Real Clear Politics, 13 August.*

de Meijer, C., Koopmanschap, M., d'Uva, T. B., and Van Doorslaer, E. (2011). Determinants of long-term care spending: Age, time to death or disability? *Journal of health economics*, 30(2):425–438.

De Nardi, M., French, E., and Jones, J. B. (2010). Why do the elderly save? the role of medical expenses. *Journal of Political Economy*, 118(1):39–75.

De Nardi, M., French, E., and Jones, J. B. (2016). Medicaid insurance in old age. *The American Economic Review*, 106(11):3480–3520.

Emanuel, E. J. and Emanuel, L. L. (1994). The economics of dying–the illusion of cost savings at the end of life. *New England Journal of Medicine*, 330(8):540–544.

Fahle, S., McGarry, K., and Skinner, J. (2016). Out-of-pocket medical expenditures in the united states: Evidence from the health and retirement study. *Fiscal Studies*, 37(3-4):785–819.

Feenberg, D., Skinner, J., et al. (1994). The risk and duration of catastrophic health care expenditures. *The Review of Economics and Statistics*, 76(4):633–47.

French, E., De Nardi, M., Jones, J. B., Baker, O., and Doctor, P. (2006). Right before the end: Asset decumulation at the end of life.

French, E. and Jones, J. B. (2004). On the distribution and dynamics of health care costs. *Journal of Applied Econometrics*, 19(6):705–721.

Gardner, L. and Gilleskie, D. (2006). The effects of state medicaid policies on the dynamic savings patterns of the elderly. Technical report, National Bureau of Economic Research.

Hoover, D. R., Crystal, S., Kumar, R., Sambamoorthi, U., and Cantor, J. C. (2002). Medical expenditures during the last year of life: findings from the 1992–1996 medicare current beneficiary survey. *Health services research*, 37(6):1625–1642.

Kaiser Family Foundation (2013). *Medicaid: A Primer*.

Levinsky, N. G., Yu, W., Ash, A., Moskowitz, M., Gazelle, G., Saynina, O., and Emanuel, E. J. (2001). Influence of age on medicare expenditures and medical care in the last year of life. *Jama*, 286(11):1349–1355.

Lubitz, J., Cai, L., Kramarow, E., and Lentzner, H. (2003). Health, life expectancy, and health care spending among the elderly. *New England Journal of Medicine*, 349(11):1048–1055.

Meara, E., White, C., and Cutler, D. M. (2004). Trends in medical spending by age, 1963–2000. *Health Affairs*, 23(4):176–183.

Pashchenko, S. and Porapakkarm, P. (2016). Medical spending in the us: facts from the medical expenditure panel survey data set. *Fiscal Studies*, 37(3-4):689–716.

Polder, J. J., Barendregt, J. J., and van Oers, H. (2006). Health care costs in the last year of life – the dutch experience. *Social science & medicine*, 63(7):1720–1731.

Scitovsky, A. A. (1984). " the high cost of dying": what do the data show? *The Milbank Memorial Fund Quarterly. Health and Society*, pages 591–608.

Scitovsky, A. A. (1994). " the high cost of dying" revisited. *The Milbank Quarterly*, pages 561–591.

Seshamani, M. and Gray, A. M. (2004). A longitudinal study of the effects of age and time to death on hospital costs. *Journal of health economics*, 23(2):217–235.

Spillman, B. C. and Lubitz, J. (2000). The effect of longevity on spending for acute and long-term care. *New England Journal of Medicine*, 342(19):1409–1415.

Wagstaff, A. and Van Doorslaer, E. (2000). Equity in health care finance and delivery. *Handbook of health economics*, 1:1803–1862.

Yang, Z., Norton, E. C., and Stearns, S. C. (2003). Longevity and health care expenditures: the real reasons older people spend more. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 58(1):S2–S10.

Young, K., Garfield, R., Musumeci, M., Clemans-Cope, L., and Lawton, E. (2012). Medicaid's role for dual eligible beneficiaries. *Kaiser Commission on Medicaid and the Uninsured*, pages 1–19.

Zweifel, P., Felder, S., and Meiers, M. (1999). Ageing of population and health care expenditure: a red herring? *Health economics*, 8(6):485–496.

2A Appendix

2A.1 Supplementary Tables

Table 2A.1 shows mean Medicaid payments conditional on payment percentile for both the MCBS and the MSIS (from Young et al., 2012[22]). Our calculations for this table use the subset of the MCBS that receives both Medicare and Medicaid, the subset most similar to the subset of the MSIS data used by Young et al. The table shows that in both data sets, the least costly 50% of total Medicaid enrolees account for less than 1% of total Medicaid payments, whereas the most costly 5% are responsible for over 40% of the total. But even though the MCBS Medicaid data match the MSIS expenditure shares, they understate the level of spending at all parts of the distribution.

Spending Percentile	% of Medicaid enrollees	% of Medicaid spending (MSIS)	Average Spending per Enrollee (MSIS)	% of Medicaid spending (MCBS)	Average Spending per Enrollee (MCBS)
All	100%	100%	15,880	100%	8,760
95–100%	5%	40.9%	118,490	43.9%	76,880
90–95%	5%	20.4%	59,420	26.8%	46,910
70–90%	20%	32.4%	25,980	26.2%	11,480
50–70%	20%	5.5%	4,370	2.6%	1,140
0–50%	50%	0.9%	280	0.4%	90

Note: 2008 MSIS data, adjusted to 2014 dollars.

Table 2A.1. Medicaid Enrolment and Expenditures by Enrollee Spending Percentile: MSIS versus MCBS

Table 2A.3 presents a different measure of expenditure ratios: construct the ratios for each individual, then average over all individuals. This differs from the measures used in Tables 2.2 and 2.4, where expenditures are averaged across all individuals and then used to calculate ratios. As it turns out, changing the method of calculating ratios has significant effects. For example, the share of aggregate medical expenditures covered by Medicaid is 9.4%, but the average individual Medicaid share is 4.1%. The difference arises because taking the ratio of the means weights more heavily those with high medical spending. Medicaid spending is concentrated amongst a small number of individuals who consume a very large amount of medical resources. Most individuals receive no Medicaid assistance at all. Among expenditure types, nursing home care represents 20.6% of medical spending in the aggregate, versus 5.3% when averaged across individuals. Again, the key difference is the weighting: the small share of people in nursing homes consume a great deal of medical resources, meaning that nursing home expenditures are responsible for a large share of total resources.

[22] Young, K., Garfield, R., Musumeci, M. B., Clemans-Cope, L. and Lawton, E. (2012), 'Medicaid's role for dual eligible beneficiaries', Kaiser Commission on Medicaid and the Uninsured, Issue Brief.

	All	Men	Women
Fraction by Payer			
Out-of-Pocket	28.5%	28.0%	28.9%
Private Insurance	18.2%	19.3%	17.4%
Uncollected liabilities	2.2%	2.3%	2.2%
Government	51.1%	50.4%	51.5%
Medicaid	4.1%	2.6%	5.2%
Medicare	43.5%	42.5%	44.2%
Other government	3.5%	5.3%	2.2%
Fraction by Type of Expenditure			
Nursing Home Care	5.3%	3.6%	6.5%
Hospitals	19.6%	20.8%	18.7%
Inpatients	9.7%	10.5%	9.0%
Outpatients	9.9%	10.3%	9.6%
Professional Services	43.0%	44.3%	42.0%
Drugs	30.2%	29.7%	30.6%
Home Help and Hospice	1.9%	1.5%	2.2%

Note: This table reports expenditure ratios for each individual, averaged over all individuals.

Table 2A.3. Percentage of Total Expenditures, by Payer and Expenditure Type: MCBS data

Type of Spending	A: Spending in Levels			B: Spending in Logs		
		t+1	t+2		t+1	t+2
All	All	0.57	0.40	All	0.61	0.53
	Men	0.49	0.33	Men	0.57	0.50
	Women	0.61	0.45	Women	0.64	0.55
		t+1	t+2		t+1	t+2
All (excl. nursing homes)	All	0.45	0.28	All	0.56	0.48
	Men	0.39	0.25	Men	0.54	0.47
	Women	0.49	0.31	Women	0.57	0.49
		t+1	t+2		t+1	t+2
Hospitals	All	0.27	0.19	All	0.30	0.25
	Men	0.28	0.17	Men	0.29	0.24
	Women	0.25	0.20	Women	0.31	0.25

Table 2A.4. Correlation of Medical Spending in Year t with Spending in Years t+1 and t+2, by Gender

Quintile current year	*Quintile next year*				
	Bottom	Second	Third	Fourth	Top
Bottom	61.3	17.8	8.8	6.6	5.5
Second	23.4	36	19.3	12.1	9.1
Third	9.4	24.5	31.1	21.3	13.6
Fourth	6.3	13.2	25.4	31.7	23.5
Top	4.8	8.4	14	26.6	46.3

Quintile current year	*Quintile two years ahead*				
	Bottom	Second	Third	Fourth	Top
Bottom	57.4	18.4	10.2	7.5	6.6
Second	25.2	31.4	19.4	12.8	11.2
Third	11.7	24.8	27.8	20.5	15.3
Fourth	7.6	14.6	24.3	30.1	23.4
Top	6.3	10.4	15.6	26.1	41.7

Table 2A.5. Transition Matrices for Total Medical Expenditure Excluding Nursing Home Costs

Quintile current year	*Quintile next year*				
	Bottom	Second	Third	Fourth	Top
Bottom	39.5	24.8	13.6	11.3	10.9
Second	27.1	28.9	17.9	13.7	12.4
Third	13.9	20.9	29	20.5	15.7
Fourth	10.9	14.3	23.1	30	21.7
Top	10	12.1	16.1	23.9	37.8

Quintile current year	*Quintile two years ahead*				
	Bottom	Second	Third	Fourth	Top
Bottom	36.4	23.7	14.2	12.8	12.9
Second	28.4	27.3	17	13.8	13.5
Third	15.1	21.8	26.9	20	16.2
Fourth	11	14.8	23.4	28.7	22.1
Top	11.4	13.7	17.3	23	34.6

Table 2A.6. Transition Matrices for Hospital Expenditure

Year	Mean spending ($)	Standard Errors
1996	808	(36.5)
1997	824	(37.5)
1998	872	(40.3)
1999	876	(40.3)
2000	937	(40.3)
2001	1,051	(41.7)
2002	1,123	(44.4)
2003	1,183	(43.6)
2004	1,130	(44.9)
2005	1,121	(49.2)
2006	1,136	(42.7)
2007	1,113	(42.8)
2008	1,202	(45.1)
2009	1,193	(44.6)
2010	1,128	(42.6)

Table 2A.8. Out-of-Pocket Insurance Premia Spending, by Year

	2008 data only Mean spending ($)		1996–2010 data Mean spending ($)	
Year of death	Unweighted	Population weighted	Unweighted	Population weighted
All spending	41,801	43,033	37,889	38,514
Out-of-pocket	5,128	4,439	5,773	5,182

Table 2A.9. Last Year of Life Spending, From Data

Chapter 3

The Effect of Disability Insurance Receipt on Mortality

3.1 Introduction

This paper estimates the effect of Disability Insurance (DI) and Supplemental Security Income (SSI) benefit receipt on mortality, for those persons who would receive benefits if their case is heard by a lenient Administrative Law Judge (ALJ), but not if their case is heard by a stricter judge. We compare mortality rates of individuals who applied for and received disability insurance benefits to the mortality rates of those who applied for benefits but were denied. Those receiving benefits receive large cash transfers, and health insurance from Medicare or Medicaid. However, beneficiaries also face important work disincentives. Each of these factors could affect mortality. The income and health insurance benefits, taken together, likely reduce mortality, but the work disincentive could increase mortality. Identifying the overall mortality effect is difficult, however, because those allowed benefits may be unobservably less healthy than those denied.

Using Social Security administrative data, we exploit the essentially random assignment of DI cases to ALJs. We document large differences in allowance rates across judges, and show that these differences are unrelated to the health or earnings potential of DI applicants. We use judge specific allowance rates to construct an instrumental variable, which we call "judge leniency". We use our judge leniency instrument to predict the allowance of individual cases. We then use predicted allowance to estimate the effect of allowance on mortality. Because we have the population of DI applicants whose case was heard by an ALJ over 1995-2004, we can obtain precise estimates, even for relatively small subgroups of this population. We find heterogeneous effects. For persons aged 55-64 when assigned to an ALJ, DI benefit allowance increases the mortality rate, 10 years after assignment, by a statistically significant 2.81%, relative to a baseline 10-year mortality rate of 22.0%. This increase in mortality is surprising given that benefit allowance provides a cash benefit, which is likely to improve mortality, and health insurance, which may also improve mortality for this population. Balanced against this, allowance creates a large work disincentive. We find evidence suggesting that all three effects are important, with the net effect on mortality varying based on age and prior health.

This point estimate is a weighted average of Marginal Treatment Effects (MTEs) among those impacted by variations in judge leniency. We also estimate the distribution of MTEs within the observed range of judge leniency. Thus we can identify MTEs for healthier individuals who would be denied by almost all judges (apart from the most lenient) as well as less healthy individuals who would be allowed by almost all judges (apart from the strictest). Among the healthier individuals, DI receipt increases mortality. However, among the less healthy individuals, DI receipt tends to reduce estimated mortality, especially for those aged 55-64. Therefore, inframarginal individuals (who would be allowed benefits, whether seen by a strict or lenient judge and are the majority of DI recipients) likely benefit from DI receipt. Thus, our findings are consistent with the view that DI receipt reduces mortality on average.

Our results suggest that measured by impact on mortality, the current DI screening threshold is close to optimal. Making it less strict (and thus increasing the allowance rate) will likely increase mortality for marginal applicants, especially for those assigned to more lenient judges. Conversely, a modest tightening of the allowance rules, which brings them closer to the current decisions of stricter judges, should reduce mortality.

We find heterogeneous effects based on recipients' health conditions, although estimates are not precise. Benefit receipt lowers mortality among those with cancer, which is the highest mortality rate condition, and often requires expensive medical treatment that health insurance could help to fund. Benefit receipt also lowers mortality for recipients with respiratory and nervous system conditions, which are also high mortality conditions. Conversely, for recipients with conditions with lower medical spending, but likely strong labor supply effects, such as musculoskeletal disorders, benefit receipt predicts higher mortality.

We rely on Social Security Administration (SSA) mortality records. We compare these records to mortality records from the National Death Index (considered to be the best available source US mortality rates). SSA mortality records have historically been of suspect quality, but as we show, they have substantially improved in recent years. Mortality rates in the two databases are extremely similar for the age 55+ population: the SSA data appear to understate mortality rates for these persons by less than 1%, with larger understatement for younger age groups. Thus, our results are unlikely to be materially affected by any tendency for SSA being more likely to record deaths for DI recipients than for non-recipients. We assess the robustness of our results to potential under-reporting of mortality among those denied benefits, and find that allowing for potential under-reporting only slightly reduces our estimates of the effect of DI receipt on mortality, especially for persons aged 55-64.

Section 3.2 gives a literature review, section 3.3 describes the DI system, section 3.4 describes our estimation methods, and section 3.5 shows the data and discusses the data quality. In section 3.6 we present our main results, with Marginal Treatment Effect estimates being displayed in section 3.6.4. Section 3.7 discusses some channels by which DI receipt could impact mortality and how the effects vary by health condition. Section 3.8 shows that our estimates are robust to other specifications and methods of handling the data. Section 3.9 concludes.

3.2 Literature Review

Despite the great cost of the Disability Insurance program, relatively little research has been done on how the program affects the health and mortality of the disabled population. We might think that receiving benefits would impact health and mortality, since being allowed benefits impacts the health insurance, income, and employment of those receiving benefits. There is an active literature assessing the separate effects of health insurance, income, and work on health. In this section we review the evidence.

3.2.1 Disability Insurance

To the best of our knowledge, the only other paper to estimate the effect of DI on mortality is Gelber et al. (2017). They estimate the effect of Disability Insurance benefit income on mortality rates. They exploit the kinks in the DI benefit formulas. They measure the effect of benefit generosity on mortality, whereas we measure the effect of receiving benefits versus not receiving them. Receiving benefits not only affects income, but also affects health insurance, and affects labor supply incentives in a different way than receiving a slightly larger or smaller benefit. They find evidence that higher income benefits lead to lower mortality at the lower bend point of the DI benefit formula, but find no robust evidence of an effect at the upper bend point.

3.2.2 Health Insurance

Several important studies, including results from the RAND Health Insurance Experiment (Brook et al., 1983), analyses of Medicare (Finkelstein and McKnight, 2008; Card et al., 2009), and Medicaid (the Oregon Health Insurance Experiment) (Finkelstein et al., 2012) find that for the adult and elderly population, the near-term effect of health insurance on subsequent health outcomes is small. Card et al. (2009) find overall small, statistically insignificant effects of turning 65 and becoming Medicare eligible, but find that access to Medicare does modestly reduce mortality after emergency visits. Some studies do find significant effects of health insurance on mortality. For example, Sommers et al. (2014) finds that after the Massachusetts 2006 health care reform, which attained near-universal insurance coverage in the state, all-cause and health care-amenable mortality decreased when compared with similar counties in other states. Hernandez-Pizarro (2016) estimates the effect of a Spanish system of publicly-allocated long-term care benefits on mortality. She finds that access to greater benefits reduces mortality, particularly for those with only moderate needs. None of these studies focus on the disabled.

To the best of our knowledge, Weathers and Stegman (2012) is the only study that focuses on the value of health insurance for the disabled. They exploit a randomized experiment that reduced the wait time before DI recipients received Medicare benefits from 2 years to 0 years. They find no significant effect of immediate versus delayed receipt of health insurance on mortality; their point estimates imply higher mortality among those who received Medicare immediately.

These studies focus on short run effects, have limited sample sizes, or both. Thus, it is difficult to know if there is no average effect of health insurance on mortality, or if the sample size is too small or the sample period too short to detect an effect. Black et al. (2017) study longer-term effects, but also have a limited sample and use pure observational study methods, rather than a true or natural experiment. Using our data, we can estimate 10 year mortality rates for a large sample.

3.2.3 Income and Employment

Most papers that estimate an effect of income on mortality are estimating the joint effect on mortality of income from employment, and employment itself.

For example, Sullivan and von Wachter (2009) find that job loss significantly increases mortality, potentially reflecting loss of health insurance and loss of income. Several papers using European administrative data, such as Rege et al. (2009) and Eliason and Storrie (2006), find similar results.

Multiple papers have found that reductions in employment lead to poorer health and higher mortality. Fitzpatrick and Moore (2016) document a two percent increase in overall male mortality immediately after age 62, and suggest decreasing labor force participation as the possible key factor. To similar effect, Snyder and Evans (2006) assess the mortality effect of Social Security benefits for members of the "Social Security notch" cohort (those born in the years before 1917), who received benefits at a younger age than those born afterwards. They find that the notch cohort had higher mortality rates and lower employment levels, and conclude that greater work effort has beneficial health impacts, which more than offset any mortality gains from greater Social Security income.

Several recent papers use European retirement reforms to estimate the impact of employment on mortality. While the evidence is mixed, the bulk of the evidence suggests that early retirement increases mortality. For example, Kuhn et al. (2017) find that an early retirement scheme in Austria led to higher mortality among males, with the higher mortality concentrated among heart diseases, diseases related to alcohol consumption, and vehicle accidents. This evidence suggests adverse changes in health behavior as a causal mechanism.

3.3 The Disability Insurance System

Social Security Disability Insurance (SSDI or DI) is one of America's largest social insurance programs. Furthermore, many disabled individuals with low income receive Supplemental Security Income (SSI) benefits. In 2014, 6.4% of people ages 18-64 and 16.3% of those aged 55-64 were receiving either DI or SSI benefits (U.S. Social Security Administration, 2014a).[1] Most DI and SSI beneficiaries also receive health insurance benefits through Medicare (for DI beneficiaries) or Medicaid (for SSI beneficiaries). The combined cost of these programs was $428 billion in 2008 (Livermore et al., 2011), making these programs several times more expensive than unemployment insurance. The costs have risen rapidly, generating many policy proposals to reform the system (Autor and Duggan, 2010; Burkhauser and Daly, 2011; Burkhauser et al., 2014).

3.3.1 Exit Rates from the DI Program

Relatively few people lose disability benefits for reasons other than death.[2] For example, of 7.1 million individuals (DI worker beneficiaries) drawing DI benefits in 2007, 0.5% had benefits terminated because they earned above the Substantial Gainful Activity (SGA) limit for an extended period of time in 2007. Another 0.3% had benefits terminated because they were deemed medically able to work after a continuing disability review, which is a periodic review conducted by SSA of the health of DI beneficiaries (U.S. Social Security Administration, 2007). Thus, the disability allowance decision is high stakes. If the individual is allowed benefits, that individual is typically given disability benefits until normal retirement age (age 65 during the 1990s and now 66), when the person becomes eligible for regular Social Security benefits.

[1] The percentage for persons aged 55-64 is based on authors' calculations using statistics from (U.S. Social Security Administration, 2014a) for the number receiving DI, (U.S. Social Security Administration, 2014b) for the number receiving SSI, and (U.S. Census Bureau, 2015) for population estimates. The total number of people in this age group receiving both DI and SSI is not reported by the SSA. We assume that the percentage of people receiving both is not dependant on age and therefore use the same percentage of 9.6% for those aged 18-64, which is reported in (U.S. Social Security Administration, 2014a).

[2] DI benefits are converted into retiree benefits once the beneficiary turns the normal retirement age. The statistics above are for DI benefits before the conversion to retiree benefits.

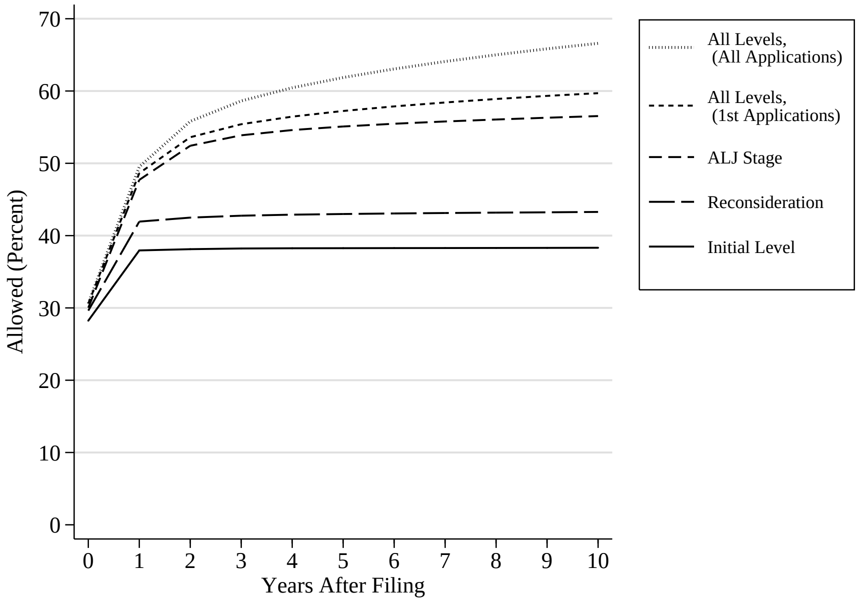


Figure 3.1. Allowance at Different Stages of the Applications and Appeals Process

3.3.2 Determining Eligibility for DI Benefits

An individual is deemed eligible for benefits if they meet certain work requirements and are deemed medically disabled. Although the exact algorithm is complex[3], one of two conditions must be met for the individual to be deemed disabled.

The first is a "listed impairment". Individuals who have one of over 100 specific listed impairments are given immediate benefits. Examples include statutory blindness (i.e., corrected vision of 20/200 or worse in the better eye) and multiple sclerosis.

The second condition is inability to work, either at their past work or other work. Eligibility under this condition turns on a combination of medical impairment and vocational factors such as education, work experience, and age. These cases can be especially difficult to evaluate. Myers (1993), a former Social Security Administration Deputy Commissioner, points out that if a worker "can do only sedentary work, then disability is presumed in the case where the person is aged 55 and older, has less than a high school education, and has worked only in unskilled jobs, but this is not so presumed in the case of a similar young worker. Clearly, borderline cases arise frequently and are difficult to adjudicate in an equitable manner!"

The disability determination is a multi-step process. Figure 3.1 shows the share of applicants who are allowed at different steps during our sample period. After an initial 5-month waiting

[3] See Hu et al. (2001) or Benitez-Silva et al. (1999) for details.

period, DI applicants have their case reviewed by a Disability Determination Service review board. Figure 3.1 shows that 39% of applicants are allowed and 61% are denied at this stage. At this stage the most clear-cut cases are allowed, such as those with a listed impairment. Cases that are harder to judge (such as musculoskeletal problems) are usually denied at this stage. About half of all applicants who are initially denied appeal at the disability determination service reconsideration stage. About 7.5% of those that appealed in 2013 were allowed benefits at this stage (U.S. Social Security Administration, 2014a). Sixty days after the disability determination service decision, a DI appeal can be requested. DI appeals are reviewed by Administrative Law Judges (ALJs) after a delay of about one year.[4] 14% of all initial claims, or 59% of all claims that are appealed, are allowed at the ALJ level.[5] If the case is denied at the ALJ level, the applicant can appeal to the SSA Appeals Council. If the applicant is denied at this level, she can then appeal after 60 days to Federal Court. However, Figure 3.1 shows that appeals at the higher levels are rarely successful: only about 2% of all initial claimants receive benefits at the Appeals Council or Federal Court level. Lastly, denied applicants can re-apply for benefits. The last line on Figure 3.1 includes those who re-apply for benefits. Another 7% of all initial claims are eventually allowed benefits through a re-application. 33% do not get benefits at any stage after 10 years.

Because we identify the causal effect of DI on mortality using variation at the ALJ level, the estimated effect applies only to marginal cases. The least healthy individuals, such as those with listed impairments, will almost always be allowed at the Disability Determination Service stage. The healthiest individuals will almost always be denied, whichever ALJ they see. Thus, our results are not generalizable to all DI applicants. However, the marginal cases are of great policy interest, because these are the individuals most likely to be affected by changes in the leniency of the appeals level of the DI system.

3.3.3 Assignment of DI Cases to Judges

Judicial independence means that judges have a great deal of latitude to determine eligibility (Taylor, 2007). As a result, two different judges can have very different allowance rates even though they see similar applicants.

Administrative Law Judges (ALJs) are assigned to hearing offices, and within hearing office, hear cases on a rotating basis.[6] When a judge finishes a case, that judge received the oldest pending

[4] Judges can make one of three decisions: allowed, denied, or remand. A “remand" is a request for more information from the disability determination service. Our measure of “allowed" is the final determination at the ALJ stage, and thus includes the final decision on remands.

[5] The allowance rate varies by age, and is significantly higher, at 84%, for those age 55-64, who are the principal focus of this study.

[6] Title 5, Part III, Subpart B, Chapter 31, Subchapter I, Section 3105 of the US Code states that “Administrative law judges shall be assigned to cases in rotation so far as practicable" (United States, 2007). The Social Security

case at his or her hearing office. Therefore, for applicants who apply at a given office at a given point in time, the assignment of cases to ALJs is "essentially random" (Social Security Advisory Board, 2006). Judges do not pick the cases they handle. Judges are not assigned cases based on the expertise of the judge. Furthermore, an applicant cannot choose an alternate judge after being assigned a judge.

The initially assigned judge is not necessarily the judge who decides the case. Paletta (2011) documents a judge who took assigned cases from other judges and made decisions on those cases. Thus, the cases were not randomly assigned to the deciding judge.[7] We have information on the assigned judge in addition to the deciding judge. Although the deciding judge is not necessarily randomly assigned, the initially assigned judge is. We use initial assignment to a judge as our source of exogenous variation. The initially assigned judge is the deciding judge in 96% of all cases.

As we confirm below, the assigned judge is for all practical purposes randomly assigned conditional on hearing office and day. However, individuals are not randomly assigned to hearing offices. The zip code in which a person lives determines the hearing office to which they are assigned. Applicant characteristics can vary by location (e.g., black lung disease is more common near mining towns) as well as across time (e.g., the share of DI applicants listing mental illness as the main health problem has risen over time). For this reason we condition on hearing office and day in the estimations below. In doing so, we exploit only within hearing office-day variation in judge level leniency. This variation should be essentially random.

3.4 Estimating Equations

To estimate the effect of DI allowance on mortality, we use a two-step procedure. In the first step we generate an instrumental variable that is a measure of relative judge leniency, within a given hearing office and hearing day. This variable is correlated with the probability of allowance, but is independent of applicant health and other characteristics. In the second step we use

Administration's Hearings, Appeals and Litigation Law Manual (HALLEX) Volume I Chapter 2 Section 1-55 states that "the Hearing Office Chief Administrative Law Judge generally assigns cases to ALJs from the master docket on a rotational basis, with the earliest (i.e., oldest) Request for Hearing receiving priority." (U.S. Social Security Administration, 2009). HALLEX gives 11 exceptions to this rule. For example, the exceptions include "critical cases", such as individuals with terminal conditions and military service personnel, as well as remand cases. These cases are expedited and reviewed by Senior Attorneys. If there is a clear cut decision to be made, then the Senior Attorney will make the decision without a hearing. If the case is not clear cut, then the case is put back in the master docket and is assigned to a judge in rotation. We can identify cases that were decided without a hearing and delete them from our sample. We study the remaining cases where there was a hearing.

[7] Furthermore, an individual can potentially reject the assigned judge. For example, if an individual misses her court case, she may be reassigned to a different judge. Also, some cases in remote areas are held via video conference where the judge and claimant are not in the same room. Claimants can demand that the judge be present at a hearing, and thus the judge must travel to the claimant. Some judges refuse to travel, and thus another judge will be reassigned to the case.

instrumental variables procedures to estimate the effect of DI on mortality, as well as other factors that potentially affect mortality, such as employment, earnings and benefits. We focus principally on applicants age 55-64 at time of application, because SSA death records are more accurate for older applicants, which we explain in more detail in section 3.5, but also present results for younger applicants.

3.4.1 Basic Specification

Our basic estimating approach is a modified instrumental variables regression where in a first stage we estimate

$$A_i = j_i\gamma + X_i\delta_A + e_i. \tag{3.1}$$

where A_i is a 0-1 indicator equal to 1 if individual i is allowed benefits by the ALJ, j_i are judge indicator variables (equal to 1 if judge j heard individual i's case), and X_i are hearing office-day indicators (equal to 1 if individual i's case is assigned on that hearing office-day pair). In some specifications we add further covariates such as gender, age, race, past income, legal representation, application type (SSDI or SSI), education, and main health condition of the individual. For the second stage we adopt the random coefficients model of Bjorklund and Moffitt (1987):

$$y_{i\tau} = A_i\phi_{i\tau} + X_i\delta_{y\tau} + u_{i\tau} \tag{3.2}$$

where $y_{i\tau}$ is mortality (or another outcome variable such as earnings, participation, appeals or allowance), τ years after after assignment to an ALJ. We allow for heterogeneity in the parameter $\phi_{i\tau}$ to capture heterogeneity in the effect of benefit receipt on outcomes, both across individuals and over time. We allow the variables $u_{i\tau}$ and $\phi_{i\tau}$ to be potentially correlated with A_i, and with each other.

We focus on the effect of ALJ allowance *at first hearing* on mortality and other outcomes after 5 years and 10 years. ALJ allowance after a first hearing and eventual allowance can differ because some people denied by an ALJ are allowed upon reapplication or appeal (as shown in Figure 3.1). We use ALJ allowance at first hearing rather than eventual allowance because those who die soon after this hearing cannot reapply or appeal: eventual allowance is thus itself a function of mortality, creating a spurious correlation between eventual allowance and mortality. This problem is circumvented by using ALJ allowance.

3.4.2 Estimating Equations

When estimating equation (3.2) we are confronted with three concerns. First, we wish to allow for heterogeneity in the parameter $\phi_{i\tau}$. Second, we have 1,404 judges in our sample, each of whom is a potential instrument. IV estimators can suffer from small sample bias when both the number of instruments and the number of observations is large (e.g., Hausman et al. (2012)). Third, we have just under 200,000 hearing office-day interactions in the covariate set X_i.

To solve these three concerns, we first construct the judge-specific allowance rate of the judge who heard individual i's case, averaged over all cases other than individual i's case. Formally this is

$$Z_i = \frac{1}{N_j - 1} \sum_{s \in \{J\}, s \neq i} A_s \tag{3.3}$$

where N_j is the number of cases heard by judge j_i over the sample period, and $\{J\}$ is the set of cases heard by judge j_i. This has been used as an instrument by Maestas et al. (2013) Dahl et al. (2014), and Autor et al. (2015), for example. We then de-mean this object by hearing office and day, creating $\widetilde{Z}_i$. In what follows "$\sim$" represents a de-meaned variable (e.g., $\widetilde{Z}_i = Z_i - \overline{Z_i}$ where $\overline{Z_i}$ is the mean value of Z_i on all cases that were assigned on the same day and at the same hearing office as case i).

Thus our instrument compares the fraction of cases allowed by judge j with the corresponding average probability for all other judges in the same office-day.[8] We refer to our instrument as judge leniency. Judge leniency will be positive (negative) to the extent that a judge is more (less) likely to allow than other judges making decisions in that same office-day. Because we remove observation i, estimated judge leniency is independent of e_{it} or $u_{i\tau}$, even in a small sample.

Finally, we estimate the equations

$$\begin{aligned} \widetilde{A}_i &= \lambda \widetilde{Z}_i + \epsilon_i, & (3.4) \\ \widetilde{y}_{i\tau} &= \phi_\tau \widehat{\widetilde{A}}_i + \widetilde{u}_{i\tau} & (3.5) \end{aligned}$$

jointly using two stage least squares. In specifications where we include additional covariates we also demean each covariate by hearing office and day.

Given the above assumptions, the estimated effect can be interpreted as a Local Average Treatment Effect (LATE). The object we identify is not technically a LATE, since a LATE assumes a binary instrument, whereas our instrument is continuous. However, some papers refer to this as a LATE. More precisely, our procedure identifies a weighted average of $\phi_{i\tau}$ for the individuals affected by the instrument (see Heckman et al. (2006) and French and Taber (2011) for more details).

We identify the LATE if three conditions are met. First, if judges are randomly assigned to cases, conditional on date and hearing office, then assignment satisfies the "independence assumption".

[8] Doyle Jr (2007) and French and Song (2014) construct a slightly different judge leniency variable–this alternative approach is described in appendix 3A.3.2. When we replace $\widetilde{Z}_i$ with their instrument we obtain similar results (see Section 3.8).

Second, if judges differ only in leniency and rank applicants the same with respect to relative severity of their disability, then the Imbens and Angrist (1994) “monotonicity assumption” is satisfied. The monotonicity assumption implies that a case allowed by a strict judge will always be allowed by a more lenient one. Third, we assume that the instrument causes variation in allowance rates, sometimes known as the rank or existence condition. Sections 3.6.1 and 3.6.2 provide evidence on the extent to which these assumptions hold.

3.4.3 Marginal Treatment Effects

We are interested both in the LATE – the average effect of allowance for the marginal cases for which we can identify this effect – and also how the treatment effect varies with judge leniency, within the range of leniencies that we observe. Section 3.6.4 presents estimated Marginal Treatment Effects (MTEs), which measure how the mortality response varies with (de-meaned) allowance rates. We use a polynomial estimating equation to estimate the MTE. Heckman et al. (2006) experiment with different approaches to estimating the MTE, such as local polynomial smoothers. They find that the polynomial approach works about as well as other procedures.[9] We estimate the equations

$$\widetilde{A}_i = \sum_{k=1}^{K} \lambda_k (\widetilde{Z}_i)^k + \eta_i, \tag{3.6}$$

$$\widetilde{y}_{i\tau} = \sum_{k=1}^{K} \varphi_{k\tau} \widetilde{(\widehat{\widetilde{A}}_i)^k} + \mu_{i\tau} \tag{3.7}$$

where $\widehat{\widetilde{A}}_i$ in equation 3.7 is the predicted value of $\widetilde{A}_i$ from equation (3.6), and K is the order of the polynomial.

As shown by Heckman et al. (2006) and French and Taber (2011), as well as appendix 3A.3, the estimated MTE(a) is

$$\sum_{k=1}^{K} k\varphi_{k\tau} \widetilde{(\widehat{\widetilde{A}}_i)^{k-1}} = \hat{E}[\phi_{i\tau} | \text{allowed only if } \widehat{\widetilde{A}}_i \geq a, \text{not allowed if } \widehat{\widetilde{A}}_i < a,] \tag{3.8}$$

where a is a particular realization of (de-meaned) judge observed health of an applicant who would be allowed by a fraction a of all judges. Equation (3.8) shows that MTE(a) is the mean value of $\phi_{i\tau}$ for those who would be allowed if their assigned judge allowed slightly higher than a share a of cases, and would be denied if assigned to a judge allowing slightly lower than this share.

[9] Our Monte Carlo simulations suggest there is very little bias when using polynomials. Furthermore, the polynomial procedure is computationally feasible with large numbers of covariates, such as a full set of hearing office-day interactions.

As a increases, we estimate the effect of the instrument for individuals with less severe disability. Appendix section 3A.3.1 provides more details on interpretation and estimation of the MTE.

3.5 Data

Our initial sample is all individuals aged 25-64 who appealed either a DI or SSI initial benefit denial, and were assigned to an ALJ during 1995-2004. Using Social Security Numbers, we match together data from the SSA 831 file, the Office of Hearings and Appeals Case Control System (OHACCS), the Hearing Office Tracking System (HOTS), the Appeals Council Automated Processing System (ACAPS), the Litigation Overview Tracking System (LOTS), the Master Earnings file (MEF), and mortality data from the Numerical Identification file (NUMIDENT). These data are described in greater detail in the appendix. We study mortality outcomes up to 10 years following assignment to a judge. Thus, our mortality data run from 1995 to 2014.

We drop all observations heard by a judge who heard less than 200 cases during the sample period. We also drop cases with missing education information. Table 3A.1 in appendix 3A.1 presents more details on sample selection criteria.

Those who die before their case was heard may possibly be recorded as "not allowed," which could inflate near-term mortality for those denied benefits. To address this problem we drop all cases where the individual died before her case was heard. In addition, to address any mismeasurement in whether a case was heard before death, we also drop 30,807 cases where the individual died in the year of assignment to an ALJ. This selection decision has only a modest effect on our estimates, which is shown in robustness checks in section 3.8. Our full estimation sample has 2,759,907 DI or SSI cases heard by 1,436 judges, with a mean allowance rate at the ALJ stage of 70.8%. Our main estimation subsample of those ages 55-64 includes 610,231 cases, with a mean allowance rate at the ALJ stage of 84.1%. All dollar amounts below are in 2014 dollars, deflated by the CPI.

Cases in our sample were heard on 195,935 hearing office-day pairs. Thus, on an average $2,759,907/195,935 = 14.1$ cases were heard at each hearing office-day pair. Although we have a large number of hearing office-day fixed-effects, consistency in fixed effects estimators depends on the number of observations going to infinity, not the number of observations per fixed effect going to infinity. A non-trivial number of cases were heard when there was only a single judge at the hearing office on that day. These observations do not contribute any identifying variation.

Figure 3.2 plots the distribution of judge specific allowance rates, both unconditional (left panel) and also the judge leniency variable constructed in section 3.4.2 , which is conditional on hearing office-day (right panel). There is less variation in allowance rates after conditioning on hearing

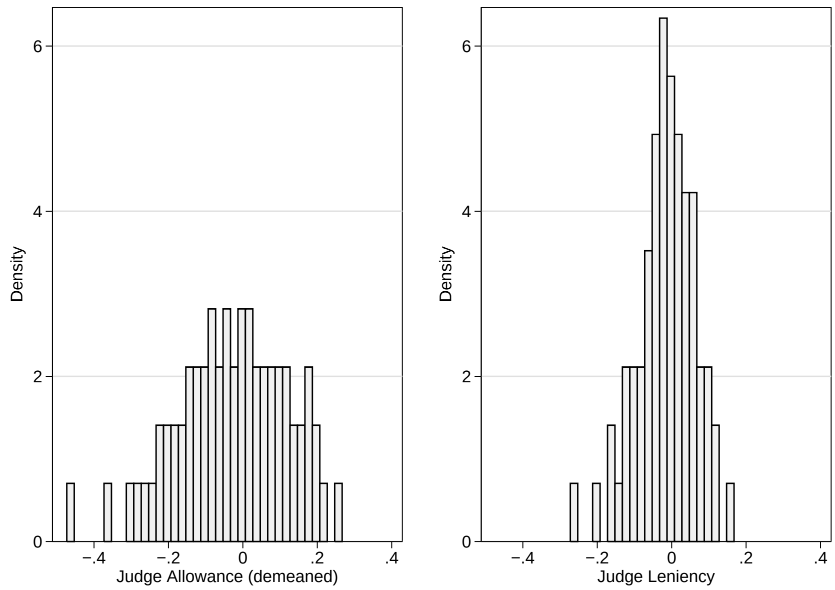


Figure 3.2. Allowance Rate of ALJs, De-meaned (Left Panel), and De-meaned by Hearing Office and Day (Right panel).

office and day; the standard deviation for the unconditional judge allowance rate is 0.149, but the standard deviation of the judge leniency variable is .096 (weighted by the number of cases handled by each judge). This means that being assigned to a judge one standard deviation more lenient than the office-day average increases the probability of allowance by 9.6 percentage points.

3.5.1 SSA Mortality Data

A core data issue for this study is the quality of the SSA mortality data, which comes from SSA's confidential NUMIDENT file. The SSA uses these data to process DI, SSI, and Social Security benefits. These data have been extensively used in previous research, but differ from the data used to construct the official mortality statistics for the US. The SSA obtains death records from various sources, including states, family members, funeral directors, post offices, financial institutions, and other federal agencies. It has a financial incentive to record deaths, especially if it was paying benefits to that individual.

One concern is that because the SSA has a greater financial incentive to record deaths of beneficiaries than non-beneficiaries, it will do a better job of capturing deaths of those who are allowed benefits than those who are denied benefits. Furthermore, it is easier for the SSA to measure deaths of beneficiaries, because if it sends payments to a deceased beneficiary, institutions such as banks (if benefit payments are electronically deposited) or post-offices (if benefit payments are sent by mail) will be more likely to report the death to the SSA (see GAO (2013) for information on how death information is collected). Any undercounting of deaths of those denied benefits will bias down their estimated mortality and could make it appear as if receiving benefits causes higher mortality. As we show below, however, SSA undercounting of deaths, which was formerly a major concern, is no longer an important issue, and should have at most a small effect on our estimates.

Older studies using older versions of the SSA data have shown that the SSA mortality data understates the National Death Index (NDI) data (which are considered the "gold standard" of US mortality data).[10] For example, Hill and Rosenwaike (2002) show that, in the years 1995-1997, the SSA capture approximately 80% of all deaths in the 55-64 year old population, and 95% of all deaths of those 65 and older. However, in separate work (Black et al., 2016), summarized below, we show that the SSA data have greatly improved in recent years, including retroactive updating for prior years.

We estimate the ratio of deaths in the SSA data to NDI deaths over 1995-2014, by age group. We construct these statistics to be as comparable as possible to the SSA data. Thus, we adjust the NDI data to include deaths of people in US territories and exclude foreign residents in the US, because the SSA data includes deaths of US nationals living abroad, whereas the NDI data does not. See Black et al. (2016) for details.

The top panel of Table 3.1 shows our estimates for the years 1995-2014, which is our sample period. Estimates broken down by each year can be found in appendix Table 3A.2. The estimates show that the SSA data have improved considerably relative to the estimates shown in Hill and Rosenwaike (2002). The SSA data capture 98% of all deaths over the 1995-2014 period, and is very close to complete for those over age 55. Indeed, in recent years, SSA data capture somewhat more deaths than the NDI for persons over 65. However, the ratio of SSA/NDI deaths is lower for those under 55. Thus, we focus our principal analyses on applicants age 55-64, where the quality of the SSA mortality data is excellent.

We provide further results on the quality of the SSA mortality data in appendix section 3A.2.2.

3.5.2 Correction for Underreporting in the SSA Mortality Data

While any underreporting of mortality for those denied benefits should be small, nonetheless, to account for possible underreporting, we calculate a correction factor, p, which is the probability that a denied individual's death is observed. We assume that SSA captures all deaths of allowed

[10] The NDI is maintained by National Center for Health Statistics, and made available to researchers through the National Association for Public Health Statistics and Information Systems (NAPHSIS). These data are used to construct the Vital Statistics data for the US.

	All (20+)	20-44	45-54	55-64	65+
Estimated Ratio of Deaths in SSA to NDI data					
1995-1999	0.970	0.944	0.965	0.969	0.973
2000-2004	0.981	0.945	0.975	0.989	0.983
2005-2009	0.991	0.947	0.976	0.993	0.996
2010-2014	0.995	0.957	0.976	0.990	1.000
Average	0.948	0.948	0.973	0.985	0.985
Estimated Ratio of Non-Beneficiary Deaths that are Reported (p)					
1995-1999	--	0.929	0.948	0.955	--
2000-2004	--	0.919	0.962	0.984	--
2005-2009	--	0.918	0.961	0.989	--
2010-2014	--	0.928	0.957	0.983	--
Average	--	0.923	0.957	0.978	--

Notes: Estimated ratio of deaths in the SSA Numident data to adjusted National Death Index deaths over 1995-2014, by age group. Total (20+) column excludes children (age 0-19). The estimated ratio is calculated as D_{kt}/O_{kt} where D_{kt} represents the number of deaths reported in the SSA data for age group k occurring in year t and O_{kt} represents the official number of deaths of U.S. residents reported in the NDI for age group k during year t. Estimated ratio of non-beneficiary deaths that are reported (p) is calculated as in equation (3.9).

Table 3.1. Estimated Percentage of U.S. Deaths Included in the SSA Death Data and Underreporting Correction, by Age Group

individuals but misses a fraction $(1-p)$ of non-beneficiaries' deaths – thus assuming that all of the SSA undercount comes from non-beneficiaries. This is a worst case bound – there are other reasons why SSA may count fewer deaths than NDI – but it gives a sense of how important underreporting among those denied could be for our results.[11] To see why this is likely a worst case bound, consider the fact that if both those allowed and those denied had the same underreporting probabilities then the bias would only come from usual attenuation bias. In appendix 3A.3.3 we show that p can be calculated as:

$$p = \frac{\#\text{of deaths in the SSA data} - \#\text{of deaths of beneficiaries in SSA data}}{\#\text{of deaths in the NDI data} - \#\text{of deaths of beneficiaries in SSA data}} \tag{3.9}$$

We calculate the average of p for each individual in our sample, using their age and year of application, over the sample period in which we observe them which we define as $\bar{p_i}$.[12] This

[11] Although we made several adjustments to the data to make SSA mortality records comparable to the NDI, we cannot fully match the two. For example, illegal immigrants who lack an Social Security number should be captured in the NDI statistics if they die in the US. But SSA records deaths only for persons with Social Security numbers. Thus, the difference between NDI recorded deaths and SSA recorded deaths, likely overstates the number of missing deaths in the SSA data.

[12] In practice we calculate p for each year for the following age groups: 25-44, 45-54 and 55-64. Using these values, we then calculate the two values $\bar{p}_5$ and $\bar{p}_{10}$ for each age and year of application combination using the

approach allows us to reflect in our estimate of p the higher quality of the mortality data at older ages (when most deaths occur) and in more recent years.

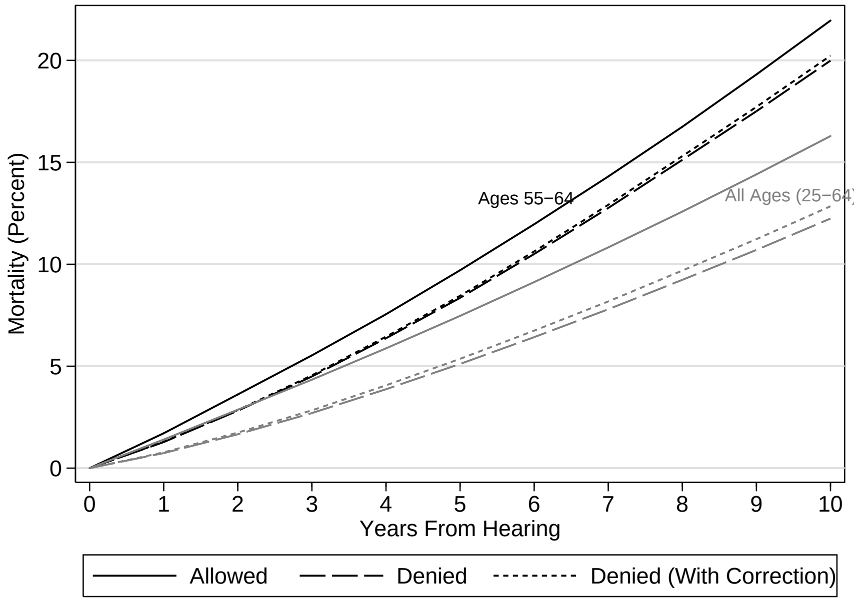
In appendix 3A.3.3 we show how we use $\bar{p}_i$ to calculate a lower bound for the effect of receiving benefits on mortality by multiplying the observed mortality rate for persons denied benefits by $\frac{1}{\bar{p}_i}$ and using the estimation procedures shown in section 3.4.

3.5.3 Mortality Rates of Those Denied and Allowed

In this section we document some basic facts about mortality rates of those allowed versus denied. Figure 3.3 shows cumulative mortality rates conditional on assignment to an ALJ. For those aged 55-64 at time of application, the cumulative mortality rates in the year after assignment to an ALJ are 1.3% for those denied, versus 1.7% for those allowed, respectively. In the subsequent year the rates are 2.8% for those denied and 3.6% for those allowed. Over time, the mortality of those allowed rises faster than those denied, with a 10-year cumulative mortality rate of 22.0% for those allowed and 20.0% for those denied, a difference of 2.0%. For the full sample (aged 25-64), the 10-year cumulative mortality rate is 16.3% for those allowed and 12.2% for those denied, a difference of 4.1%. These differences should not be taken as causal, since those allowed may be less healthy. Our IV strategy seeks to address this issue. The mortality rates for those denied, with and without the correction for underreporting described in the previous section, can been seen in Figure 3.3. The underreporting correction has only a modest effect on our estimates: the estimated difference in 10 year cumulative mortality rates for applicants aged 55-64 between those allowed and denied falls from 2.0% to 1.8%.

Our estimated mortality rates are lower than Parsons (1991). He reports a six year mortality rate for all applicants of 12.9% for those denied and 17.5% for those allowed at ALJ stage. Our estimated six year mortality rates for all ages are 5.8% for those denied versus 8.6% for those allowed, and for the aged 55-64 are at 10.0% for those denied versus 11.1% for those allowed. Our estimates are likely lower because Parsons' cohort is from 1970 whereas ours is from 1995-2005. We also find a much smaller gap than Parsons between mortality for those allowed and denied. This could reflect more complete SSA capture of deaths of those denied benefits. More recent DI beneficiaries tend to be healthier than older ones and have primary diagnoses less related to mortality, as shown in Autor and Duggan (2006). Note that Parsons (1991) shows that mortality rates of those allowed at the initial stage is much higher than mortality rates of those who are allowed at subsequent stages of the adjudication process. Our sample is limited to those who are initially denied and thus get to an ALJ hearing, and our IV estimates of the effect of benefit

mean values of the observations for the 5 or 10 years periods from the application year. We assume p is equal to 1 for those ages 65+, and therefore $\bar{p}_x$ is calculated as: $\bar{p}_{x,age,birthyear} = (\sum_{age=a}^{a+x} p_{g(a),a+birthyear})/x$ where $g(a)$ is the age group 25-44, 45-54, 55-64 or 65+, and $x \in \{5, 10\}$.



Notes: Cumulative mortality rates for applicants aged 55-64, and aged 25-64, at time of hearing with separate mortality rate curves for those allowed benefits, those denied benefits, and those denied benefits with correction for underreporting of mortality.

Figure 3.3. Cumulative Mortality Rates, Allowed versus Denied

receipt on mortality apply to those on the margin for being allowed at the ALJ stage. These estimates should not be extrapolated to applicants who receive benefits at the initial stage. Those who apply at the ALJ stage are healthier than those allowed at the initial stage (but should be less healthy than all persons denied at the initial stage, some of whom do not appeal to an ALJ). Nevertheless, we think that our sample is particularly interesting from a policy perspective, since these are the individuals whose allowance rates are likely to be affected by policy reforms that affect which persons receive benefits.

3.6 Results

3.6.1 Establishing the validity of the Randomization

In previous sections we claimed that the assignment of cases to judges is random, conditional on hearing office and day. Random assignment implies that we should not be able to predict judge leniency using observable characteristics of the applicants who appear before that judge. Table 3.2 presents tests of this hypothesis for persons aged 55-64 when they apply. For similar tests on the full sample see Table 3A.3 in the appendix.

First, we consider which variables predict allowance. Column 1 of Table 3.2 presents estimates from regressing an allowance indicator (de-meaned by hearing office and day) on the gender, age, race, labor force and earnings histories, legal representation, application type, education and health conditions of individuals in our estimation sample. Women, older individuals, whites, those with strong attachment to the labor market, high earners, those represented by a lawyer, and those who did not complete high school are more likely to be allowed benefits. Column 2 presents $t - statistics$ (all standard errors throughout are clustered by judge). Almost all of the covariates are highly statistically significant, due to the large sample size. The R^2 shows that the covariates explain 1.3% of the variation in allowance rates.

Our instrumental variable is judge leniency, $\tilde{Z}_i$. Column 3 presents estimates from a regression of judge leniency on the same covariates. Column 4 provides $t - statistics$.

Of the 20 covariates, only one has a coefficient that is statistically different than 0 at the 5% level, and not strongly so. For the full sample of those aged 25-64 we again only find one covariate that has a coefficient that is statistically different than 0 at the 5% level (see Table 3A.3). All the estimated coefficients are small in comparison to the coefficients on the same variables in the allowance equation. The R^2 shows that the covariates explain 0.22% of the variation in judge specific allowance rates. These results could easily arise by chance, and are consistent with random assignment, which satisfies the independence assumption described in section 3.4. The next section provides some evidence on whether the rank and monotonicity conditions hold.

3.6.2 First Stage Estimates: The Effect of Judge Leniency on Allowance

Table 3.3 in the text and Table 3A.4 in the appendix present estimates of the effect of judge leniency on allowance rates for the main estimation sample and the full sample, respectively. Column 1 shows the number of observations for different subsamples. Column 2 shows the allowance rate at the ALJ stage for that group. It shows, for example, that older individuals, high earners, and those represented by lawyers have relatively high allowance rates.[13] For health conditions, those with neoplasms (e.g., cancer), circulatory problems (e.g., heart disease), and musculoskeletal disorders (e.g., back pain) have high allowance rates, whereas those with mental disorders or retardation have lower allowance rates. Nevertheless, differences in allowance rates across subgroups are small.

Column 3 shows the estimated first stage regression coefficient $\hat{\lambda}$ from a regression of allowance on judge leniency using equation (3.4). The estimated value of $\hat{\lambda}$ for the main estimation sample is .68, meaning that the probability that case i is allowed at assignment rises .68 percentage

[13]The high allowance rate of cases represented by lawyers could be the result of lawyers representing only the most disabled claimants or lawyers causing the allowance probability to rise. We cannot distinguish between these two hypotheses.

	Dependent Variable: Allowed		Dependent Variable: Judge Leniency	
	Coefficient	t-stat	Coefficient	t-stat
Covariate	(1)	(2)	(3)	(4)
Sex				
Female	0.0074	7.3	0.0007	1.9
Age				
55 to 59	-0.0089	-9.5	-0.0019	-2.2
Race				
Black	-0.0170	-10.2	-0.0016	-1.0
Other (non-black, non-white) or unknown	-0.0079	-4.2	-0.0013	-0.9
Labor force participation and income				
Average participation rate, years -11 to -2	0.0068	7.8	0.0006	1.0
Average earnings/billion, years -11 to -2 ($2006)	0.0004	8.9	0.0000	1.1
Represented by lawyer				
Represented by lawyer	0.0185	3.1	-0.0075	-1.8
Application type				
SSDI	-0.0134	-5.3	0.0010	0.5
Education				
High school graduate, no college	-0.0109	-10.8	-0.0012	-1.0
Some college	-0.0234	-14.9	-0.0019	-0.8
College graduate	-0.0269	-12.7	-0.0029	-1.2
Health conditions (by diagnosis group)				
Neoplasms (e.g., cancer)	0.0347	12.2	0.0031	1.2
Mental disorders	0.0019	0.9	0.0003	0.3
Mental retardation	0.0186	3.3	0.0001	0.1
Nervous system	0.0155	7.1	0.0011	1.0
Circulatory system (e.g., heart disease)	0.0325	17.5	0.0031	1.3
Musculoskeletal disorders (e.g., back pain)	0.0281	16.4	0.0031	1.6
Respiratory system	0.0194	8.8	0.0009	0.6
Injuries	0.0218	9.5	0.0016	0.9
Endocrine system (e.g., diabetes)	0.0281	12.8	0.0017	1.0
Standard deviation of dependent variable	0.2887		0.0955	
R^2	0.0127		0.0022	
Number of Applicants = 610,231	Number of Judges =1,436			

Notes: Column (1) is from a regression of de-meaned allowance on all the covariates listed. Column (3) is from a regression of judge leniency on all the covariates listed. Omitted category is male, 60-64s, white, not represented by a lawyer, applying for SSI or SSI and DI concurrently, not a high school graduate, with a health condition other than those listed above. The sample includes applicants aged 55 to 64, and we exclude applicants who died the year of application. Standard errors clustered by judge.

Table 3.2. Predictors of Allowance and Judge Leniency, Aged 55-64

	Obs.	Allowance Rate at ALJ Stage	Coefficient on Judge Leniency	Std. Error	T-Ratio	Relative Likelihood*
	(1)	(2)	(3)	(4)	(5)	(6)
All groups						
All groups	610,231	0.841	0.676	(0.008)	81	1.000
Sex						
Male	291,994	0.839	0.670	(0.010)	64	0.991
Female	318,237	0.843	0.682	(0.010)	71	1.009
Age						
55 to 59	390,600	0.836	0.686	(0.009)	77	1.015
60 to 64	219,631	0.850	0.657	(0.011)	60	0.972
Race						
White	415,125	0.853	0.653	(0.009)	72	0.966
Black	98,698	0.823	0.695	(0.016)	44	1.028
Other or unknown	96,408	0.806	0.747	(0.014)	55	1.104
Income						
Average earnings $<$ $10000	283,146	0.785	0.765	(0.012)	62	1.131
Average earnings $\geq$ $10000	327,085	0.889	0.578	(0.010)	56	0.855
Represented by lawyer						
Represented by lawyer	385,118	0.854	0.652	(0.011)	59	0.964
Not represented by lawyer	225,113	0.820	0.727	(0.017)	44	1.076
Application type						
SSDI	352,991	0.856	0.647	(0.010)	66	0.956
SSI or Concurrent (both SSDI and SSI)	257,240	0.821	0.713	(0.011)	68	1.054
Education						
Less than high school	218,871	0.841	0.664	(0.011)	62	0.982
High school graduate, no college	267,634	0.847	0.668	(0.010)	69	0.988
Some college	77,685	0.830	0.706	(0.015)	46	1.044
College graduate	46,041	0.823	0.740	(0.018)	41	1.094
Health conditions (by diagnosis group)						
Neoplasms (e.g., cancer)	20,000	0.871	0.609	(0.025)	24	0.901
Mental disorders	61,508	0.795	0.817	(0.017)	47	1.209
Mental retardation	3,193	0.812	0.693	(0.056)	12	1.024
Nervous system	34,444	0.828	0.671	(0.022)	30	0.993
Circulatory system (e.g., heart disease)	103,725	0.861	0.637	(0.013)	50	0.942
Musculoskeletal disorders	231,391	0.856	0.648	(0.011)	62	0.959
Respiratory system	30,066	0.845	0.656	(0.020)	32	0.971
Injuries	27,091	0.840	0.689	(0.029)	24	1.019
Endocrine system (e.g., diabetes)	39,331	0.841	0.674	(0.018)	38	0.997
All other	59,482	0.793	0.719	(0.020)	35	1.063

Notes: Column (3) displays the first stage estimate of the coefficient λ from the regression of de-meaned allowance rates on judge leniency for those aged 55-64. Average earnings is calculated on income between 11 and 2 years before application. Standard errors clustered by judge.
*Relative likelihood is the ratio of the group specific coefficient on judge leniency (presented in column 3) to the full sample coefficient.

Table 3.3. First Stage Estimates: Regression of Allowance Rates on Judge Leniency Variable, by Demographics, Aged 55-64

points for every 1 percentage point increase in judge leniency (the de-meaned allowance rate for all other cases heard by case i's judge). Column 4 shows the standard error and column 5 the t-statistic: the estimate of $\hat{\lambda}$ is highly statistically significant for all subgroups. For the full sample in appendix Table 3A.4 the estimate of $\hat{\lambda}$ is .97. The difference in the two estimates arises because we measure judge leniency using the full sample. There is more dispersion in allowance rates in the full sample than for the 55-64 sample, so a judge who is 1 percentage point more lenient on the full sample is only .68 percentage points more lenient for the 55-64 sample, who already have high allowance rates.

Column 3 shows that the estimated coefficient $\hat{\lambda}$ is larger for younger individuals, those with lower labor force participation and earnings prior to appealing, those not represented by a lawyer, and those whose primary health problem is a mental disorder. Abadie (2003) shows that the ratio of the group specific estimate of $\hat{\lambda}$ to the full sample estimate of $\hat{\lambda}$ is informative for understanding the characteristics of those allowed due to a small increase in the ALJ allowance rate. This ratio, shown in column 6, provides the relative likelihood that someone with a given characteristic is allowed given a small increase in judge leniency. Thus, an increase in the allowance threshold of all judges would increase the allowance rate of those with low participation and earnings, those not represented by a lawyer, and those with mental disorders more than for other groups, holding the applicant pool and the rest of the re-applications and appeals process constant. However, all relative likelihoods are close to 1, implying that more lenient judges are lenient across all applicants, to a similar extent.

The monotonicity assumption described in section 3.4 implies that the probability of allowance is non-decreasing in judge leniency for all subgroups of the population. Column 6 provides evidence supporting the monotonicity assumption. Furthermore, all estimates are highly significant, so the rank condition holds.

3.6.3 Second Stage: The Effect of Disability Recipiency on Mortality

Panel (a) of Table 3.4 presents estimates of the effect of disability recipiency on mortality 5 and 10 years after assignment to an ALJ for our main estimation sample. For example, the first two rows show that 21.95% of those allowed benefits in our sample die within 10 years, whereas 19.99% of those denied benefits die within 10 years. This difference of 1.97% is shown in the third row. These estimates suggest that those allowed benefits are more likely to die. An equivalent, way of obtaining this difference is to take the coefficient on allowance from a regression of mortality on allowance. This approach produces a standard error, reported in the fourth row. The difference in mortality between those allowed and denied is statistically significant. However, these are simple OLS estimates, without covariates, which do not address selection effects and do not provide causal estimates.

	Panel (a): Aged 55-64 Mortality (Percent)				Panel (b): Aged 45-54 Mortality (Percent)			
	5 years		10 years		5 years		10 years	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Without Covariates:								
Allowed	9.71		21.95		8.02		17.41	
Denied	8.35		19.99		6.14		14.78	
Coef on allowance	1.35		1.97		1.88		2.64	
(Std. Error)	(0.11)		(0.19)		(0.08)		(0.13)	
Coef on demeaned allowance*	1.35	1.81	1.87	1.93	1.87	1.47	2.71	2.59
(Std. Error)	(0.12)	(0.44)	(0.19)	(0.76)	(0.08)	(0.63)	(0.13)	(0.99)
With Covariates:								
Coef on demeaned allowance*	1.94	2.30	2.77	2.81	2.29	1.49	3.45	2.60
(Std. Error)	(0.12)	(0.50)	(0.18)	(0.91)	(0.08)	(0.62)	(0.12)	(0.94)
With Covariates and Underreporting Correction:								
Coef on demeaned allowance*	1.76	2.12	2.51	2.54	2.05	1.26	3.02	2.17
(Std. Error)	(0.12)	(0.50)	(0.18)	(0.90)	(0.08)	(0.62)	(0.12)	(0.94)

	Panel (c): Aged 25-44 Mortality (Percent)				Panel (d): All Ages (25-64) Mortality (Percent)			
	5 years		10 years		5 years		10 years	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Without Covariates:								
Allowed	5.23		10.91		7.47		16.29	
Denied	3.56		8.47		5.11		12.24	
Coef on allowance	1.67		2.44		2.36		4.05	
(Std. Error)	(0.06)		(0.11)		(0.10)		(0.20)	
Coef on demeaned allowance*	1.63	1.18	2.46	2.26	2.27	2.17	3.93	4.30
(Std. Error)	(0.06)	(0.42)	(0.09)	(0.70)	(0.07)	(0.80)	(0.14)	(1.64)
With Covariates:								
Coef on demeaned allowance*	1.78	1.10	2.76	2.16	2.21	1.64	3.55	2.96
(Std. Error)	(0.05)	(0.38)	(0.09)	(0.64)	(0.05)	(0.58)	(0.10)	(1.05)
With Covariates and Underreporting Correction:								
Coef on demeaned allowance*	1.80	1.02	2.70	1.94	1.97	1.39	3.09	2.49
(Std. Error)	(0.05)	(0.51)	(0.09)	(0.81)	(0.06)	(0.58)	(0.10)	(1.05)

Notes: N= 610,231 in Panel (a), N= 1,048,344 in Panel (b), N=1,101,332 in Panel (c), and N= 2,759,907 in Panel (d). Instrument is judge leniency. Covariates are those in Table 3.2; they include race, sex, age and education groups, health (disability category), average earnings and participation prior to disability, representation by an attorney, and an indicator of concurrent SSDI application. Standard errors clustered by judge. *For de-meaned allowance, all variables are de-meaned from the hearing office-day average.

Table 3.4. Estimated Effect of DI Recipiency on Mortality

Panel (d) of Table 3.4 presents the same estimates as Panel (a), but for the full sample (ages 25-64). Panels (b) and (c) display the estimates for the populations aged 45-54 and 25-44, respectively. Perhaps surprisingly, the full sample (Panel (d)) coefficient on allowance is larger than the coefficients on allowance for each of the subsamples (Panels (a)-(c)). The reason for this is that the coefficient on allowance is the raw difference in mortality between those allowed and denied. Older individuals have higher mortality and are more likely to be allowed. Thus allowed individuals are older, higher mortality individuals, and denied individuals are younger, lower mortality individuals. By separating the full sample into age subgroups, we condition away some of the age-related differences in mortality rates between those allowed and those denied.

The next rows show OLS and IV estimates of de-meaned (by hearing office and day) mortality on similarly de-meaned allowance and the associated standard error. De-meaning the data has very little effect on the OLS estimates. In Panel (b), the IV estimates show that being allowed benefits increases the 5 year and 10 year mortality rate by 1.81 and 1.93 percentage points, respectively. Surprisingly, the IV estimates are close to the OLS estimates.

What can we learn from the similarity of the OLS and IV estimates? Less than one might think. The average allowed applicant is likely in worse health than the average denied applicant; thus, higher overall mortality for those allowed benefits in OLS, without covariates, is expected. The OLS estimate also assumes homogeneous treatment effects across all applicants, regardless of health. This seems unlikely. IV, in contrast, estimates the average effect of allowance for the subsample of applicants who are on the margin for being allowed or denied, and hence affected by the judge leniency instrument. Given the 84% average allowance rate, this subsample is likely healthier than the average for all applicants. The IV estimate is based on random assignment, so the marginal allowed and denied applicants should be in similar health. IV provides a credible estimate of the effect of allowance on mortality, but only for those on the margin to be allowed or denied.

The next rows provide OLS and IV estimates which include the covariates listed in Table 3.2. Adding covariates to this specification has only a small effect on the IV estimates. Recall that our IV estimation procedure should deliver consistent estimates, with or without covariates. Thus, it is reassuring to see that adding covariates has only a small effect on the IV estimates. The IV estimates are strongly statistically significant at both 5 and 10 years.

More surprisingly, adding covariates increases the estimated OLS effect of benefit receipt on mortality. On closer look, adding some covariates increases the estimated effect of benefit receipt on mortality, whereas adding others decreases this estimate. Some groups with higher mortality rates (shown in Table 3.5) also have high allowance rates (shown in Table 3.3). For example, those with cancer, and older (age 60+) individuals have both higher mortality rates and higher allowance rates. Conditioning on these variables moves the OLS estimates closer to 0. However, other groups with higher mortality, such as blacks and those with low prior earnings, have lower allowance rates. Conditioning on these variables produces larger OLS estimates.

The OLS estimates with covariates would have a causal interpretation only under two strong assumptions: that unobservables do not predict both allowance and mortality (no omitted variable bias); and treatment effects are homogeneous. Since accounting for selection on observables somewhat increases the estimated mortality effect, it is plausible that inability to account for unobservables does not necessarily lead to upward biased estimates. However, below, we find evidence of heterogeneous treatment effects.

The final rows in each panel of Table 3.4 display the estimates with covariates, after including the underreporting correction described in section 3.5.3. As expected, the estimates fall slightly but remain broadly the same.

For the full sample, in Panel D, the IV estimates with covariates show that being allowed benefits increases the 5 year and 10 year mortality rate by 1.64 and 2.96 percentage points, respectively. For the full sample, the IV estimates with covariates are somewhat smaller than the OLS estimates, but this comparison should be made cautiously, because these estimates apply to different populations.

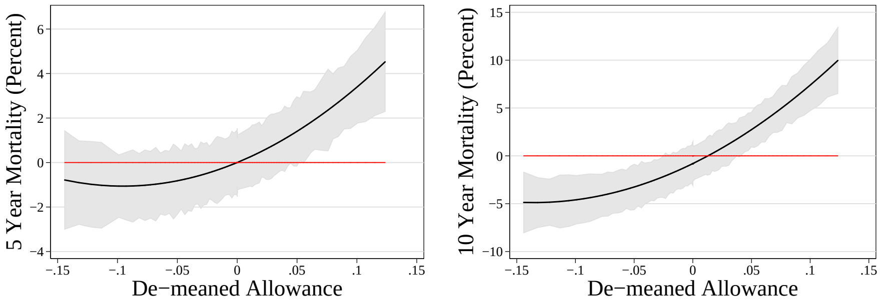
3.6.4 Heterogeneity in the Mortality Effect Based on Judge Leniency: Marginal Treatment Effects

Using the Marginal Treatment Effects approach described in section 3.4.3 and the appendix section 3A.3.1, this section shows how the predicted effect of DI benefit allowance varies with predicted de-meaned allowance.

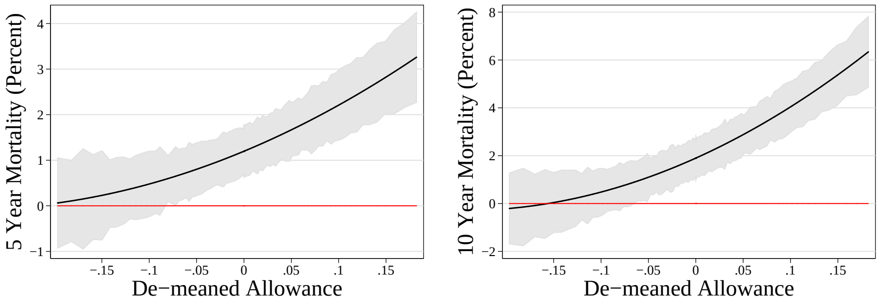
Figure 3.4 presents four panels, all showing how the MTE (i.e., the mortality response for the marginal case allowed) varies with predicted de-meaned allowance. The left panels show 5 year mortality responses. The right panels show 10 year mortality responses. The top panels show estimates (without covariates) for our main estimation sample, ages 55-64, whereas the bottom panels show estimated mortality responses for full sample, ages 25-64.

Using the Marginal Treatment Effects approach described in section 3.4.3 and the appendix 3A.3.1, this section shows how the predicted effect of DI benefit allowance varies with predicted de-meaned allowance.

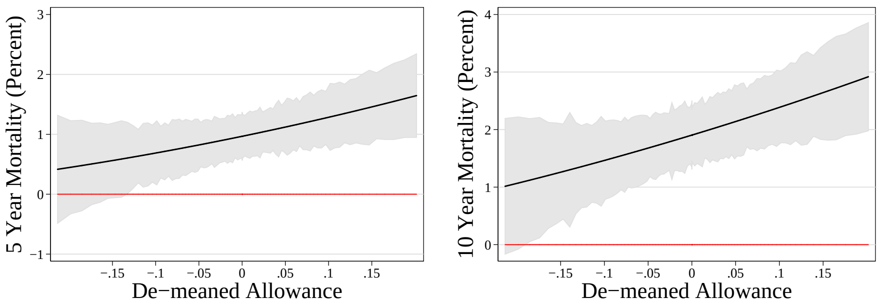
Figure 3.4 presents how the MTE (i.e., the mortality response for the marginal case allowed) varies with predicted de-meaned allowance for different age groups. In each age panel the left



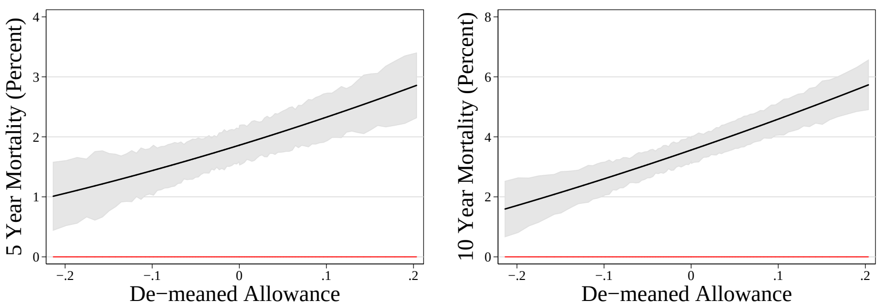
(a) Aged 55-64



(b) Aged 45-54



(c) Aged 25-44



(d) All Ages (25-64)

Notes: This figure displays the estimated mortality response as a function of predicted de-meaned allowance. We control for the covariates listed in Table 3.2. Within each panel: the left figure displays 5 year mortality, and the right figure displays 10 year mortality. Mean allowance rate is 0.84 for those aged 55-64, 0.71 for those aged 45-54, 0.63 for those aged 25-44, and 0.71 for those aged 25-64.

Figure 3.4. Marginal Treatment Effects: Mortality Response by De-meaned Allowance

figure displays 5 year mortality, and the right figure displays 10 year mortality, controlling for covariates.

We use third order polynomials for both the instrument and the endogenous variable (de-meaned allowance) when estimating equations (3.6) and (3.7). The cubic specification is flexible, although visual inspection of Figure 3.4, as well as both the Akaike and Bayesian information criterion show that there is little gain from going beyond the quadratic specification. In appendix 3A.2.3 we show that these results change only modestly when excluding covariates or when using a local polynomial smoother following Maestas et al. (2013).

Since polynomial smoothers have poor endpoint properties, we show estimated MTEs over the middle 90% of the distribution of de-meaned allowance rates. In Monte Carlo experiments, we found our procedure produced little bias over this range. Figure 3.4 also shows bootstrapped 95% confidence intervals.

Consider first panel (a), which shows estimates for our main estimation sample, ages 55-64. The estimated MTE is close to zero at the average predicted allowance rate, at both 5 and 10 years, but there is strong heterogeneity in the responses. Being allowed benefits reduces 5 year mortality by an estimated 0.8 percentage points for the marginal applicant heard by an ALJ who is stricter than 95% of all judges. These judges have allowance rates that are twelve percentage points below the average (de-meaned by hearing office and day). However, allowance increases mortality by 4.5 percentage points for the marginal applicant heard by an ALJ who is more lenient than 95% of all judges. These judges have allowance rates that are ten percentage points above the average. The 10 year mortality response of those 55-64 is qualitatively similar to the 5 year response. The magnitudes are larger, which is unsurprising, given that the impacts of allowance have more years to accumulate.

Estimates for the other age groups follow the same basic patterns as for those aged 55-64: greater leniency implies higher recipient mortality for the marginal applicant.

In summary, our results suggest that making the DI screening threshold significantly less strict (and thus increasing the allowance rate) will increase mortality of the marginal applicants, at least for those assigned to more lenient judges. Interestingly however, for the 55-64 year olds, our 5 year mortality estimates suggest that increasing the screening threshold for the strictest judges would not increase mortality. This provides some evidence, at least for the 55-64 year olds, that current screening thresholds are about right. Our evidence also suggests that the screening threshold could be made stricter for younger age groups without worsening – and likely increasing – their longevity.

Our results are consistent with the notion that as allowance rates rise, more healthy individuals are allowed DI. Healthier individuals benefit less from Medicare and Medicaid insurance from

DI allowance. These individuals also have a bigger decline in labor supply in response to DI receipt (see French and Song, 2014). This is due to more of them being able to work in the absence of DI receipt. Since working potentially has beneficial effects on mortality, any adverse effect of not working will be stronger; younger recipients will also lose more working years. Thus, any beneficial effect of DI allowance is smaller, and the adverse effect larger, for healthier (and younger) individuals.

Given that the average allowance rate is 71% for those aged 25-64 (and is 84% for those aged 55-64), and this excludes the sickest individuals who are allowed at the initial level, the average recipient is substantially less healthy than the marginal applicant. For this reason we would expect the average recipient to be positioned well off to the left of the MTE graphs. If the MTE curve continues to slope down and to the left − which is plausible, but unprovable − this suggests that receiving disability benefits reduces mortality, perhaps strongly so, for the average applicant. This is true even though DI receipt increases mortality, on average, for the applicants who are affected by our judge leniency instrument.

3.6.5 Heterogeneity in the Mortality Effect Based on Observables

Table 3.5 disaggregates the 5 and 10 year mortality response by demographics, prior earnings, and health conditions. The left panel shows 5 year mortality estimates and the right panel shows the 10 year mortality estimates, for applicants aged 55-64. Each panel reports the unadjusted mean mortality for allowed and denied individuals, the OLS estimate of allowance on mortality with covariates, the IV estimate of allowance on mortality with covariates, and the standard error. Table 3.5 shows that the effect of DI allowance on 10 year mortality does not vary in a dramatic way across subgroups. Other than the subgroups for specific health conditions (bottom rows), all subgroup IV estimates are positive, most are statistically significant, and the 95% confidence interval for related subgroups generally overlap. The principal difference across subgroups is that the higher mortality for whites is smaller at both 5 and 10 years than for other racial groups.

The subgroups based on health condition listed in the disability application are listed in order of decreasing 5 year mortality rates. Sample sizes are generally much smaller and standard errors are much larger, but there are some suggestive differences. Individuals diagnosed with neoplasms (e.g. cancer) have the highest overall mortality rates, and have *higher* mortality rates when denied, in both the OLS estimates and the 10 year IV estimates (the 5 year IV estimate is close to zero). This is potentially evidence that DI, and the associated health care benefits, are more valuable to those with cancer than other disabilities. Perhaps health insurance is of special value to this group, given both the high cost of treating cancer, and the high mortality of those with cancer. Note too that the second highest mortality group, with respiratory disease, has a negative IV estimate at 5 years, and a near-zero estimate at 10 years. We investigate these hints

		Panel A: 5 Year Mortality (Percent)						Panel B: 10 Year Mortality (Percent)					
		Mortality Rates		OLS		IV		Mortality Rates		OLS		IV	
	Obs	Allowed	Denied	Diff.	SE	Diff.	SE	Allowed	Denied	Diff.	SE	Diff.	SE
All groups	610,231	9.71	8.35	1.94	(0.12)	2.30	(0.50)	21.95	19.99	2.77	(0.18)	2.81	(0.91)
Sex													
Male	291,994	12.42	10.96	2.20	(0.19)	2.73	(0.98)	27.14	25.65	2.59	(0.27)	3.46	(1.50)
Female	318,237	7.23	5.91	1.67	(0.15)	1.88	(0.58)	17.22	14.67	2.89	(0.22)	2.16	(0.83)
Race													
White	415,125	9.64	8.78	1.67	(0.15)	1.60	(0.64)	22.07	20.72	2.59	(0.22)	2.32	(1.28)
Black	98,698	11.09	9.33	2.70	(0.28)	3.39	(1.50)	23.83	21.83	3.45	(0.40)	4.17	(2.09)
Other	96,408	8.55	6.04	2.29	(0.24)	4.05	(0.93)	19.48	15.92	3.17	(0.37)	3.90	(1.33)
Education Group													
Less than high school	218,871	9.76	8.40	1.73	(0.20)	2.63	(0.72)	22.51	20.44	2.30	(0.29)	2.12	(1.04)
High school graduate	267,634	9.64	8.24	2.18	(0.18)	2.54	(1.00)	21.73	19.86	3.03	(0.27)	3.45	(1.70)
Some college	77,685	9.78	8.45	1.86	(0.32)	0.33	(1.29)	21.89	20.00	3.34	(0.46)	2.85	(2.08)
College graduate	46,041	9.72	8.50	1.72	(0.40)	2.70	(1.87)	20.71	18.69	2.75	(0.55)	2.84	(2.24)
Income													
Average earnings $<$ \$10000	283,146	10.78	9.29	1.80	(0.17)	2.64	(0.62)	23.90	21.79	2.28	(0.24)	2.26	(0.98)
Average earnings $\geq$ \$10000	327,085	8.89	6.77	2.23	(0.17)	1.99	(0.79)	20.47	16.95	3.61	(0.25)	3.74	(1.43)
Health conditions													
Neoplasms	20,000	25.80	28.26	-1.55	(1.02)	0.33	(8.15)	40.93	43.22	-0.43	(1.21)	-1.59	(9.66)
Respiratory system	30,066	14.53	12.09	2.84	(0.57)	-3.72	(2.70)	32.71	30.10	3.75	(0.82)	0.49	(3.55)
Endocrine system	39,331	14.19	10.51	3.38	(0.49)	0.56	(2.99)	32.11	25.42	5.97	(0.68)	1.89	(5.75)
Circulatory system	103,725	11.95	9.53	2.38	(0.30)	4.69	(1.28)	27.73	23.76	2.99	(0.43)	3.96	(2.12)
Mental retardation	3,193	11.00	9.98	4.49	(1.72)	4.76	(7.26)	22.26	21.63	3.89	(2.37)	9.59	(12.13)
Nervous system	34,444	9.41	8.60	2.38	(0.47)	2.04	(2.16)	21.88	21.20	2.69	(0.66)	0.71	(2.75)
Mental disorders	61,508	8.33	7.26	1.81	(0.32)	3.56	(1.27)	19.20	17.88	2.58	(0.45)	4.22	(1.91)
Injuries	27,091	7.62	6.22	1.51	(0.16)	1.82	(1.11)	17.95	15.92	2.32	(0.25)	2.59	(1.26)
Musculoskeletal disorders	231,391	5.94	5.01	2.48	(0.48)	2.29	(1.87)	15.02	13.72	3.67	(0.73)	4.11	(3.26)
All other	59,482	12.11	11.00	2.26	(0.39)	3.08	(2.16)	25.03	24.00	2.84	(0.54)	2.95	(2.34)

Notes: This table displays the estimated effect of DI recipiency on mortality (with covariates) for those aged 55-64. Mortality Rates do not adjust for covariates. OLS and IV estimates are demeaned and include covariates. Covariates are those in Table 3.2; they include race, sex, age and education groups, health (disability category), average earnings and participation prior to disability, representation by an attorney, and an indicator of concurrent SSDI application. IV estimates use demeaned variables and judge leniency as the instrument. Standard errors clustered by judge.

Table 3.5. Estimated Effect of DI Recipiency on Mortality, Aged 55-64, Disaggregated

of differential effects based on health condition, and the cost of treating that condition, in the next section.

3.7 The Channels by which DI Affects Mortality

We do not find any adverse effect of being denied benefits on mortality for the marginal applicant. Yet, as we show below, cash income and health insurance transfers to the disabled are large, which would suggest lower mortality, other factors equal. This leaves the effect of receiving benefits on labor supply as a potential offsetting effect. As noted previously in Section 3.2.3, many studies have shown that DI receipt reduces employment, and other studies suggest that employment reductions can increase mortality. In this section we discuss some channels by which allowance could impact mortality.

We summarize the quantitative magnitude of these channels in Table 3.6. In this table we display several outcomes for individuals denied by an ALJ and calculate the difference in these outcomes between those allowed versus denied.

3.7.1 Allowance

We estimate the effect of ALJ allowance on mortality. However, many individuals who are initially denied are eventually allowed upon reapplication or appeal. In this sense we have an "intent to treat" estimate, rather than a "treatment effect on the treated" estimate. We estimate the impact of initial allowance by an ALJ, rather than final allowance, because final allowance depends on mortality: only still-living persons can receive benefits after appeal. However, appeals and re-applications are important for understanding the magnitude of the effect of allowance on benefits received.

Panel (1) of Table 3.6 shows outcomes for those denied by an ALJ 1, 3, and 5 years after assignment to an ALJ. Row A shows that 54% of those denied by an ALJ are allowed within 5 years.

Panel (2) displays the difference in outcomes between those allowed versus denied at the ALJ stage. For example, virtually 100% of those allowed benefits are still receiving benefits 5 years later, whereas 54% of those denied by an ALJ are allowed 5 years later, therefore the difference is 100-54=46%. This can be seen in the 5 year OLS estimate of row A.

The results of Table 3.6 take into account that many persons who are denied benefits at the ALJ stage are later allowed. We present calculation details in appendix section 3A.4, and provide more information on the data sources behind our estimates.

	(1) Outcomes for denied			(2) Difference in outcomes between those allowed versus denied at ALJ stage						
	Years after ALJ stage			1 year later		3 years later		5 years later		Total discounted
	1	3	5	OLS	IV	OLS	IV	OLS	IV	benefits up to age 65*
A. Prob of being allowed at future times	0.24	0.42	0.54	0.76	0.71	0.58	0.51	0.46	0.40	
B. Prob of earnings > 0	0.22	0.20	0.16	-0.12	-0.12	-0.11	-0.10	-0.08	-0.07	
C. Prob of earnings > SGA	0.08	0.08	0.06	-0.06	-0.06	-0.06	-0.06	-0.05	-0.04	
D. Cash income	4,265	5,733	6,471	5,969		4,204		2,958		24,039
D(i). Cash benefits	2,313	4,139	5,433	7,866		6,046		4,804		33,799
D(ii). Average earnings, before taxes	2,402	1,950	1,314	-2,182	-2,365	-2,111	-2,256	-1,943	-1,586	-11,103
D(iii). Average earnings, net of tax	1,953	1,594	1,037	-1,897		-1,843		-1,846		-9,760
E. Prob of receiving Medicare/Medicaid	0.12	0.40	0.50	0.16		0.48		0.39		
F. Annual Medicare/Medicaid payments	1,457	4,809	6,037	1,926		5,784		4,736		23,038
G. Total dollar value	5,723	10,541	12,507	7,665		8,793		6,182		47,077

Notes: Panel (1) displays the predicted average outcomes for those denied at the ALJ stage. Panel (2) displays the difference in outcomes between those allowed versus denied at ALJ stage. OLS and IV esimtates in panel (2) control for covariates. The average predicted outcomes for those allowed at the ALJ stage is the sum of the relevants cells in panels (1) and (2). The calculations assume individuals were under age 60 at assignment. All dollar amounts in 2014 dollars.
Row A: probability of being allowed at future times. Source: French and Song (2014).
Row B: probability of having positive earnings. Source: Our data.
Row C: probability of earning above the Substantial Gainful Activity Level ($12,480 per year in 2014). Source: Our data.
Row D: predicted cash benefits plus after tax income. Source: Our data.
Row D(i): predicted cash benefits received after deducting the average reduction in benefits due to work. Source: French and Song (2014).
Row D(ii): predicted average earnings before tax. Source: Our data.
Row D(iii): predicted average earnings after tax. Source: Source: Our data and French and Song (2014).
Row E: probability of receiving Medicare and/or Medicaid. Source: Rupp and Riley (2012) and the appendix.
Row F: average annual medical payments from Medicare and/or Medicaid. Source: Section 2.5 of this thesis and the appendix of De Nardi et al., 2016b.
Row G: total dollar value difference of predicted cash income, benefits, taxes, and medical payments from Medicare and/or Medicaid.
* The total discounted values assume that an individual is first seen by an ALJ at age 58, which is the median age at assignment in our sample. Benefits are cumulated through age 65 (7 years later), and discounted using an interest rate of 3% and the observed mortality rate for those allowed by an ALJ in our sample.

Table 3.6. Key Outcome Differences Between Those Allowed versus Denied

3.7.2 The Income Benefit and Labor Supply Incentives

One potentially important determinant of mortality is income. There are many possible channels through which income can affect health, including through investment in health through better food, shelter, and health care. In this section we discuss how income responds to benefit allowance. Specifically, we focus on the response of taxable earnings and DI/SSI benefits to benefit allowance.

Both income effects (through the high replacement rate) and substitution effects (beneficiaries will lose benefits if they earn above the SGA amount) causes DI recipients to reduce labor supply. DI/SSI benefits likely also reduce labor supply through a third channel – health insurance, which greatly reduces the value of employer-provided health insurance, which can be an important work incentive (French and Jones, 2011).

Row B of Table 3.6 presents estimates of the employment response to being allowed disability benefits. Panel (1) shows that 16% of all individuals denied by an ALJ have positive earnings 5 years after assignment to an ALJ. The OLS estimates in Panel (2) show that being allowed benefits by an ALJ reduces employment rates by 8 percentage points after 5 years, with similar IV estimates. The OLS estimates in row C show that being allowed by an ALJ reduces the probability that earnings exceed the SGA limit (of $12,480 in 2014) by 5 percentage points after 5 years; IV estimates are again similar. These reductions in employment lead to significant declines in earnings: pre-tax earnings fall when allowed by $1,943 after 5 years (see row D(ii)), although the post-tax earnings loss is somewhat smaller (see row D(iii)).

Total cash income rises after allowance, since the cash value of DI/SSI benefits exceed the decline in income. The average extra value of these benefits for those allowed at the ALJ stage averages $5,969 1 year after being allowed by an ALJ, but falls to $2,958 5 years after. This fall occurs because many of those initially denied are later allowed upon appeal or re-application or because they are old enough to receive Social Security benefits.

We should note that we cannot assess all channels by which DI/SSI receipt may affect household income. For example, Autor et al. (2015) show that show that in Norway disability benefit receipt also leads to reductions in spouse's earnings and other benefits (such as unemployment insurance).

3.7.3 Health Insurance Benefits

Individuals receiving DI benefits are eligible for Medicare after a two year waiting period. Individuals drawing SSI are often also immediately eligible for Medicaid, the government health insurance program for the poor. Livermore et al. (2011) show that federal and state governments spend more on health care than on cash benefits for the disabled.

Rupp and Riley (2012) report the percentage of DI beneficiaries receiving either Medicare or Medicaid over a period covering 12 months before they were awarded DI until 6 years after. They show that immediately following DI/SSI benefit receipt, 24.7% receive either Medicaid or Medicare, the majority being SSI beneficiaries who receive Medicaid. The total jumps to 89.7% just after 2 years when DI beneficiaries become eligible for Medicare, and reaches 96.8% after 6 years.

Using the values from Rupp and Riley (2012) and the calculations explained in appendix 3A.4 we calculate the difference in the probability of receiving Medicare or Medicaid between those allowed versus denied at ALJ stage, taking into account that many of those denied by an ALJ are later allowed. These results are shown in row E of Table 3.6. The higher probability of receiving Medicare and/or Medicaid is fairly small 1 year later at only 16 percentage points, peaks at 3 years later when almost everyone allowed by the ALJ is receiving Medicare, and then decline as many of the initially denied are later allowed.

Using data from Section 2.5 of this thesis (also found in the appendix of De Nardi et al., 2016b) we calculate that the average Medicare and/or Medicaid recipient receives \$12,012 worth of medical transfers from Medicare/Medicaid per year. Row F of Table 3.6 calculates the difference in the average annual medical payments by multiplying \$12,012 by the difference in probability of receiving Medicare/Medicaid (row E). This means that 1 year later those allowed are receiving on average \$7,665 more in medical transfers. After 5 years this difference is \$6,182 per year.

3.7.4 Total Discounted Value of Income and Benefits

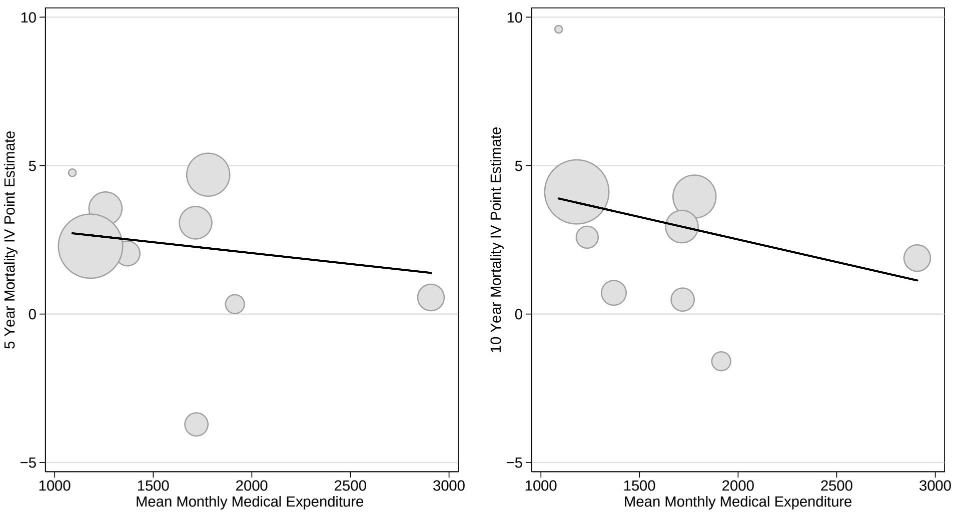
The final column in Table 3.6 shows the present discounted value of all income and benefits that arise from being allowed DI by an ALJ up to age 65, when everyone should become eligible for Medicare and Social Security benefits

To calculate this we assume that everyone in the age 55-64 group is age 58, which is the median age for this group in our sample. We discount future benefits and income using an interest rate of 3%, taking into account that not everyone lives to age 65, using the mortality rates for those allowed by an ALJ in our sample. We estimate that the average total discounted value of income and benefits of being awarded DI by an ALJ is \$47,077. Of this, 51% is in cash income and 49% in medical transfers. These are substantial amounts which, other factors equal, would be expected to reduce mortality.

3.7.5 Effects Disaggregated by Health Condition

In Table 3.5 we find some evidence that mortality responses vary for different reported applicant health conditions. In Figure 3.5, we investigate further the hints from that table that the effect of benefit allowance on mortality is more favorable (less adverse) for more expensive health conditions and for conditions that predict higher near-term mortality.

Figure 3.5 plots the 5 and 10 year mortality point estimates by health condition from Table 3.5 against mean monthly health care spending for that condition (in thousands of 2014 US dollars) by health condition from the Medicare Current Beneficiary Survey (MCBS).[14] We calculate mean medical spending for disabled Medicare beneficiaries under age 65.[15] The size of the circles represents the number of observations.



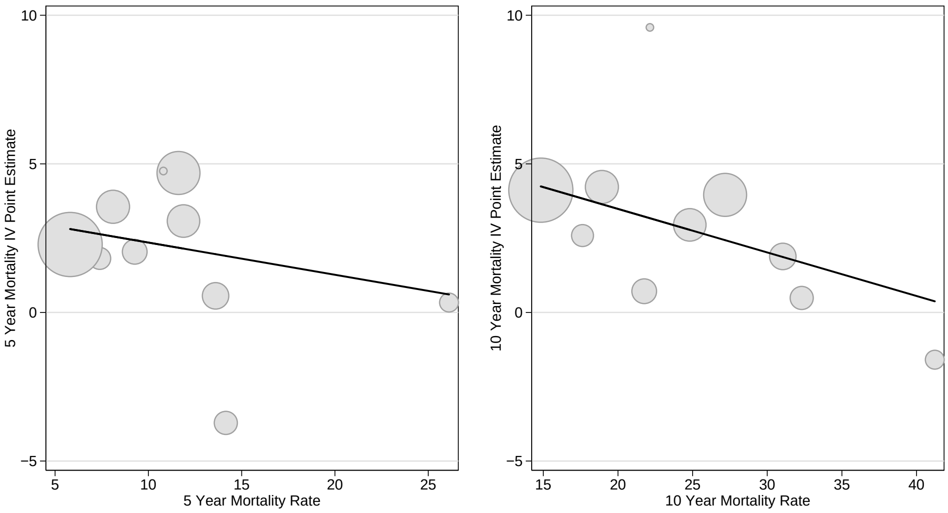
Notes: This figure displays a scatter plot of the 5 year (left graph) and 10 year (right graph) mortality IV point estimates by health condition from Table 3.5 plotted against mean monthly spending for that condition (in thousands of 2014 US dollars), from the MCBS. Circle size is proportional to number of disability applicants with that condition. The line represents predicted mortality from a regression of the health condition specific mortality point estimates against mean monthly spending, weighted by the number of individuals in each condition group from the SSA.

Figure 3.5. Estimated 5 Year and 10 Year Mortality Effect of Allowance by the Medical Expenditure for Each Health Condition

Over both 5 year and 10 year periods, we find a general tendency, albeit with substantial scatter, for benefit allowance to be less adverse to mortality (averaged over the range of judge leniency we observe) for higher-cost medical conditions. This is consistent with the view that access to health insurance, and thus potential access to better healthcare, reduces mortality for those with more expensive conditions, and can offset any adverse effect of work disincentives.

[14]We use estimates from Section 2.5 of this thesis.

[15] More precisely, we use those receiving Medicare benefits who are younger than 65. Virtually everyone under age 65 who receives Medicare also receives disability benefits. The MCBS has high quality medical spending data since it uses administrative Medicare records for Medicare spending and a mixture of survey data and reconciliation of survey, Medicaid participation, and Medicare records to infer payments by other payors. De Nardi et al. (2016b) find that the MCBS captures approximately 80% of total medical spending for its target population and French et al. (2017b) find that out of pocket spending and private insurance information match up well between MCBS and the Health and Retirement Study. An attractive aspect of the MCBS data is that respondents are asked about the main health condition that caused them to be eligible for Medicare benefits. Thus we can match the condition that led to allowance in both the Social Security data and the MCBS data.

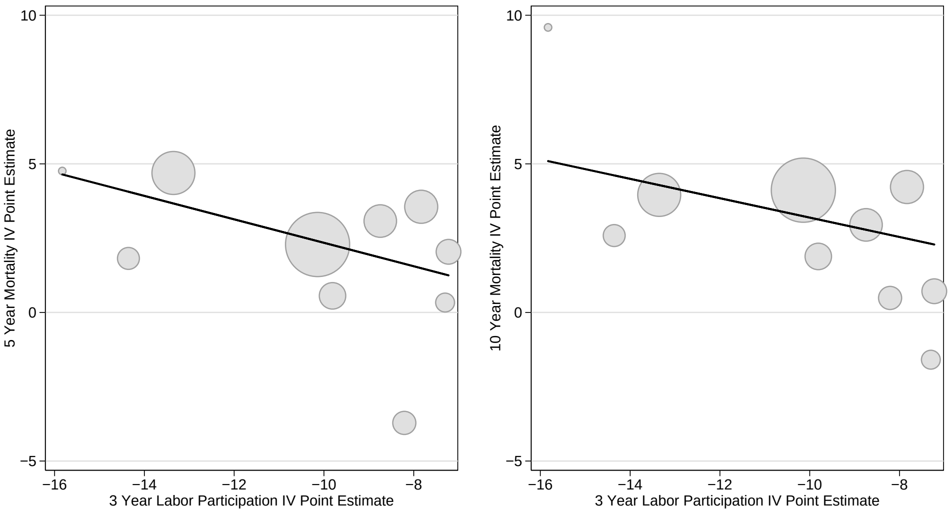


Notes: This figure displays a scatter plot of the 5-year and 10-year mortality IV point estimates by health condition from Table 3.5 plotted against the unconditional mortality rate of the individuals with each condition from SSA mortality records.

Figure 3.6. Estimated 5 Year and 10 Year Mortality Effect of Allowance by the Mortality Rate for Each Health Condition

Figure 3.6 plots the 5 and 10 year mortality point estimates by health condition from Table 3.5 against the average 5 and 10 year mortality rate for that condition. Over 5 years, we find either no increase or a predicted decline in mortality, among those with neoplasms (e.g., cancer), respiratory conditions and problems with the endocrine system, which are the highest mortality rate conditions. These conditions are also amongst the most expensive in terms of medical treatment. Conversely, conditions with relatively low mortality, which also tend to have lower medical spending, such as mental retardation, mental disorders, and musculoskeletal disorders, have increased average marginal mortality following benefit receipt. Similar to Figure 3.5, Figure 3.6 is consistent with the view that improved access to health care can reduce mortality for expensive, high mortality conditions. The negative slopes in Figures 3.5 and 3.6 are not statistically significant, however, and should only be taken as suggestive evidence.

Any effects due to improved access to health care for high cost conditions are likely offset by increases in mortality associated with reduced employment. This can be seen in the relationship plotted in Figure 3.7, which shows that the health conditions where receipt of benefits increases mortality by the most are also the conditions where the receipt of benefits decreases labor supply the most. This is consistent with the view that the work disincentive from receiving benefits could increase mortality.



Notes: This figure displays a scatter plot of the 5 year mortality IV point estimates by health condition from Table 3.5 plotted against the 3 year labor supply IV point estimates for each condition group from the SSA.

Figure 3.7. Estimated 5 Year and 10 Year Mortality Effect of Allowance by the Labor Supply Response for Each Health Condition

3.8 Robustness

Our results for the main estimation sample (those aged 55-64 at time of application) are robust to a number of other modifications to sample selection and functional form. Table 3.7 provides robustness checks. The left panel shows 5 year mortality estimates and the right panel shows the 10 year mortality estimates. In each panel, odd-numbered columns reports the estimates with no covariates; even columns present estimates with covariates.

The first two rows display OLS estimates. In the second row, we include the 10,006 individuals who died within a year of seeing a judge. As discussed in Section 3.5, we exclude these individuals as we are concerned that some of our sample who are denied are likely just those who die before being heard by a judge. Including these individuals decreases the coefficients by a small amount without covariates, but increases the coefficients slightly with covariates; thus, our choice to exclude these individuals does not meaningfully affect our overall findings.

The remaining rows provide IV estimates in various different specifications or with different sample selections. Including persons who died within a year of seeing a judge has only a small effect on our IV estimates. The next two rows change the number of judges we exclude due to them seeing a limited number of cases. Whereas in our baseline specification we exclude judges who heard less than 200 cases, here we consider a lower threshold of 50 cases and a higher

	Panel A: 5 Year Mortality (Percent)		Panel B: 10 Year Mortality (Percent)	
	No Covariates	With Covariates	No Covariates	With Covariates
OLS				
Baseline	1.35	1.94	1.87	2.77
	(0.12)	(0.12)	(0.19)	(0.18)
Inc. those who die year of app.	1.15	1.99	1.69	2.82
	(0.14)	(0.13)	(0.19)	(0.18)
IV				
Baseline	1.81	2.30	1.93	2.81
	(0.44)	(0.50)	(0.76)	(0.91)
Inc. those who die year of app.	1.73	2.30	1.82	2.81
	(0.44)	(0.50)	(0.75)	(0.90)
Drop Judges who saw $<$ 50 cases	1.80	2.28	1.91	2.78
	(0.45)	(0.51)	(0.78)	(0.93)
Drop Judges who saw $<$ 500 cases	1.76	2.27	1.80	2.73
	(0.46)	(0.54)	(0.79)	(0.95)
Drop Middle Third of Judges	1.81	2.29	1.94	2.79
	(0.43)	(0.49)	(0.76)	(0.91)
Doyle's Instrument	1.41	1.88	1.53	2.38
	(0.45)	(0.51)	(0.70)	(0.84)
Demean by hearing office-year	1.94	2.43	2.35	3.18
	(0.49)	(0.63)	(0.96)	(1.24)
Underreporting Correction				
Baseline Correction	1.62	2.12	1.66	2.54
	(0.44)	(0.50)	(0.76)	(0.90)
Double Size of Base Correction	1.43	1.92	1.18	2.07
	(0.44)	(0.49)	(0.74)	(0.89)
Set p = .94 for Denied	1.38	1.87	0.86	1.75
	(0.44)	(0.50)	(0.76)	(0.91)
Set p =.88 for Denied	0.73	1.24	-0.73	0.18
	(0.44)	(0.50)	(0.75)	(0.90)

Notes: Baseline instrument is judge leniency. Covariates are those in Table 3.2; they include race, sex, age and education groups, health (disability category), average earnings and participation prior to disability, representation by an attorney, and an indicator of concurrent SSDI application. For details on how the correction for underreporting of mortality is calculated, see the discussion in section 3.5.2. Standard errors clustered by judge.
In the Drop Middle Third of Judges row we only keep judges in the top and bottom thirds of the distribution of judge leniency.
In the Doyle's Intrument row we replace the baseline instrument with the one constructed in the appendix section 3A.3.2
In the baseline we demean by hearing office-day and drop judges who saw less than 200 cases, which gives N=610,231.
In the rows where we include those who die within 1 year of seeing a judge, N = 620,237.
In the rows where we drop judges who saw $<$ 50 cases, N = 616,599.
In the rows where we drop judges who saw $<$ 500 cases, N =601,042.

In the rows where we drop the middle third of judges, N = 408,853.

Table 3.7. Robustness Checks

threshold of 500 cases - neither of which change our estimates by much. In the next row we see if we can increase the strength of our instrument by only keeping judges in the top and bottom third in the distribution of judge leniency, our instrumental variable. Our estimates and standard errors are almost unchanged. The next row displays the results from the IV regression where we use the instrument proposed by Doyle Jr (2007) instead of our judge leniency instrument. With Doyle's instrument, our coefficients are somewhat smaller than the baseline but inference is similar. In the next row, instead of demeaning by hearing-office and day as in the baseline, we demean by hearing-office year. The estimates tend to be a bit higher, but standard errors also increase. The bottom four rows all adjust for underreporting, using different values for the correction, p. Our baseline underreporting correction uses individual specific values for p (which average 0.95 and 0.96 at 5 and 10 years, respectively), as described in section 3.5.2. Next, we assign individual values for p, assuming the undercount is twice as large as for the baseline correction. This adjustment modestly decreases our point estimates. The next underreporting correction assumes that the value of p is the one estimated for individuals aged 25-44, whose mean value of $p = 0.94$, which is much lower than for the baseline. We assign this value of p for all denied individuals. Our final underreporting correction takes an extreme value for p by assuming the undercount for the individuals aged 25-44 is twice as large and assigning this value for everyone. As expected, with the correction the estimates fall but only slightly. It would take a very large undercount of deaths of denied individuals – much larger than is plausible – to change our core inferences.

These robustness checks, taken together, increase our confidence in our estimation strategy. In every case, barring implausibly large underreporting corrections, our estimates are positive, statistically significant, and similar in magnitude to our main estimates.

3.9 Conclusion and Policy Implications

This paper estimates the effect of Disability Insurance receipt on mortality, for persons on the margin of receiving benefits or not. Those receiving benefits receive large cash transfers, and health insurance from Medicare or Medicaid. However, beneficiaries also face important work disincentives. Each of these factors could affect mortality, but not necessarily with the same sign. We would expect higher income and access to health insurance to cause lower mortality. However, reduced employment may increase mortality. Identifying this combined effect is difficult, however, because those allowed benefits are likely to be less healthy than those not allowed, perhaps in ways observed by the ALJ, but not fully captured by our covariates. We rely for causal inference on the effectively random assignment of judges to disability cases, and on an instrumental variable that measures the tendency for each judge to allow benefits, relative to other judges in the same hearing office on the same day.

We find that benefit receipt increases mortality on average, for those on the margin to receive benefits or not, after both 5 and 10 years. This is consistent with the view that benefit receipt lowers labor supply, which in turn increases mortality. However, Marginal Treatment Effects estimates reveal strong heterogeneity in the response to benefit allowance, even within the range of leniency that we observe. For those aged 55-64, allowance reduces mortality for less healthy applicants who would be allowed by all but the strictest judges, but increases mortality for healthier applicants who would only be allowed by the most lenient judges. These results suggest that significant changes in the DI screening threshold may increase mortality of the marginal applicants. This provides some evidence that current screening thresholds are about right. For younger age groups we find that a modest tightening of the screening thresholds (i.e. making them more strict) will not decrease the longevity of applicants and might increase the longevity of marginal applicants.

Our estimates show that among the healthier individuals, DI receipt increases mortality, but among the less healthy individuals, DI receipt tends to reduce estimated mortality. All of our estimates are for *marginal* recipients, who would receive benefits if seen by a lenient ALJ, but be denied by a stricter one. However, the majority of benefit recipients are *inframarginal* cases who are less healthy than our marginal cases. Thus, our findings are consistent with the view that DI receipt reduces mortality on average.

We also find evidence that for certain expensive, high mortality health conditions such as cancer and respiratory conditions, benefit receipt is relatively more favorable for future mortality, whereas for lower cost, lower mortality conditions such as musculoskeletal disorders, benefit receipt predicts higher mortality for marginal applicants.

Our findings have important policy implications. Given the extreme cost of the disability insurance program, many reform proposals have been put forward, including making the disability criteria more stringent. Our results speak directly to how increasing stringency might impact applicants' health. In general our findings suggest that for maximizing the longevity of current DI applicants, the current disability thresholds are at about the right level. However, a modest increase in program stringency should not increase mortality, and might decrease it, especially for younger applicants.

We also find evidence for the value of health insurance for selected high-cost, high-mortality conditions. This suggests that persons with these conditions who receive benefits might gain from not being subject to the current 2 year waiting period to receive Medicare coverage.

We provide evidence that working appears to reduce mortality, at least on average, for marginal recipients. Thus, reforms of the disability insurance rules to reduce the strong work disincentives of the current rules may improve recipient mortality.

Bibliography

Abadie, A. (2003). Semiparametric instrumental variable estimation of treatment response models. *Journal of econometrics*, 113(2):231–263.

Autor, D. and Duggan, M. (2006). The growth in the social security disability rolls: A fiscal crisis unfolding. *The Journal of Economic Perspectives*, 20(3):71–96.

Autor, D., Kostol, A. R., and Mogstad, M. (2015). Disability benefits, consumption insurance, and household labor supply.

Autor, D. H. and Duggan, M. (2010). Supporting work: a proposal for modernizing the us disability insurance system. *Center for American Progress and The Hamilton Project*.

Benitez-Silva, H., Buchinsky, M., Chan, H. M., Rust, J., and Sheidvasser, S. (1999). An empirical analysis of the social security disability application, appeal, and award process. *Labour Economics*, 6(2):147–178.

Bjorklund, A. and Moffitt, R. (1987). The estimation of wage gains and welfare gains in self-selection models. *The Review of Economics and Statistics*, pages 42–49.

Black, B., Espin-Sanchez, J.-A., French, E., and Litvak, K. (2017). The long-term effect of health insurance on near-elderly health and mortality. *American Journal of Health Economics*.

Black, B., French, E., McCauley, J., Faria de Carvalho Nunes, L., and Song, J. (2016). Comparing death data from the social security administration and the national center for health statistics, 1990-2013.

Brook, R., Ware, J. E., Rogers, W. H., Keeler, E. B., Davies, A. R., Donald, C. A., Goldberg, G. A., Lohr, K. N., Masthay, P. C., and Newhouse, J. P. (1983). Does free care improve adults' health? results from a randomized trial. *New England Journal of Medicine*, 309(23):1426–1434.

Burkhauser, R. V. and Daly, M. (2011). *The declining work and welfare of people with disabilities: What went wrong and a strategy for change*. AEI Press.

Burkhauser, R. V., Daly, M. C., McVicar, D., and Wilkins, R. (2014). Disability benefit growth and disability reform in the us: lessons from other oecd nations. *IZA Journal of Labor Policy*, 3(1):4.

Card, D., Dobkin, C., and Maestas, N. (2009). Does medicare save lives? *The Quarterly Journal of Economics*, 124(2):597–636.

Dahl, G. B., Kostol, A. R., and Mogstad, M. (2014). Family welfare cultures. *The Quarterly Journal of Economics*, 129(4):1711–1752.

De Nardi, M., French, E., Jones, J. B., and McCauley, J. (2016b). Medical spending of the us elderly. *Fiscal Studies*, 37(3-4):717–747.

Doyle Jr, J. J. (2007). Child protection and child outcomes: Measuring the effects of foster care. *The American Economic Review*, 97(5):1583–1610.

Eliason, M. and Storrie, D. (2006). Lasting or latent scars? swedish evidence on the long-term effects of job displacement. *Journal of Labor Economics*, 24(4):831–856.

Finkelstein, A. and McKnight, R. (2008). What did medicare do? the initial impact of medicare on mortality and out of pocket medical spending. *Journal of Public Economics*, 92(7):1644–1669.

Finkelstein, A., Taubman, S., Wright, B., Bernstein, M., Gruber, J., Newhouse, J. P., Allen, H., Baicker, K., and Group, O. H. S. (2012). The oregon health insurance experiment: Evidence from the first year. *Quarterly Journal of Economics*, 127(3):1057–1106.

Fitzpatrick, M. and Moore, T. (2016). The mortality effects of retirement: Evidence from social security eligibility at age 62. mimeo.

French, E. and Jones, J. B. (2011). The effects of health insurance and self-insurance on retirement behavior. *Econometrica*, 79(3):693–732.

French, E., Jones, J. B., and McCauley, J. (2017b). The accuracy of economic measurement in the health and retirement study. In *Forum for Health Economics and Policy*. De Gruyter.

French, E. and Song, J. (2014). The effect of disability insurance receipt on labor supply. *American Economic Journal: Economic Policy*, 6(2):291–337.

French, E. and Taber, C. (2011). Identification of models of the labor market. *Handbook of Labor Economics*, 4:537–617.

GAO (2013). "Preliminary Observations on the Death Master File", Testimony Before the Comittee on Homeland Security and Governmental Affairs.

Gelber, A., Moore, T., and Strand, A. (2017). Disability insurance income saves lives. mimeo.

Hausman, J. A., Newey, W. K., Woutersen, T., Chao, J. C., and Swanson, N. R. (2012). Instrumental variable estimation with heteroskedasticity and many instruments. *Quantitative Economics*, 3(2):211–255.

Heckman, J. J., Urzua, S., and Vytlacil, E. (2006). Understanding instrumental variables in models with essential heterogeneity. *The Review of Economics and Statistics*, 88(3):389–432.

Hernandez-Pizarro, H. (2016). The effect of long-term care benefits on mortality. mimeo.

Hill, M. E. and Rosenwaike, I. (2002). The social security administration's death master file: The completeness of death reporting at older ages. *Social Security Bulletin*, 64:45–51.

Hu, J., Lahiri, K., Vaughan, D. R., and Wixon, B. (2001). A structural model of social security's disability determination process. *Review of Economics and Statistics*, 83(2):348–361.

Imbens, G. W. and Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, 62(2):467–475.

Kuhn, A., Wuellrich, J.-P., and Zweimuller, J. (2017). Fatal attraction? early retirement and mortality.

Livermore, G., Stapleton, D. C., and O'Toole, M. (2011). Health care costs are a key driver of growth in federal and state assistance to working-age people with disabilities. *Health Affairs*, 30(9):1664–1672.

Maestas, N., Mullen, K. J., and Strand, A. (2013). Does disability insurance receipt discourage work? using examiner assignment to estimate causal effects of ssdi receipt. *The American Economic Review*, 103(5):1797–1829.

Myers, R. J. (1993). *Social security*. Univ of Pennsylvania Pr.

Paletta, D. (2011). Disability-claim judge has trouble saying ‘no’; near perfect approval record; social security program strained. *Wall Street Journal*.

Parsons, D. O. (1991). Self-screening in targeted public transfer programs. *Journal of Political Economy*, pages 859–876.

Rege, M., Telle, K., and Votruba, M. (2009). The effect of plant downsizing on disability pension utilization. *Journal of the European Economic Association*, 7(4):754–785.

Rupp, K. and Riley, G. F. (2012). Longitudinal patterns of medicaid and medicare coverage among disability cash benefit awardees. *Soc. Sec. Bull.*, 72:19.

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

Social Security Advisory Board (2006). Improving the social security administration’s hearing process.

Sommers, B. D., Long, S. K., and Baicker, K. (2014). Changes in mortality after massachusetts health care reform: a quasi-experimental study. *Annals of internal medicine*, 160(9):585–593.

Sullivan, D. and von Wachter, T. (2009). Job displacement and mortality: An analysis using administrative data. *Quarterly Journal of Economics*, pages 1265–1306.

Taylor, M. H. (2007). “refugee roulette” in an administrative law context: The" déjà vu" of decisional disparities in agency adjudication. *Stanford Law Review*, pages 475–501.

United States (2007). US Code Title 5, Part III, Subpart B, Chapter 31, Subchapter I, Section 3105.

U.S. Census Bureau (2015). Annual estimates of the resident population for selected age groups by sex for the united states, states, counties, and puerto rico commonwealth and municipios: April 1, 2010 to july 1, 2014.

U.S. Social Security Administration (2007). Annual statistical report on the social security disability insurance program.

U.S. Social Security Administration (2008). Social security bulletin annual statistical supplement.

U.S. Social Security Administration (2009). Hearings, appeals and litigation law manual (HALLEX).

U.S. Social Security Administration (2014a). Annual statistical report on the social security disability insurance program.

U.S. Social Security Administration (2014b). Ssi annual statistical report.

Weathers, R. R. and Stegman, M. (2012). The effect of expanding access to health insurance on the health and mortality of social security disability insurance beneficiaries. *Journal of health economics*, 31(6):863–875.

3A Appendix

3A.1 Data

We use the universe of all DI or SSI appeals heard by ALJs, 1995-2004. We merge data from the following sources: the Office of Hearings and Appeals Case Control System (OHACCS), the Hearing Office Tracking System (HOTS), the Appeals Council Automated Processing System (ACAPS), the Litigation Overview Tracking System (LOTS), the SSA 831 file, SSA Master Earnings file (MEF), the Master Beneficiary Record (MBR), the Supplemental Security Record (SSR), and the SSA Numerical Identification (NUMIDENT) file.

The OHACCS data contain details of Social Security DI and SSI cases adjudicated at the ALJ level (plus limited information on cases heard at the Appeals Council or in federal court). The OHACCS data also include cases involving Retirement and Survivors Insurance and Medicare Hospital insurance. We keep only the SSI and DI cases. The OHACCS data are used for administering DI and SSI cases, and are thus very accurate. They include information on the judge assigned to the case, the hearing office, the date of assignment, and the case outcome (such as allowed or denied), the claimant's Social Security number and type of claim (DI versus SSI). Because the SSA mortality data is less complete prior to 1995, we use OHACCS data only for 1995-2014.

Until 2004, individual hearing offices maintained their own data, called the Hearing Office Tracking System (HOTS). These data were then uploaded to the OHACCS system. We found some missing cases in the OHACCS system, apparently the result of HOTS data not being properly uploaded. The problem occurs in about 1% of all cases. For these cases we augment the OHACCS data with HOTS. After 2004, all uploading of data is automatic, and thus there are no problems with missing data.

Although OHACCS also contains Appeals Council records, Appeals Council decisions are sometimes missing from OHACCS. Thus we use the Appeals Council Automated Processing System (ACAPS) data, which the Appeals Council uses to administer cases, to track outcomes for cases heard at the Appeals Council level.

The Litigation Overview Tracking System (LOTS) data are used by SSA to administer cases that were denied by the Appeals Council but then reach federal courts. We combine the LOTS data with information provided by the Federal Court to determine whether the cases was eventually allowed or denied. The SSA 831 data have information on the details of the DI application received by the Disability Determination Service. The data include the type of application (whether DI or SSI or concurrent) and whether the claim is based on one's own earnings history or on the history of a spouse or parent. It also has all the information relevant for determining whether the application should be allowed at the initial level, before reaching an ALJ, based on the applicant having a listed medical condition or the vocational grid. Thus we have detailed information on applicants' health at time of application. Because of the vocational grid, we have information on age, education, industry and occupation. We also have some other demographic information such as sex. Since a new 831 record is established whenever a new application is filed and adjudicated, we use information in the 831 file to identify those who reapplied for benefits.

The Master Earning File (MEF) includes annual longitudinal earnings data for the US population, taken directly from W-2 filings, starting from 1978. Wage earnings are not top-coded. Self-employment earnings are top coded until 1992. Our earnings measure is the sum of wage earnings and self employment earnings, which we topcode at $200,000 per year, the topcoding affects only a very small percentage of applicants.

The Master Beneficiary Record (MBR) includes beneficiary and payment history data for the entire Social Security OASDI program. The Supplemental Security Record (SSR) contains information on individuals applying for SSI benefits. We use the MBR and SSR to identify disability benefit award status of individuals.

Lastly, we use the SSA NUMIDENT file for information on date of death. The NUMIDENT file includes information from the Social Security Number application form such as name, date of birth and Social Security number, and once the individual dies, the date of death.

For Figure 3.1 we use all cases filed 1989-1999. For all other figures and tables, we begin with the universe of all cases adjudicated by an ALJ and make the sample restrictions, described in Table 3A.1. We drop a relatively small number of cases who died within the year of assignment to the judge, had missing education data, or where the judge handled fewer than 200 cases. This leaves an estimation sample with 2,759,907 cases. In many analyses we further restrict the sample to persons age 55-64 at application, which is 610,231 cases.

3A.2 Additional Tables and Figures

3A.2.1 Main Tables: All Ages

Table 3.2 in the text provides evidence for random assignment for our main estimation sample (age 55-64 at time of application). Table 3A.3 provides a similar table for the full sample. The last two columns show whether the instrument (judge leniency) significantly predicts our covariates which it should not, if assignment is random. Of the 20 covariates in the table, only one takes a coefficient with a t-statistic > 2.0, and only mildly so (female t-statistic = 2.2). This is consistent with random assignment.

	Sample Size
Original data	3,368,017
Number of drops	
Age at application <25 or >64	339,515
Died the year of application	30,807
Missing education data	204,859
Judge handled fewer than 200 cases	32,929
Remaining sample (Aged 25-64)	2,759,907
Age at application: 25-44	1,101,332
Age at application: 45-54	1,048,344
Age at application: 55-64 (Main Sample)	610,231

Notes: The original sample excludes those with missing judge or hearing office information, pre-viewed cases, DI Child cases, and Survivor cases.

Table 3A.1. Sample Selection

Table 3.3 in the main text provides first-stage results for our main estimation sample (age 55-64 at time of application), disaggregated by gender, income, health, and other subgroups of our 55-64 sample. Table 3A.4 provides a similar table for the full sample. The allowance rates are lower for the full sample (70.8%) than for our 55-64 subsample (84.1%). The full sample coefficient from regressing allowance on judge leniency is 0.966, and thus close to 1, as it should be since we estimate judge leniency using the full sample. The comparable estimate for applicants aged 55-64 is 0.676; thus, judge leniency has a larger effect on allowance rates for younger applicants. The monotonicity assumption again cannot be rejected, with most relative likelihoods close to 1.

3A.2.2 Quality of SSA Mortality Data In section 3.5 we described some of the reasons why mortality rates of those denied benefits might be undercounted in the SSA data. By showing that aggregate mortality rates are very similar in both the SSA data and the National Death Index, we provided evidence that this undercount was not a serious issue. In this appendix we provide further evidence that the SSA data accurately measures mortality of those denied benefits.

To provide an alternative approach to measuring mortality undercounts, we note that most individuals who are denied DI benefits will take regular retirement benefits at either ages 60, 62, or 65.[16] And once individuals are receiving benefits, we should expect SSA data to have high accuracy in recording deaths, both for those allowed DI benefits, and those who are receiving regular retirement benefits. Thus, we should expect any undercount of mortality for those who are denied benefits to occur principally prior to these ages. If there is significant undercounting prior to these ages, we should also expect to see a jump in mortality rates at these ages.

[16] Many widows and widowers can draw benefits at 60. The Social Security Early Retirement Age is 62. The Normal Retirement Age is 65 or 66, depending on the year.

Year	All (20+)	55+	55-59	60-64	55-64	65+
1995	96.6	96.9	96.0	96.5	96.3	97.0
1996	96.8	97.0	96.1	97.1	96.6	97.0
1997	97.0	97.2	96.8	97.1	96.9	97.2
1998	97.1	97.3	96.6	97.2	96.9	97.4
1999	97.5	97.7	97.5	97.9	97.7	97.7
2000	97.7	98.0	97.9	98.2	98.1	97.9
2001	97.9	98.2	98.7	98.8	98.8	98.1
2002	98.1	98.4	98.7	99.4	99.1	98.3
2003	98.2	98.4	98.8	99.5	98.1	98.3
2004	98.6	99.0	98.9	99.6	99.3	99.0
2005	98.8	99.2	98.8	99.6	99.2	99.2
2006	98.8	99.3	98.6	99.6	99.1	99.3
2007	99.1	99.6	98.8	99.7	99.3	99.6
2008	99.4	99.8	99.3	99.6	99.5	99.8
2009	99.4	99.8	98.8	99.7	99.3	99.9
2010	99.7	100.0	99.3	100.0	99.7	100.1
2011	99.7	100.0	99.1	99.9	99.5	100.1
2012	99.7	100.1	98.7	99.9	99.4	100.2
2013	99.7	100.0	98.5	99.5	99.0	100.2
2014	98.6	99.1	96.5	97.7	97.2	99.4
Average	98.4	98.8	98.1	98.8	98.5	98.8

Notes: Estimated ratio of deaths in the SSA Numident data to adjusted National Death Index deaths over 1995-2014, by age group. Total (20+) column excludes children (age 0-19). The estimated ratio is calculated as $100 \times D_{kt}/O_{kt}$ where D_{kt} represents the number of deaths reported in the SSA data for age group k occurring in year t and O_{kt} represents the official number of deaths of US residents reported in the NDI for age group k during year t.

Table 3A.2. Estimated Percentage of US Deaths Included in the SSA Death Data, 1990-2014, by Age Group

In Figure 3A.1 we plot mortality rates at different ages, separately for those allowed and denied by an ALJ, by age at application, for our main estimation sample (age 55-64 at application). The figure shows mortality for up to 10 years after assignment and includes data for the year of assignment. The left panel shows mortality of those allowed by an ALJ, whereas the right panel shows mortality of those denied.

Mortality rates of those allowed rise from approximately 1.3% at age 56 to 3.0% by age 74. The lines are not perfectly smooth, but this is mostly due to sampling variability. There is no noticeable jump in mortality rates after any particular age. Among those denied, mortality rates are slightly lower than for those allowed. However, unlike the allowed, the denied appear to have spikes in mortality in the year of application.

The size of the spikes in mortality rates for the denied rises progressively for those who apply

	Dependent Variable: Allowed		Dependent Variable: Judge Leniency	
	Coefficient	t-stat	Coefficient	t-stat
Covariate	(1)	(2)	(3)	(4)
Sex				
Female	0.0175	16.1	0.0008	2.2
Age				
55 to 59	-0.1073	-53.5	-0.0109	-1.5
Race				
Black	-0.0582	-26.5	-0.0025	-1.0
Other (non-black, non-white) or unknown	-0.0085	-4.1	-0.0017	-0.9
Labor force participation and income				
Average participation rate, years -11 to -2	0.0043	6.5	0.0005	0.8
Average earnings/billion, years -11 to -2 ($2006)	0.0012	19.7	0.0000	1.2
Represented by lawyer				
Represented by lawyer	0.0479	5.5	-0.0075	-1.5
Application type				
SSDI	0.0289	20.1	0.0016	0.6
Education				
High school graduate, no college	-0.0044	-4.5	-0.0008	-1.0
Some college	-0.0154	-10.7	-0.0025	-1.6
College graduate	-0.0032	-1.7	-0.0022	-1.6
Health conditions (by diagnosis group)				
Neoplasms (e.g., cancer)	0.0436	17.2	0.0018	0.9
Mental disorders	-0.0207	-9.5	-0.0012	-1.2
Mental retardation	0.0007	0.2	0.0018	0.8
Nervous system	0.0009	0.5	-0.0008	-0.7
Circulatory system (e.g., heart disease)	0.0235	14.9	0.0024	1.1
Musculoskeletal disorders (e.g., back pain)	-0.0036	-2.3	0.0003	0.4
Respiratory system	-0.0281	-13.8	-0.0011	-1.4
Injuries	-0.0090	-4.4	-0.0007	-0.7
Endocrine system (e.g., diabetes)	0.0182	10.1	0.0008	0.7
Standard deviation of dependent variable	0.4116		0.1058	
R^2	0.0361		0.0040	
Number of Applicants = 2,759,907		Number of Judges =1,436		

Table 3A.3. Predictors of Allowance and Judge Leniency, All Ages

Notes: Column (1) is from a regression of de-meaned allowance on all the covariates listed. Column (3) is from a regression of judge leniency on all the covariates listed. Omitted category is male, 55-64s, white, not represented by a lawyer, applying for SSI or SSI and DI concurrently, not a high school graduate, with a health condition other than those listed above. The sample includes all applicants aged 25 to 64, and we exclude applicants who died the year of application. Standard errors clustered by judge.

	Obs.	Allowance Rate at ALJ Stage	Coefficient on Judge Leniency	Std. Error	T-Ratio	Relative Likelihood*
	(1)	(2)	(3)	(4)	(5)	(6)
All groups						
All groups	2,759,907	0.708	0.966	(0.019)	52	1.000
Sex						
Male	1,322,817	0.704	0.947	(0.023)	41	0.980
Female	1,437,090	0.711	0.984	(0.015)	66	1.019
Age						
25 to 54	2,149,676	0.670	1.022	(0.020)	51	1.058
55 to 64	610,231	0.841	0.705	(0.012)	60	0.730
Race						
White	1,738,652	0.737	0.939	(0.011)	84	0.972
Black	546,125	0.637	1.007	(0.039)	26	1.042
Other or unknown	475,130	0.682	1.001	(0.021)	47	1.036
Income						
Average earnings $<$ \$10000	1,587,843	0.644	1.035	(0.032)	33	1.071
Average earnings $\geq$ \$10000	1,172,064	0.794	0.839	(0.007)	121	0.868
Represented by lawyer						
Represented by lawyer	1,802,345	0.732	0.946	(0.007)	130	0.979
Not represented by lawyer	957,562	0.663	1.023	(0.051)	20	1.059
Application type						
SSDI	1,144,427	0.774	0.869	(0.008)	103	0.899
SSI or Concurrent (both SSDI and SSI)	1,615,480	0.661	1.023	(0.026)	39	1.059
Education						
Less than high school	918,011	0.693	0.981	(0.025)	39	1.015
High school graduate, no college	1,287,621	0.712	0.964	(0.017)	57	0.997
Some college	399,954	0.708	0.978	(0.015)	63	1.012
College graduate	154,321	0.763	0.865	(0.014)	61	0.896
Health conditions (by diagnosis group)						
Neoplasms (e.g., cancer)	55,935	0.791	0.763	(0.023)	34	0.790
Mental disorders	506,499	0.660	1.072	(0.023)	47	1.109
Mental retardation	32,893	0.645	1.022	(0.049)	21	1.058
Nervous system	172,606	0.708	0.922	(0.022)	42	0.954
Circulatory system (e.g., heart disease)	267,349	0.765	0.853	(0.018)	47	0.883
Musculoskeletal disorders	1,008,542	0.722	0.966	(0.013)	73	1.000
Respiratory system	96,781	0.686	0.967	(0.035)	27	1.001
Injuries	159,977	0.687	0.969	(0.026)	37	1.003
Endocrine system (e.g., diabetes)	144,969	0.723	0.913	(0.024)	39	0.945
All other	314,356	0.694	0.946	(0.028)	34	0.979

Notes: Column (3) displays the first stage estimate of the coefficient λ from the regression of de-meaned allowance rates on judge leniency for the full sample. Average earnings is calculated on income between 11 and 2 years before application. Standard errors clustered by judge.
*Relative likelihood is the ratio of the group specific coefficient on judge leniency (presented in column 3) to the full sample coefficient.

Table 3A.4. First Stage Estimates: Regression of Allowance Rates on Judge Leniency, By Demographics, All Ages

		Panel (a): 5 Year Mortality (Percent)						Panel (b): 10 Year Mortality (Percent)					
		Mortality Rates		OLS		IV		Mortality Rates		OLS		IV	
	Obs	Allowed	Denied	Diff.	SE	Diff.	SE	Allowed	Denied	Diff.	SE	Diff.	SE
All groups	610,231	9.71	8.35	1.94	(0.12)	2.30	(0.50)	21.95	19.99	2.77	(0.18)	2.81	(0.91)
Age Band													
55 to 59	390,600	9.13	7.71	1.96	(0.14)	2.56	(0.60)	20.60	18.88	2.64	(0.21)	2.79	(1.03)
60 to 64	219,631	10.72	9.60	1.94	(0.20)	1.91	(0.85)	24.32	22.13	3.11	(0.29)	3.03	(1.28)
Represented by lawyer													
Represented by lawyer	385,118	9.18	7.99	1.76	(0.15)	2.16	(0.78)	21.19	19.34	2.75	(0.22)	3.13	(1.20)
Not represented by lawyer	225,113	10.65	8.85	2.22	(0.20)	2.78	(0.64)	23.31	20.89	2.86	(0.29)	2.78	(1.10)
Application type													
SSDI	352,991	8.49	8.04	1.73	(0.15)	2.40	(0.67)	19.67	18.94	2.73	(0.23)	3.61	(1.03)
SSI or SSI/SSDI concurrent	257,240	11.44	8.69	2.57	(0.19)	2.45	(0.76)	25.22	21.15	3.54	(0.27)	2.37	(1.38)

Notes: This table displays the estimated effect of DI recipiency on mortality (with covariates) for those aged 55-64. Mortality Rates do not adjust for covariates. OLS and IV estimates are demeaned and include covariates. Covariates are those in Table 3.2; they include race, sex, age and education groups, health (disability category), average earnings and participation prior to disability, representation by an attorney, and an indicator of concurrent SSDI application. IV estimates use demeaned variables and judge leniency as the instrument. Standard errors clustered by judge.

Table 3A.5. Estimated Effect of DI Recipiency on Mortality, Disaggregated cont'd, Aged 55-64

		Panel (a): 5 Year Mortality (Percent)						Panel (b): 10 Year Mortality (Percent)					
		Mortality Rates		OLS		IV		Mortality Rates		OLS		IV	
	Obs	Allowed	Denied	Diff.	SE	Diff.	SE	Allowed	Denied	Diff.	SE	Diff.	SE
All groups	2,759,907	7.47	5.11	2.21	(0.05)	1.64	(0.58)	16.29	12.24	3.55	(0.10)	2.96	(1.05)
Sex													
Male	1,322,817	9.65	6.94	2.53	(0.07)	1.85	(0.60)	20.35	15.97	3.75	(0.12)	3.25	(1.04)
Female	1,437,090	5.48	3.38	1.92	(0.06)	1.43	(0.55)	12.59	8.72	3.34	(0.10)	2.66	(1.03)
Race													
White	1,738,652	7.35	5.45	1.82	(0.06)	1.56	(0.52)	16.29	12.81	3.15	(0.11)	3.04	(1.02)
Black	546,125	8.65	5.30	3.00	(0.10)	1.70	(0.91)	17.81	12.44	4.46	(0.16)	3.00	(1.51)
Other	475,130	6.67	3.83	2.55	(0.08)	1.85	(0.62)	14.65	10.27	3.78	(0.13)	2.66	(0.97)
Education Group													
Less than high school	918,011	7.81	5.23	2.23	(0.07)	1.55	(0.50)	17.40	12.64	3.74	(0.13)	2.84	(1.05)
High school graduate	1,287,621	7.31	5.06	2.21	(0.07)	1.57	(0.65)	15.88	12.03	3.51	(0.11)	2.91	(1.11)
Some college	399,954	7.35	4.97	2.33	(0.09)	1.61	(0.68)	15.60	11.97	3.56	(0.15)	3.30	(1.22)
College graduate	154,321	7.16	5.19	1.98	(0.16)	3.06	(1.14)	15.08	12.18	2.83	(0.23)	3.38	(1.39)
Income													
Average earnings $<$ \$10000	1,587,843	7.92	5.33	2.30	(0.06)	1.58	(0.54)	17.03	12.70	3.65	(0.11)	3.02	(0.96)
Average earnings $\geq$ \$10000	1,172,064	6.97	4.59	2.04	(0.07)	1.76	(0.69)	15.47	11.17	3.34	(0.11)	2.87	(1.27)
Health conditions													
Neoplasms	55,935	23.74	24.47	-0.16	(0.65)	2.71	(6.73)	36.91	40.75	-2.42	(0.82)	2.94	(7.62)
Respiratory system	96,781	11.43	6.57	3.77	(0.21)	1.00	(1.15)	25.57	17.59	5.73	(0.32)	3.04	(2.19)
Endocrine system	144,969	12.53	7.06	5.03	(0.18)	2.96	(1.34)	27.72	17.93	8.76	(0.26)	4.72	(2.58)
Circulatory system	267,349	11.57	7.99	3.79	(0.15)	3.36	(1.02)	26.00	19.63	6.18	(0.25)	5.28	(1.90)
Mental retardation	32,893	5.60	3.51	1.76	(0.25)	1.02	(0.88)	12.29	8.87	2.73	(0.38)	0.56	(1.53)
Nervous system	172,606	6.58	5.15	1.93	(0.15)	0.94	(0.88)	14.85	12.27	3.17	(0.21)	1.12	(1.38)
Mental disorders	506,499	5.61	4.15	1.52	(0.07)	1.61	(0.29)	12.47	10.25	2.27	(0.11)	2.54	(0.42)
Injuries	159,977	5.37	3.77	1.65	(0.05)	1.12	(0.21)	12.15	8.93	2.92	(0.08)	2.41	(0.46)
Musculoskeletal disorders	1,008,542	4.59	3.10	1.55	(0.12)	1.11	(0.50)	11.12	8.25	2.92	(0.18)	3.04	(0.80)
All other	314,356	10.89	8.07	3.31	(0.18)	2.27	(1.70)	21.03	17.22	4.57	(0.28)	4.30	(2.45)

Notes: This table displays the estimated effect of DI recipiency on mortality (with covariates) for the full sample (all ages). Mortality Rates do not adjust for covariates. OLS and IV estimates are demeaned and include covariates. Covariates are those in Table 3.2; they include race, sex, age and education groups, health (disability category), average earnings and participation prior to disability, representation by an attorney, and an indicator of concurrent SSDI application. IV estimates use demeaned variables and judge leniency as the instrument. Standard errors clustered by judge.

Table 3A.6. Estimated Effect of DI Recipiency on Mortality, Disaggregated, All Ages

		Panel (a): 5 Year Mortality (Percent)						Panel (b): 10 Year Mortality (Percent)					
		Mortality Rates		OLS		IV		Mortality Rates		OLS		IV	
	Obs	Allowed	Denied	Diff.	SE	Diff.	SE	Allowed	Denied	Diff.	SE	Diff.	SE
All groups	2,759,907	7.47	5.11	2.21	(0.05)	1.64	(0.58)	16.29	12.24	3.55	(0.10)	2.96	(1.05)
Age Band													
25 to 44	2,149,676	6.67	4.67	2.25	(0.06)	1.54	(0.54)	14.27	11.18	3.66	(0.10)	3.05	(0.88)
45 to 64	610,231	9.71	8.35	2.21	(0.12)	2.55	(0.64)	21.95	19.99	3.48	(0.23)	3.70	(1.56)
Represented by lawyer													
Represented by lawyer	1,802,345	6.97	4.99	1.90	(0.05)	1.40	(0.45)	15.51	12.04	3.16	(0.08)	2.55	(0.78)
Not represented by lawyer	957,562	8.50	5.29	2.75	(0.09)	2.05	(0.75)	17.89	12.55	4.25	(0.17)	3.68	(1.44)
Application type													
SSDI	1,144,427	6.29	4.72	1.55	(0.07)	1.61	(0.67)	14.25	11.26	2.66	(0.12)	2.96	(1.21)
SSI or SSI/SSDI concurrent	1,615,480	8.45	5.29	2.61	(0.06)	1.67	(0.57)	17.98	12.70	4.08	(0.11)	2.99	(1.04)

Notes: This table displays the estimated effect of DI recipiency on mortality (with covariates) for the full sample (all ages). Mortality Rates do not adjust for covariates. OLS and IV estimates are demeaned and include covariates. Covariates are those in Table 3.2; they include race, sex, age and education groups, health (disability category), average earnings and participation prior to disability, representation by an attorney, and an indicator of concurrent SSDI application. IV estimates use demeaned variables and judge leniency as the instrument. Standard errors clustered by judge.

Table 3A.7. Estimated Effect of DI Recipiency on Mortality, Disaggregated cont'd, All Ages

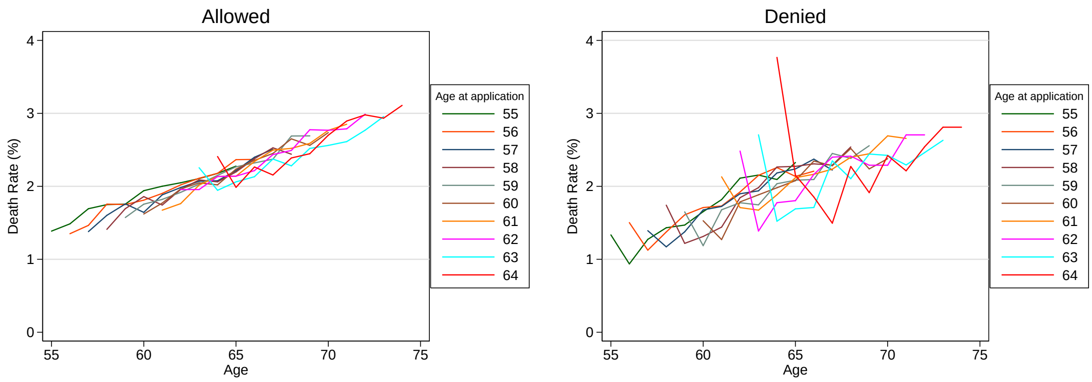


Figure 3A.1. Mortality of Those Allowed and Denied, by Age at Application

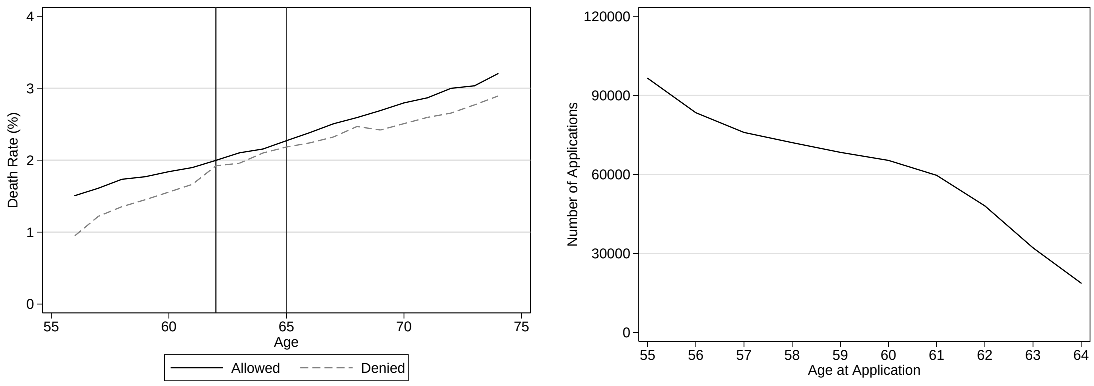


Figure 3A.2. Mortality of Those Allowed and Denied between ages 61 through 64, and there is an increase for those allowed also at ages 63 and 64. This high first year mortality rate is potentially due to the sample becoming increasingly selected towards individuals with high near term mortality. Note that the number of applicants drops sharply after age 60, as shown in the right panel of Figure 3A.2. The applicants who apply shortly before the regular retirement age of 65, and at or after the early retirement age of 62, are self-selected, potentially in different ways than those who apply at earlier ages.[17]

The high first year mortality among that we observe for the denied (but less so for the allowed) is potentially due to mismeasurement. Although we drop those whose death was recorded before an ALJ decision was made, there is the possibility that individuals who die prior to having their case heard might errantly be defined as being denied. It is for this reason that we drop sample members who die in the year of assignment in all of our main results. However, in Section 3.8 we show that whether or not we include these individuals has virtually no effect on our estimates.

Following the approach of our main analysis, in the left panel of Figure 3A.2 we drop individuals who die within the year of assignment. The left panel of Figure 3A.2 pools the mortality data from Figure 3A.1 by age to present a clearer pattern. While there is no jump in mortality rates at ages 60 or 65, it shows a small jump in mortality at age 62 for denied individuals, which is consistent with potential underreporting of mortality amongst these people before 62. We provide evidence on the size of the potential underreport below, giving us an alternative estimate of the underreporting correction p. However, we should also point out that Fitzpatrick and Moore (2016), using data from the National Center for Health Statistics Multiple Cause of Death (MCOD) files, document a two percent increase in male mortality immediately after age 62, and argue that the jump in mortality is caused by the fall in labor supply at this age. This could explain the jump that we observe, which is also consistent with the findings in this paper.

Here we make the assumption that the jump in mortality at age 62 is caused entirely by measurement error, and estimate the potential size of any undercount based on that jump. Let m^* denote the observed mortality rate, which is a function of age, and can be seen in the left panel of Figure 3A.2. Let $f(age) = \sum_{k=0}^{K} \gamma_k age^k$ denote the true age-specific mortality rate for the denied and as before p is the underreporting rate for those who are not claiming either disability or Social Security benefits. Let b_{age} denote the percentage of the population at that age that are claiming social security benefits.[18] Then

$$m^* = p(1 - b_{age})f(age) + b_{age}f(age) + \epsilon_{it}$$

We estimate p using nonlinear least-squares estimation using different functional forms for $f(age)$.

[17] One reason for applying, even though one would receive benefits for a short period of time, is if applicants have expensive high-mortality conditions, and hope to receive Medicaid (available to SSI applicants). The conditions that prompt application could also lead to high first-year mortality, relative to those who apply when younger.

[18] Note that before age 62, b_{age} is small, but jumps to 50% at 62, and is close to 95% by age 66. We are grateful to Timothy Moore for providing this data for each age.

The estimates are sensitive to the ages used and also the order of the polynomial K. If we use ages 59-68 and $K = 2$, we estimate $\hat{p} = .90$. If instead we use ages 56-74 and $K = 2$, then we estimate $\hat{p} = 1.02$. This approach provides often lower estimates of p than in our approach in the main text. However, this provides additional evidence that p is in the range of 0.9 to 1.0. Table 3.7 in Section 3.8 presents estimates which accounts for p being in this range. The key results of this paper do not markedly change when using these values of p.

3A.2.3 Marginal Treatment Effect Estimates The estimates displayed in Figure 3A.3 are calculated the same as the MTE estimates displayed in Figure 3.4, with the only difference being that Figure 3A.3 does not controls for covariates. When we control for covariates the slopes of the graphs change. For the sample aged 55-64 the slopes become more steep, and for those aged 25-64 the slope becomes less steep. The results for the aged 55-64s remain similar.

Figure 3A.4 presents MTE graphs using local polynomial smoothed estimates. While the first stage is calculated in the same way as in the MTE figures displayed in the main text (where allowance is regressed on a cubic of the instrument), the second stage is calculated using the local linear smoother from Maestas et al. (2013). Specifically, estimate a local quadratic regression of mortality (de-meaned by hearing offfice and day) and compute numerical derivatives to estimate the MTE ($\partial E[\tilde{y}]/\partial \hat{\hat{A}}$). Despite the difference in methodology, the estimated MTEs presented in Figure 3A.4 are similar to the ones presented in Figures 3A.3 and 3.4.

3A.3 Derivations

3A.3.1 Marginal Treatment Effects All derivations in this appendix are purely for completeness – they are straightforward adaptations of the results in Heckman et al. (2006) and French and Taber (2011). Define A_i as a 0-1 indicator =1 if individual i is allowed benefits, y_i is mortality. Other variables are defined in text. The outcome variable for individual i is:

$$y_i = \begin{cases} y_{1i} & \text{if} \qquad A_i = 1 \\ y_{0i} & \text{if} \qquad A_i = 0 \end{cases} \tag{3.10}$$

where

$$\begin{aligned} y_{1i} &= \phi + X_i\delta_y + u_{1i} \\ y_{0i} &= X_i\delta_y + u_i \end{aligned} \tag{3.11}$$

Combining equations (3.10) and (3.11) yields:

$$y_i = A_i\phi_i + X_i\delta_y + u_i. \tag{3.12}$$

where $\phi_i = \phi + u_{1i} - u_i$. Allowance is determined by

$$A_i = 1\{g(Q_i) - V_i > 0\} \tag{3.13}$$

where $1\{g(Q_i) - V_i > 0\}$ is the indicator function and is equal to 1 when $g(Q_i) - V_i > 0$, $g(Q_i)$ is an arbitrary function of $Q_i = (X_i, Z_i)$, where Z_i is our judge leniency measure described in the text, and V_i can be interpreted as a measure of the health of individual i. The variables u_i and ϕ_i are potentially correlated with each other but by assumption V_i is independent of Z_i and X_i. The Marginal Treatment Effect is

$$MTE(X_i = x, V_i = a) \equiv E[y_{1i} - y_{0i}|X_i = x, V_i = a] \tag{3.14}$$

where $P(Q_i) \equiv \Pr(A_i = 1|Q_i)$. Given equation (3.11), $MTE(X_i = x, V_i = a) = \phi + u_{1i} - u_{0i} = \phi_i$. Using equation (3.12), we estimate the conditional expectation function

$$\begin{aligned} E[y_i|X_i = x, P(Z_i) = a] &= E[A_i\phi_i + X_i\delta_y + u_i|X_i = x, P(Q_i) = a] \\ &= E[A_i(\phi + u_{1i} - u_i)|X_i = x, P(Q_i) = a] + X_i\delta_y + E[u_i|X_i = x, P(Q_i) = a] \\ &= E[A_i\phi|X_i = x, P(Q_i) = a] + E[(u_{1i} - u_i)|A_i = 1, X_i = x, P(Q_i) = a]a + X_i\delta_A \\ &\quad + E[u_i|X_i = x, P(Q_i) = a] \end{aligned} \tag{3.15}$$

where the step $E[A_i(u_{1i} - u_i)|X_i = x, P(Q_i) = a] = E[(u_{1i} - u_i)|A_i = 1, X_i = x, P(Q_i) = a]\Pr[A_i = 1|X_i = x, P(Q_i) = a]$ follows from the Law of Total Probability, and noting that $\Pr[A_i = 1|X_i = x, P(Q_i) = a] = a$. Continuing with the simplifications, and noting that we have already assumed that u_{1i}, u_i are independent of X_i we have:

$$\begin{aligned} E[y_i|X_i = x, P(Q_i) = a] &= \phi a + E[(u_{1i} - u_i)|A_i = 1, P(Q_i) = a] + X_i\delta_A + E[u_i|P(Q_i) = a] && (3.16) \\ &= X_i\delta_A + \phi a + E[(u_{1i} - u_i)|A_i = 1, P(Q_i) = a]a + E[u_i|P(Q_i) = a] \\ &= X_i\delta_A + K(a) && (3.17) \end{aligned}$$

where $K(a) \equiv \phi a + E[(u_{1i} - u_i)|A_i = 1, P(Q_i) = a]a + E[u_i|P(Q_i) = a]$. Differentiating equation (3.16) with respect to a yields

$$\frac{\partial E[y_i|X_i = x, P(Q_i) = a]}{\partial a} = K'(a) \tag{3.18}$$

This derivative is equal to the Marginal Treatment Effect. To see this, note that as a normalization we can let the distribution of V_i be uniform $[0,1]$, so

$$\begin{aligned}\frac{\partial E[y_i|X_i=x,P(Q_i)=a]}{\partial a} &= \frac{\partial\left[\int_0^a E[y_{1i}|X_i=x,V_i]dV_i+\int_a^1 E[y_{0i}|X_i=x,V_i]dV_i\right]}{\partial a}\\ &= E[y_{1i}|X_i=x,V_i=a]-E[y_{0i}|X_i=x,V_i=a]\\ &\equiv MTE(X_i=x,V_i=a).\end{aligned} \tag{3.19}$$

Thus estimation of equation (3.16) and taking $K'(a)$ yields the MTE. In the text we refer to $P(Q_i)$ as the plim of $\widehat{A}_i$.

3A.3.2 De-Meaning the Data and Doyle's Instrument In our estimation procedure, we have just under 200,000 hearing office-day interactions as covariates, so directly estimating equations (3.1) and (3.2) is not computationally feasible. To simplify the problem we de-mean variables by hearing office and day, and construct variables $\widetilde{A}_i = A_i - \bar{A}_i$, $\widetilde{y}_{i\tau} = y_{i\tau} - \bar{y_{i\tau}}$ where $\bar{A}_i$ and $\bar{y_{i\tau}}$ are the mean values of $A_i, y_{i\tau}$ for the hearing office-day on which case i was assignment.

Our instrument, from equation (3.3) of the text, which we rewrite below, is:

$$Z_i = \frac{1}{N_j-1}\sum_{s\in\{J\},s\neq i}\left(A_s\right) \tag{3.20}$$

which we then de-mean by hearing office and day, constructing $\widetilde{Z}_i$

As an alternative to this instrument, we also use Doyle Jr (2007) estimation procedure, also used in French and Song (2014), described below. This instrumental variable (which we term the judge allowance differential), is:

$$\widetilde{Z}_i^2 = \frac{1}{N_j-1}\sum_{s\in\{J\},s\neq i}\left(A_s-\overline{A_s}\right) \tag{3.21}$$

where $\overline{A_s}$ is the mean allowance rate by ALJs at case s's hearing office on the day case s was heard. This instrument is equivalent to the predicted allowance rate from OLS estimation of equation (3.1) where A_{i1} (the ALJ decision) is the dependent variable, controlling for a full set of hearing office$\times$ time interactions, and leaving observation i out, as in a jackknife estimator.

The instrument is $j_i\hat{\gamma}_1$ from the equation

$$A_i = j_i\hat{\gamma}_1 + X_i\delta_A + e_i \tag{3.22}$$

which implies

$$E[A_s|X_s] = E[j_s\hat{\gamma}|X_s] + X_s\delta_A \tag{3.23}$$

for any given s and so

$$E[j_s\hat{\gamma} - E[j_s\hat{\gamma}|X_s]] = E[A_s - E[A_{s1}|X_s]] \tag{3.24}$$

where the left-hand side object is $E[j_s\hat{\gamma} - E[j_s\hat{\gamma}|X_s]]$, the de-meaned instrumental variable. We approximate the right-hand side object, but using the sample analog and leaving observation i out, as in a jackknife estimator, so the constructed instrument is:

$$\widetilde{j_i}\hat{\gamma}_{-i} = \frac{1}{N_j - 1} \sum_{s\in\{J\}, s\neq i} (A_s - \overline{A_s}) \tag{3.25}$$

where N_j is the number of cases heard by judge j_i over the sample period, $\{J\}$ is the set of cases heard by judge j_i, $\overline{A_s}$ is the mean allowance rate by ALJs at case s's hearing office on the day case s was heard.

We then estimate equations (3.26) and (3.27):

$$\begin{aligned} \widetilde{A}_i &= \lambda(\widetilde{j_i}\hat{\gamma}_{-i}) + \eta_i, & (3.26)\\ \widetilde{y}_{it} &= \varphi_t(\widetilde{\widehat{\widetilde{A}_i}}) + \mu_{it} & (3.27) \end{aligned}$$

where "$\sim$" represents a de-meaned variable, e.g., $\widetilde{A}_{it} = A_{it} - \bar{A_{it}}$ and $\bar{A_{it}}$ is the mean allowance rate at the hearing office and on the day that case i was assigned and $\widetilde{j_i}\hat{\gamma}_{1,-i} = j_i\hat{\gamma}_{1,-i} - \overline{j_i\hat{\gamma}_{-i}}$ and $\overline{j_i\hat{\gamma}_{-i}}$ is the mean value of $j_i\hat{\gamma}_{1,-i}$ at the hearing office and on the day that case i was assigned. Because we remove case i from $\widetilde{j_i}\hat{\gamma}_{-i}$, as in a jackknife estimator, it should be independent of η_i and μ_{it}, even in a small sample. Based on Monte Carlo experiments with what seemed reasonable parameters, the procedure produced accurate approximations.

3A.3.3 Econometric Procedures to Address Missing Mortality Information In section 3.5.1 we showed that about 98.5% of deaths among those ages 55-64 are captured in the SSA mortality data and described some of the reasons for this discrepancy. Nevertheless, there is likely an under-count of those that die in our sample. Furthermore, this under-count is unlikely to be random. Because the SSA has a less of a financial incentive to measure deaths of non-DI/SSI recipients than non-recipients, the SSA data likely captures more than 98% of the deaths of those receiving benefits, but less than 98% of those not receiving benefits. This may make it look like non-beneficiaries are less likely to die than they are, and thus might lead us to infer that benefits do not reduce mortality when in fact they do. The larger the discrepancy between underreporting of beneficiaries relative to non-beneficiaries, the greater the potential bias in our estimates.

We assess how serious this problem is for our estimates. To construct the most extreme case, we assume that all deaths of beneficiaries are measured, but only a fraction p of non-beneficiaries' deaths that are measured. Define individual i's measured mortality at time τ as $y^*_{i\tau}$. Given the undercount of mortality amongst those denied, this will be

$$y^*_{i\tau} = \begin{cases} y_{i\tau} & \text{if} \quad A_i = 1 \\ y_{i\tau} & \text{with probability } p \text{ if} \quad A_i = 0 \\ 0 & \text{with probability 1-}p \text{ if} \quad A_i = 0 \end{cases} \tag{3.28}$$

where by assumption the probability p is independent of any of the variables that determine mortality. To address with problem we create the variable

$$\tilde{y}_{i\tau} = \begin{cases} y^*_{i\tau} & \text{if} \quad A_i = 1 \\ \frac{1}{p} y^*_{i\tau} & \text{if} \quad A_i = 0 \end{cases} \tag{3.29}$$

Writing our new variable this way, suppose that the true model is a modified version of equation (3.2)

$$y_{i\tau} = A_i \phi + X_i \delta_{y\tau} + u_{i\tau} \tag{3.30}$$

so that the coefficient on allowance is common to everyone. Using the adjusted mortality measure $\tilde{y}_{i\tau}$ the OLS estimate in equation identifies the conditional expectation of $\tilde{y}_{i\tau}$ given X_i

$$\begin{aligned} \mathbb{E}[\tilde{y}_{i\tau}|A_i, X_i] &= \mathbb{E}[\tilde{y}_{i\tau}|A_i = 1, X_i] \Pr(A = 1|X) + \mathbb{E}[\tilde{y}|A = 0, X] \Pr(A = 0|X) \\ &= \mathbb{E}[y_{i\tau}|A_i = 1, X_i] \Pr(A = 1|X) + \mathbb{E}[\frac{1}{p} y^*|A = 0, X] \Pr(A = 0|X) \\ &= \mathbb{E}[y_{i\tau}|A_i = 1, X_i] \Pr(A = 1|X) + \mathbb{E}[y|A = 0, X] \Pr(A = 0|X) \\ &= \mathbb{E}[y_{i\tau}|A_i, X_i] \end{aligned} \tag{3.31}$$

which is the conditional expectation of $y_{i\tau}$ given X_i, which is what OLS recovers.

Defining the instrument we use as Z_i, instrumental variables, using $\tilde{y}_{i\tau}$ as the left hand side variable, the conditional expectation of $\tilde{y}_{i\tau}$ given Z_i, X_i is

$$\begin{aligned} \mathbb{E}[\tilde{y}|Z, X] &= \mathbb{E}[\tilde{y}|A = 1, Z, X] \Pr(A = 1|Z, X) + \mathbb{E}[\tilde{y}|A = 0, Z, X] \Pr(A = 0|Z, X) \\ &= \mathbb{E}[\phi + X_i \delta_{y\tau} + u_{i\tau}|A = 1, Z, X] \Pr(A = 1|Z, X) \\ &+ \mathbb{E}[\frac{1}{p} p (X_i \delta_{y\tau} + u_{i\tau}) |A = 0, Z, X] \Pr(A = 0|Z, X) && (3.32) \\ &= (\phi + X_i \delta_{y\tau}) \Pr(A = 1|Z, X) + (X_i \delta_{y\tau}) \Pr(A = 0|Z, X) && (3.33) \\ &+ \mathbb{E}[u_{i\tau}|A = 1, Z, X] \Pr(A = 1|Z, X) + \mathbb{E}[u_{i\tau}|A = 0, Z, X] \Pr(A = 0|Z, X) \end{aligned}$$

Note that, by the Law of Iterated Expectations,

$\mathbb{E}[u_{i\tau}|Z,X] = \mathbb{E}[u_{i\tau}|A=1,Z,X]\Pr(A=1|Z,X) + \mathbb{E}[u_{i\tau}|A=0,Z,X]\Pr(A=0|Z,X)$ and also recall the usual IV assumption that $\mathbb{E}[u_{i\tau}|Z,X]=0$. Thus

$\mathbb{E}[\tilde{y}|Z,X] = (\phi + X_i\delta_{y\tau})\Pr(A=1|Z,X) + (\phi + X_i\delta_{y\tau})\Pr(A=0|Z,X) = (X_i\delta_{y\tau}) + \phi Pr[D=1|Z,X]$

and $\mathbb{E}[A|Z,X] = \Pr[A=1|Z,X]$. Likewise, the conditional expectation of $y_{i\tau}$ given Z_i can be derived using the same formula as in equation (3.32):

$$\mathbb{E}[y|Z,X] = (X_i\delta_{y\tau}) + \ \phi Pr[A=1|Z,X]$$

and so the IV estimator using $\tilde{y}$ as the left hand side variable should (asymptotically) yield the same values as the IV estimator using y as the left hand side variable:

$$\frac{\mathbb{E}[y|Z,X]}{\mathbb{E}[A|Z,X]} = \frac{\mathbb{E}[\tilde{y}|Z,X]}{\mathbb{E}[A|Z,X]} = \frac{(X_i\delta_{y\tau}) + \phi Pr[A=1|Z,X]}{Pr[A=1|Z,X]}$$

which is the standard formula for a IV estimator with binary endogenous variable.

Next, we describe how to measure p. Using the Law of Total Probability, the assumption $\Pr(y^*_{i\tau}=1|y_{i\tau}=1,A_i=1)$ and the definition $p \equiv (y^*_{i\tau}=1|y_{i\tau}=1,A_i=0)$ we get:

$$\begin{aligned}\Pr(y^*_{i\tau}=1|y_{i\tau}=1) &= \Pr(y^*_{i\tau}=1|y_{i\tau}=1,A_i=1)\Pr(A_i=1|y_{i\tau}=1)\\ &\quad +(y^*_{i\tau}=1|y_{i\tau}=1,A_i=0)\Pr(A_i=0|y_{i\tau}=1)\\ &= \Pr(A_i=1|y_{i\tau}=1) + p\Pr(A_i=0|y_{i\tau}=1) \end{aligned} \tag{3.34}$$

Using Bayes rule we know that:

$$\Pr(A_i=1|y_{i\tau}=1) = \frac{\Pr(y_{i\tau}=1|A_i=1)\Pr(A_i=1)}{\Pr(y_i=1)}, \tag{3.35}$$

$$\Pr(A_i=0|y_{i\tau}=1) = \frac{\Pr(y_{i\tau}=1|A_i=0)\Pr(A_i=0)}{\Pr(y_i=1)}, \tag{3.36}$$

Combining equations (3.34)- (3.36) yields

$$p = \frac{[\Pr(y^*_{i\tau}=1|y_{i\tau}=1)\Pr(y_i=1)] - [\Pr(y_{i\tau}=1|A_i=1)\Pr(A_i=1)]}{\Pr(y_{i\tau}=1|A_i=0)\Pr(A_i=0)}. \tag{3.37}$$

Using the Law of Total Probability and straightforward algebra shows that

$$\Pr(y_{i\tau}=1|A_i=0) = \frac{\Pr(y_{i\tau}=1) - \Pr(y_{i\tau}=1|A_i=1)\Pr(A_i=1)}{\Pr(A_i=0)} \tag{3.38}$$

Combining equations (3.37) and (3.38) yields:

$$p = \frac{[\Pr(y^*_{i\tau} = 1|y_{i\tau} = 1)\Pr(y_i = 1)] - [\Pr(y_{i\tau} = 1|A_i = 1)\Pr(A_i = 1)]}{[\Pr(y_i = 1)] - [\Pr(y_{i\tau} = 1|A_i = 1)\Pr(A_i = 1)]}. \tag{3.39}$$

Since, using the definition of a joint probability and the fact that anytime a death is observed in the SSA data are also observed in the National Death Index, $[\Pr(y^*_{i\tau} = 1|y_{i\tau} = 1)\Pr(y_i = 1)] = [\Pr(y^*_{i\tau} = 1, y_{i\tau} = 1)] = \Pr(y^*_{i\tau} = 1)$. Thus equation (3.39) can be rewritten as

$$p = \frac{[\Pr(y^*_{i\tau} = 1)] - [\Pr(y_{i\tau} = 1|A_i = 1)\Pr(A_i = 1)]}{[\Pr(y_i = 1)] - [\Pr(y_{i\tau} = 1|A_i = 1)\Pr(A_i = 1)]}. \tag{3.40}$$

Assuming that $\Pr\widehat{(y^*_{i\tau} = 1)}$ = #of deaths in the SSA data/population and $\Pr\widehat{(y_{i\tau} = 1)}$ = # of deaths in the NDI data/population, equation 3.40 can be estimated as

$$\begin{aligned} p &= \frac{\text{\#of deaths in the SSA data/population} - \text{\#of deaths of beneficiaries in SSA data/population}}{\text{\#of deaths in the NDI data/population} - \text{\#of deaths of beneficiaries in SSA data/population}} \\ &= \frac{\text{\#of deaths in the SSA data} - \text{\#of deaths of beneficiaries in SSA data}}{\text{\#of deaths in the NDI data} - \text{\#of deaths of beneficiaries in SSA data}}. \end{aligned} \tag{3.41}$$

We can estimate all the probabilities in equation (3.40). For those ages 55-64, $\Pr(y^*_{i\tau} = 1|y_{i\tau} = 1) = .98$ as we calculated previously, $\Pr(y_i = 1)$ is the annual mortality rate of all members in this age group, which we take from aggregate life tables, $\Pr(y_{i\tau} = 1|A_i = 1)$ we calculate from internal Social Security Administration documents. adjudication We calculate $\Pr(A_i = 1)$, the probability of receiving benefits, again using Social Security Administration data.

3A.4 Calculations of the Impact of ALJ Allowance on Subsequent Allowance, Income, and Benefits

In section 3.7 we present evidence on how receipt of DI benefits affects labor supply, earnings, health insurance, and the dollar value of those health care benefits. In this appendix we further document the calculations in that section.

3A.4.1 Allowance Many denied applicants continue to appeal and reapply for benefits until they are allowed. Figure 3.1 shows that 35% of all applicants denied by an ALJ were allowed benefits within three years. French and Song (2014) show both IV and OLS estimates of subsequent allowance rates, where the IV estimates use our judge leniency instrument. The difference between OLS and IV estimates are that the IV estimates measure the subsequent allowance rates for the marginal individual, whereas the OLS estimates measures subsequent allowance rates for the average. French and Song find that IV estimates are slightly higher than OLS. For example, the IV estimate of allowance is 42% three years after assignment, versus 35% from the OLS estimates. This finding is consistent with the view that those affected by the instrument are likely the marginal cases who have a better chance of final allowance than others denied benefits.

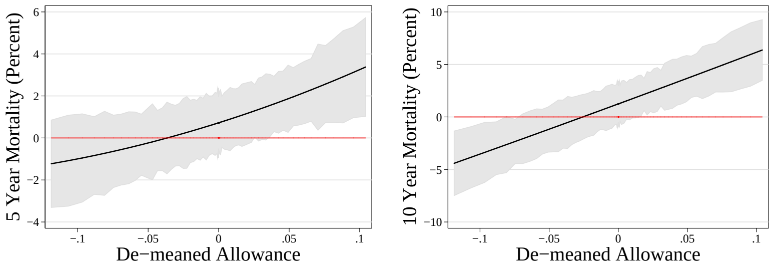
3A.4.2 Income Benefits If an applicant is allowed DI benefits, the dollar amount of benefits depends on previous labor earnings. Disabled worker benefits averaged $1,004 per month among DI beneficiaries in 2007 (U.S. Social Security Administration, 2008). Because the benefit schedule is progressive, disability benefits replace 60% and 40% of labor income for those at the 10th and 50th percentile of the earnings distribution, respectively (Autor and Duggan, 2006). Those receiving benefits can earn up to the Substantial Gainful Activity level (SGA), which was $500 per month (in current dollars) during the 1990s and $900 per month in 2007. Those earning more than this amount for more than a nine month Trial Work Period lose their benefits. Disabled individuals with especially weak earnings histories and low asset levels are eligible for a related program called Supplemental Security Income (SSI). SSI benefits are not a function of previous labor income. The Federal Maximum SSI benefit level was $386 per month in 1990 and $623 in 2007. Some states supplement this benefit. Benefits are reduced by 50 cents for every dollar of labor income. Many people draw both DI and SSI benefits concurrently. We take DI/SSI benefit calculations from French and Song (2014), which use the distribution of post-tax wages plus DI/SSI benefits for everyone in our data using the federal, state, and local tax schedule shown in French and Jones (2011). Detailed information on earnings histories and state of residence allow for accurate measurement of individual benefits. Our main limitation on these measurements is that ideally we should know family structure and all sources of income to calculate taxes. Unfortunately, we do not have this information, so we assume that the individual can claim no dependants for the DI/SSI.

3A.4.3 Health Insurance DI/SSI beneficiaries usually receive either Medicare or Medicaid health insurance. DI beneficiaries almost always receive Medicare benefits after a 2 year waiting period. For SSI beneficiaries, things are more complicated. If they meet certain requirements, SSI beneficiaries are immediately eligible for Medicaid. In certain states all SSI beneficiaries receive Medicaid benefits, whereas in other states the Medicaid eligibility criteria are more stringent.[19] Thus some SSI beneficiaries never get health insurance benefits. See Rupp and Riley (2012) for more information. We know whether the individual is applying for DI versus SSI benefits, and also state of residence. Thus we can exploit these variables and the estimates in Rupp and Riley (2012) in whether an individual has Medicaid and/or Medicare. They estimate the share of DI and SSI beneficiaries with Medicaid or Medicare benefits at different points in time.

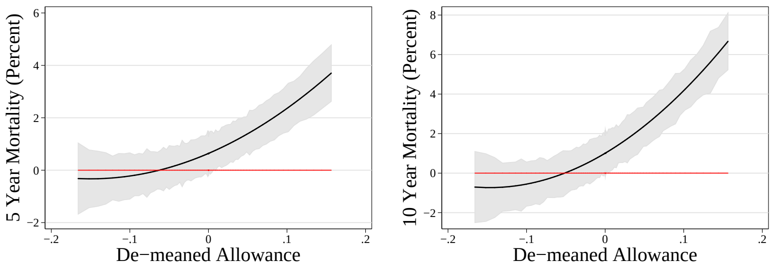
3A.4.4 Employment and Earnings Both income effects (through the high replacement rate) and substitution effects (beneficiaries will lose benefits if they earn above the SGA amount) causes DI recipients to reduce labor supply. Likewise, the income benefit and the clawback of benefits for SSI beneficiaries also causes SSI beneficiaries to reduce labor supply. Furthermore,

[19] In 32 states and DC, SSI beneficiaries are automatically eligible for Medicaid. In another seven states, SSI beneficiaries are eligible for Medicaid but must file a separate application. The remaining states have rules for Medicaid eligibility that differ from the eligibility rules for SSI.

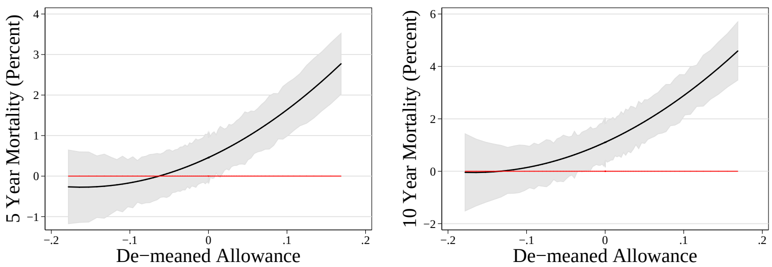
DI/SSI benefits likely reduce labor supply through a third channel – health insurance eligibility. Medicare and Medicaid largely eliminate the value of employer-provided health insurance. For those working at a firms providing health insurance, the health insurance from work is potentially a powerful incentive to stay at that job. The employment and earnings losses for our sample are reported in Table 3.6.



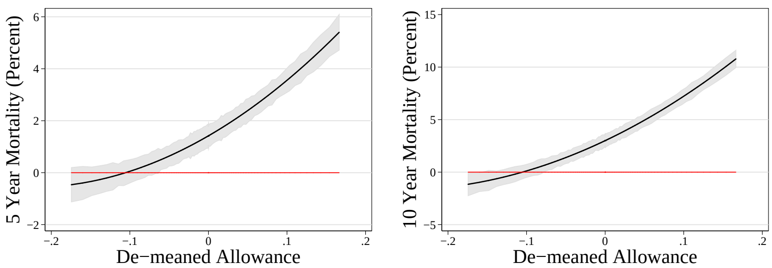
(a) Aged 55-64



(b) Aged 45-54



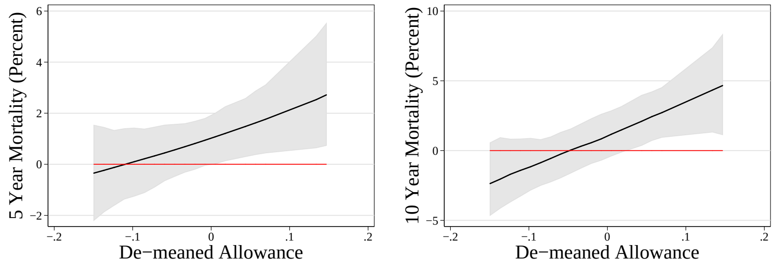
(c) Aged 25-44



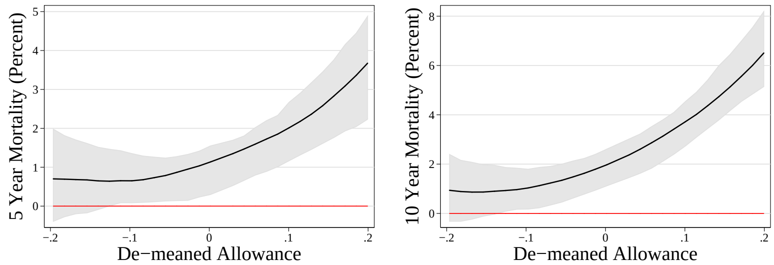
(d) All Ages (25-64)

Notes: This figure displays the estimated mortality response as a function of predicted de-meaned allowance. This is the same as Figure 3.4, but the estimates were calculated without covariates. Within each panel: the left figure displays 5 year mortality, and the right figure displays 10 year mortality. Mean allowance rate is 0.84 for those aged 55-64, 0.71 for those aged 45-54, 0.63 for those aged 25-44, and 0.71 for those aged 25-64.

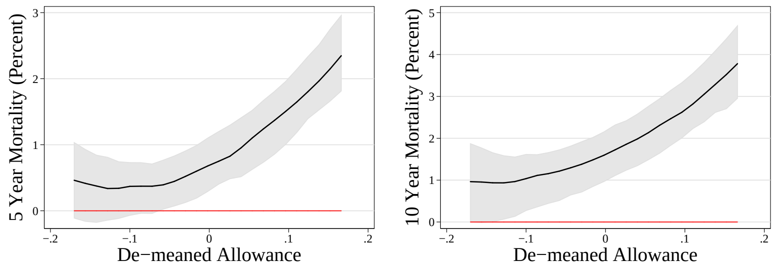
Figure 3A.3. Marginal Treatment Effects: Mortality Response by De-meaned Allowance, Without Covariates



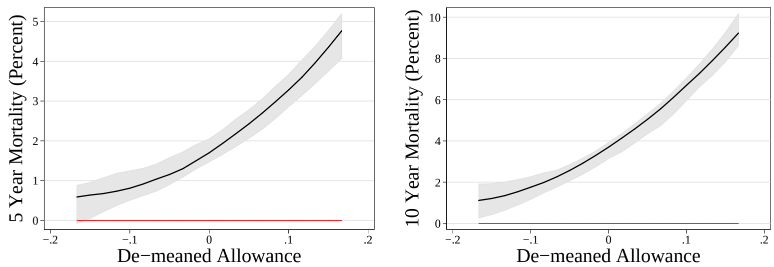
(a) Aged 55-64



(b) Aged 45-54



(c) Aged 25-44



(d) All Ages

Notes: This figure displays the estimated mortality response as a function of predicted de-meaned allowance. This is the same as Figure 3.4, but the estimates were calculated using local polynomial smoothed estimates without covariates. Within each panel: the left figure displays 5 year mortality, and the right figure displays 10 year mortality. Mean allowance rate is 0.84 for those aged 55-64, 0.71 for those aged 45-54, 0.63 for those aged 25-44, and 0.71 for those aged 25-64.

Figure 3A.4. Marginal Treatment Effects: Mortality Response by De-meaned Allowance, Using Local Polynomial Smoothed Estimates

Chapter 4

The Role of Information in Explaining the Lack of Welfare-Induced Migration

4.1 Introduction

There exists a substantial literature assessing whether geographic differences in the generosity of welfare programs induce internal migration for those seeking more generous benefits or welfare. However, there is little consensus on the size of any effect, with most studies finding no effect or a very modest one.[1] Given that sometimes the differences in welfare generosity between areas can be large this result is surprising.[2]

In the UK the majority of welfare policies are set at the national level (e.g. unemployment and housing benefits). However, personal social services provision (which includes social care) is decentralized to local authorities. The local authorities have a considerable degree of autonomy in the share of total revenues they assign to it, to the quality of services they provide, and to any user-fees they charge. This has created large variations in the spending and quality of social services across areas. While most of the social services do not take the form of direct money transfers, they still offer considerable benefits, particularly to the elderly. The benefits

[1] Brueckner (2000) provides an excellent overview of this literature.

[2] Kennan and Walker (2010) look at Aid to Families with Dependent Children programme and find that even large differences in benefit levels provide surprisingly weak migration incentives for young welfare-eligible women. Similarly, Schwartz and Sommers (2014) find no significant migration effects of low-income people following recent expansions of Medicaid in certain states.

include helping clients in their own homes (e.g. home help and meals), in nursing and residential establishments, and in day care facilities. Around £7 billion (2015 £) was spent by all local authorities on personal social services for the elderly in 2001, but the range in spending across local authorities showed a wide variation, with the highest spending around four times more (per capita) than the lowest.

It is not clear whether the elderly are aware of the differences in spending and quality of social services across areas or how their area compares to others, which could help to explain why even though welfare differences exist (in this case in social services provision), there does not appear to be much welfare-induced migration.

This paper seeks to establish the role of information as part of the explanation for the lack of welfare-based migration. I focus on the elderly, who are likely users of local social services, and in particular I focus on claimants of state pension, who I refer to throughout this paper as *pensioners*. I use a policy called the Social Services Performance Review (SSPR), which was introduced into England in 2002, where the national government gave a publicly-released rating of each local authority's personal social services on a scale from zero to three stars based on a series of accounting and performance measures. I treat this public release of the ratings as an "information shock" and analyze the number of pensioners within each area before and after the release based on their star rating. The hypothesis is that if information does play a role then we would expect that areas that receive a higher star rating would see an increase in the size of the pensioner population compared to the areas which perform poorly in the ratings, after this information is made public. Using a difference-in-differences approach I find that a one increase in the publicly-released star rating led to a 0.01 percentage points increase in the percentage of pensioners living in that area relative to other areas. This corresponds to a 1.3 % increase.

A potential problem with this estimate is that the counterfactual cannot be observed; the areas with the highest star rating may have seen an increase in pensioners even without the ratings being made public. This would be the case if pensioners already knew what the highest quality areas were or if the quality was correlated with some other underlying area attributes. To address this, I perform a series of robustness checks and also exploit the fact that the ratings were only conducted in England, whereas in Wales, where the structure of local government is the same, no ratings were produced.[3] To this end, I predict the ratings that local authorities in Wales would have received using detailed accounting and performance measures, which were collected for both England and Wales, and then use the Welsh local authorities as a control group when assessing the impact of the SSPR release on the treatment group, the English councils.

[3]Lockwood and Porcelli (2013) use Wales as a control group for England when assessing the impact of Comprehensive Performance Assessment (CPA). The CPA was similar to the ratings system to the SSPR however it rated all local government services, not just their social services, and offered incentives for councils to perform well. The social services star ratings judgements fed directly into the local government CPA. A council had to receive a good star rating for their social services in order to receive the highest CPA rating.

Using a difference-in-difference-in-differences approach I find very similar results to the original difference-in-differences estimates, which provides confidence in the results.

To confirm that the increase in the number of pensioners in areas that performed well in the ratings was caused by migration and not other reasons, I analyze migration data for the population aged over 60. I find, somewhat surprisingly, that the release of information does not appear to have had an effect on the in-migration rate of population (the number of people moving into an area as a percentage of the population), but I do find a significant effect for the out-migration rate (the number of people moving out of an area as a percentage of the population). A one increase in publicly-released star rating led to a -0.04 percentage point decrease in the out-migration rate and a 0.02 percentage point increase in the net migration rate of the area.

I use the results from my empirical investigation to motivate a search model with nested learning. I draw inspiration from the literature on equilibrium search unemployment, and in particular Moscarini (2005) and Papageorgiou (2014), but instead of workers searching for jobs, pensioners are searching for the area that provides the best social services. The economy is fully populated with pensioners, who receive utility from only two sources: a state pension (which is not area specific), and social services (which is area specific). Each pensioner has an individual-specific quality "match" for the social services in each area, however this is unobservable. Social services are experience goods, which are observed with noise by individuals. While the individual-specific match quality are unknown to pensioners, they have beliefs about the quality of every council and update their beliefs about the quality of their current council's services based on the observed services received. Each period, pensioners assess their beliefs and decide if and where to move. An "information shock" is then introduced into the model to replicate the introduction of the SSPR ratings by allowing, at the time of the shock, pensioners' beliefs to move closer to the true mean qualities of each council. I estimate the model using indirect inference, and target moments characterizing migration, the distribution of pensioners across councils of different qualities, and a difference-in-differences estimate from an information shock, mimicking the main empirical result in the paper.

Estimates suggest that there is a lot of noise in the learning process, and that it takes pensioners a long time to learn the true quality of their area. The model shows that to generate a migratory response of the same magnitude as that observed in the data, the information shock would have to bring pensioners' beliefs about the quality of each area closer to the true mean quality of those areas by 58%. The shock is met with an increase in net migration for the best councils and a decrease in net migration for the worst. This is in a large part caused by a reduction in the gross migration of pensioners living in the best councils. Overall, I find that the information release had a positive effect on average utility, worth about £240 per year, which persists into the future. I treat the SSPR as an impersonal information shock (as it only gave information on the mean quality of each area), and show that if the information shock instead conveyed personalized information the migration response is very different and can result in much greater utility gains.

This paper makes the following contributions. This is one of the first papers that considers how information releases can directly affect the migration decisions of individuals, and shows that information can offer a partial explanation for why there is a lack of internal welfare-migration. I exploit a natural experiment to find causal estimates of a response to an information shock, and show the extent to which this shock affected the migration decisions of pensioners. My estimates are shown to be robust by finding a control group. Using the empirical evidence as a motivation I develop a search model with nested learning, where pensioners search for the areas with the best social services and gradually learn about their unobserved quality. The model is very tractable and can provide insights about how beliefs, the speed of learning, and information shocks can affect migration. The model can be applied to a wide range of applications where individuals do not have perfect information about areas. This has consequences for the current literature on migration, which generally assumes individuals have perfect information about the quality of area-based amenities. If individuals instead have imperfect beliefs, this could affect the estimation of other parameters, such as moving costs. The findings in this paper are not only useful for other researchers but have important implications for policymakers regarding information releases, especially those relating to the quality of local services or benefits. I show that the form that the information release takes, and whether it is personalized or not, can lead to very different outcomes. This is increasingly relevant as central and local governments are becoming more transparent and information is becoming easier to access on the internet.

The paper unfolds as follows. Section 4.2 surveys related literature. Section 4.3 discusses the context, data and descriptive statistics of social services in England and the release of the SSPR in 2002. Section 4.4 presents and critically assesses the identification strategy. Section 4.5 presents the main empirical findings, discusses their economic significance, and reports results from a number of robustness checks. Section 4.6 introduces and estimates a search model with nested learning. The final section offers some concluding remarks.

4.2 Related Literature

4.2.1 Mobility Patterns of the Elderly

Graves and Knapp (1988) is one of the earliest studies to consider the mobility patterns of the elderly, which they model as a subcase of a more general migration model which interacts individual-specific traits (e.g. health and retirement status) and location-specific traits (e.g. amenities, rents, and wages). Their theory suggests that the spatially invariant incomes of the retired should lead to migration toward areas where the wage and rent compensation for amenities occurs primarily in the labor market, rather than in the land market. Empirical evidence appears to be consistent with their theoretical expectations. For example, Chen and Rosenthal (2008), using the 1970–2000 US Census, find that couples near retirement tend to move away from places with favorable business environments and towards places with highly valued consumer amenities.

The findings in this literature suggest that the elderly should factor the quality of local social services in their migration decisions. Indeed, some authors have found that the size and quality of the public sector is an important determinant of elderly migration.[4] However, none of these papers account for whether individuals are actually fully informed of the spatial differences in the quality of the areas.

4.2.2 Welfare Migration

The question of whether or not differences in social services (which can be viewed as a welfare benefit) may induce elderly to move, is linked to the idea of welfare migration. Most of the welfare migration literature explores the issue of welfare competition and "race to the bottom". The theory speculates that in the presence of welfare recipients' mobility, decentralized welfare is set strategically by each authority as they consider neighboring authorities' generosity levels before setting their own to avoid becoming a welfare magnet, leading to a race to the bottom in generosity of welfare. The majority of the existing empirical studies on welfare competition looks for strategic interaction among US states in setting welfare generosity. Brueckner (2000) provides a survey of these studies, most of which suggest that strategic interaction does occur. Looking at England, Revelli (2006), Moscone et al. (2007) and Fernandez and Forder (2015) all identify spatial interdependencies in social services expenditure levels between neighboring local authorities.

However, evidence on whether people are induced to move in order to obtain more generous welfare benefits is less clear.[5] Again, Brueckner (2000) provides a survey of the empirical evidence which is mixed, with some studies finding small effects, and most others finding an absence of welfare-induced migration.[6] One recent paper to look at this issue is Schwartz and Sommers (2014) who study low-income residents of states that forgo the Affordable Care Act's expansion of Medicaid, and who would be eligible if they moved to a state that did choose to expand coverage. They use a difference-in-differences analysis of migration in expansion and control states, and find no significant effects on migration. However, it is not clear whether or not the general public were aware about the welfare differences across areas or the size of any differences.

[4]Conway and Houtenville (1998; 2001) explore whether the elderly migrate to states with government policies that treat them favorably. Using state-level migration data from the 1990 Census, they estimate out migration and in migration equations that suggest that the public sector is an important determinant of elderly migration.

[5]As Brueckner points out, strategic interaction (and thus the race to the bottom) amongst states can occur even without the presence of welfare induced migration. All that is required is the *perception* by state governments that more generous benefits may attract welfare migrants.

[6]Meyer et al. (1998) also provides a good summary of the literature. Some studies do find evidence of welfare-induced migration. For example, Gelbach (2004), finds that among women likely to use welfare in the US, those who move tend to move to higher-benefit states. Similarly, McKinnish (2007), finds estimates that are consistent with the presence of welfare migration effects.

Kennan and Walker (2010) provide a systematic analysis of welfare-induced migration. They estimate a job search model based on a modified version of Kennan and Walker (2011), whereby young welfare-eligible women search across states for the optimal wage. The women know the wage in the current location, but to determine the wage at other locations it is necessary to move there. In each location, welfare acts as a fallback option, and the value of this is known by individuals. Performing counterfactual analysis on the model, it is found that that equalizing welfare benefits has a negligible effect on migration, regardless of whether the national benefit is set at either the lowest or the highest state benefit level.

4.2.3 Information Disclosures

This paper is also related to a broader literature on how people respond to information disclosures. In an experimental setting, Wiswall and Zafar (2015) provide information to college students regarding the population distribution of earnings of different college major choices. They find that college students are substantially misinformed about population earnings, and revise their self-earnings beliefs in response to being provided with information. The revisions are systematically related to the informativeness of the signal, with students exhibiting larger revisions when the information is more specific.

The majority of papers in the information disclosure literature consider the effect of information releases on house prices; looking at the effect of releasing information on school quality (Figlio and Lucas, 2004; Fiva and Kirkebøen, 2011; Carrillo et al., 2013; Imberman and Lovenheim, 2016; Haisken-DeNew et al., 2016) and even the effect of releasing information on earthquake vulnerability (Brookshire et al., 1985). To the extent that these changes in house prices are caused by shifting demand, most of these studies provide indirect evidence of the effect of information releases on migration. However, to my knowledge this is the first paper to directly consider what role information plays in migration decisions.

4.3 Context, Data and Descriptive Statistics

4.3.1 Local Governments in England

4.3.1.1 Structure Local authorities in England are split into two main types of councils, unitary and two-tier. Each two-tier council is made up of a single county council with a number of lower-level district councils. Unitary and county authorities are responsible for the provision of social services, primary and secondary education, transport and waste disposal. There are 148 unitary and county authorities that cover the whole of England and do not overlap in social

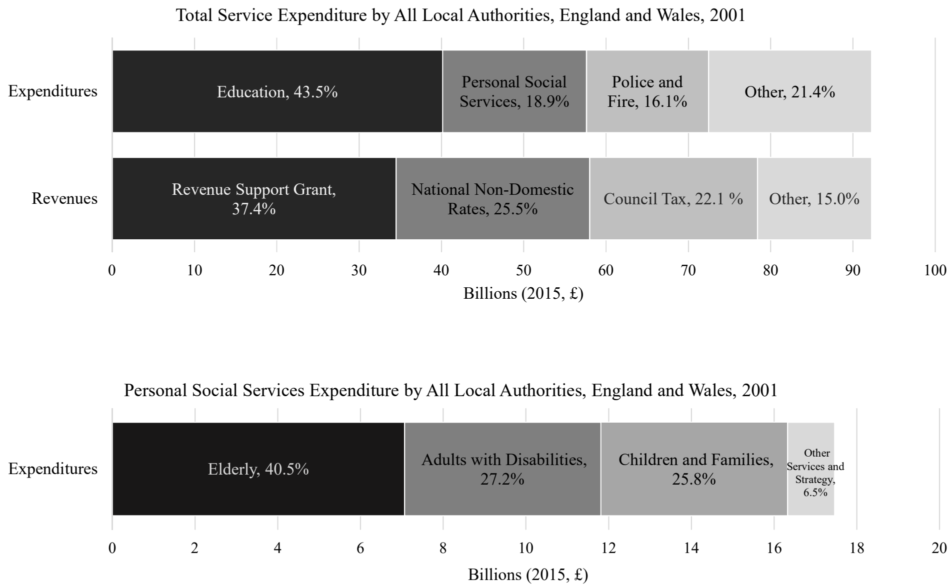


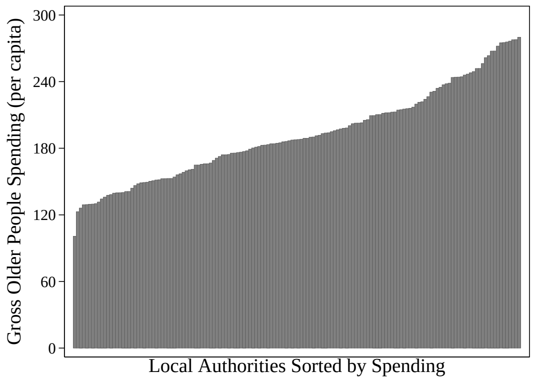
Figure 4.1. Revenues and Expenditures of All Local Authorities in England and Wales, 2001
Notes: The figure displays the total spending and revenue in 2001 of 148 local authorities in England and 22 local authorities in Wales. All amounts in 2015 £.
Source: Finance and General Statistics from the Chartered Institute of Public Finance and Accountancy (CIPFA).

service provision.[7] In addition, there are 22 Welsh local authorities which perform the same functions.

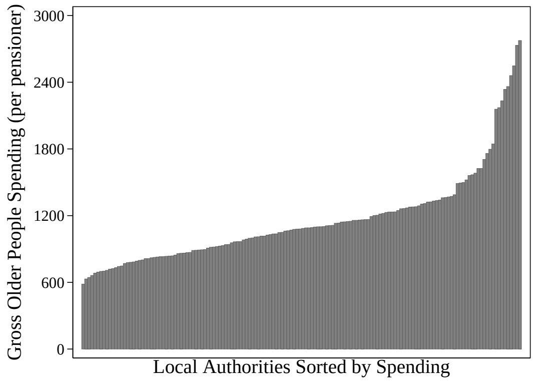
4.3.1.2 Revenues and Expenditure Figure 4.1 displays the revenues and expenditures of all local authorities in England and Wales in 2001. As can be seen in the figure, over one third of all revenues come from national government revenue support grants. Most of this grant revenue is not ring fenced, so it can be re-prioritized locally across services. The local authorities also use local taxes to complement the national government funds, with over a fifth of the revenue coming from local council tax. The largest area of expenditure for local authorities is education, which accounts for over 43% of their total expenditures in 2001. Personal social services is the second largest area of expenditure, accounting for almost 19% of their total expenditure in 2001.

Of the total expenditure on personal social services in 2001, almost 45% was spent on services for the elderly, with the rest split between services for the disabled and services for children and families. Across these user groups, supporting individuals into residential accommodation and nursing homes was the largest form of expenditure, accounting for 46% of total spending, compared to 39% for day care and home help provision.

[7]The Isles of Scilly and City of London are excluded due to their unique local government structures.



(a) Gross Spending on the Elderly (per capita)



(b) Gross Spending on the Elderly (per pensioner)

Figure 4.2. Spending on Personal Social Services for the Elderly by Local Authorities in England, 2001
Notes: The figure displays the spending in 2001 of each of the 148 local authorities in England ordered along the x-axis by their spending. All amounts in 2015 £.
Source: Local Government Comparative Statistics (LGCS) from the Chartered Institute of Public Finance and Accountancy (CIPFA).

Around £7 billion (2015 £) was spent by all local authorities on the elderly in 2001, but the range in spending across councils showed a wide variation. Figure 4.2 displays the ordered distribution of the local authority gross expenditure on personal social services in 2001. Panel (a) displays spending per capita of the whole population, whereas panel (b) displays spending per (national government) pensioner, who are a key subpopulation that would likely make use of local social services, which is discussed in section 4.3.3. The figure highlights a heavy skew in the distribution, with the highest spending council spending almost four times more (per capita) than the lowest. The distribution becomes slightly more unequal in panel (b) when we consider spending per pensioner. Fernandez and Forder (2015) show that the skew in the distribution persists even when spending is averaged across many years.[8]

4.3.2 Social Service Provision and the SSPR

In late 2001, then Secretary of State Alan Milburn announced the introduction of publicly-released star ratings for social services in England, called the Social Services Performance Review (SSPR). In April 2002, a letter was sent to the Directors of Social Services from the Chief Inspector of the which described how the ratings would be produced. According to the Social Services Inspectorate (SSI) the purpose of the SSPR ratings were to "*improve public information about the current performance of services, and the prospects for improvement*" (Social Services Inspectorate, 2002).

[8]Moscone et al. (2007) find a similar distribution for spending on services for mental health by local authorities in England.

The SSPR ratings were to be similar to 'school league tables', which began in 1992 (Department for Education, 2018), and 'NHS hospital ratings' which began in 1995 (NHS Executive, 1995). Whereas those ratings provided information on the quality of education and hospitals in the area, the SSPR was to provide information on the quality of social services.

In 1995, the national government began measuring the performance of all local government services with the introduction of annual performance indicators (Audit Commission, 1995).[9] The indicators provided select comparative statistics for each local authority on the performance of all local services, including education and social services. The main purpose of this exercise was for the national government to ensure that local governments were getting value for money for their services. While some indicators were published, they were in general not easy for the public to access or interpret. The 2002 SSPR ratings used performance indicators related to social services, along with inspectors' judgments, to produce an overall star rating of the quality of social services in each area that was easy to interpret and widely publicized.

Councils were awarded either '3' stars ('excellent'), '2' stars ('good'), '1' star ('adequate') or '0' stars ('inadequate'). The best performing councils were given more freedom in the way they use national government provided grant funds. Councils with '0' stars were subject to more rigorous and frequent monitoring. The ratings that councils received were widely reported both in the local media and in the national media, with the BBC providing a dedicated web page listing the full results from every council ordered by star rating (BBC, 2002). Ratings were first published in May 2002, and were "refreshed" with additional information in November 2002 and then new ratings were released November of subsequent years. In this paper, the focus is on the initial ratings released in May 2002 as these ratings should have conveyed the most information to the public. Subsequent ratings could have been affected if more users moved to the area putting a strain on the service. It is also possible that after the initial ratings councils may have been able to "play" the system once they were aware what criteria were affecting their rating. Data on the SSPR ratings from 2002 come from the Department of Health's SSI website.[10]

Figure 4.3a displays the 2002 SSPR ratings across local authorities in England. While a large set of performance evidence were used in constructing the star ratings, to try to ensure consistency a smaller set of Key Performance Indicators were also chosen. For these, a council could not be awarded the highest star rating if they failed to meet the "desired" level in even one of the Key Performance Indicators.[11] Spending levels were also factored into the star ratings, however higher spending did not necessarily lead to higher ratings, as some councils were penalized if they

[9]The performance indicators began being collected from the 1993/94 tax year and were renamed Best Value Performance Indicators (BVPIs) in 1997. BVPIs were discontinued in 2008.

[10]The Social Services Inspectorate was replaced by the Commission for Social Care Inspection in 2004 which was subsequently replaced by the Care Quality Commission in 2009. Their website can still be accessed through the UK Government Web Archive.

[11]There were a total of 11 Key Performance Indicators, which included statistics such as "Percentage of older people helped to live at home" and "Percentage of adults and older people receiving a statement of needs".

were deemed to be spending too much on certain services. This led, in some cases, to councils receiving fairly arbitrary star ratings. It is important to note that whether or not the star ratings convey information regarding the true underlying quality of social services in the area is not of prime importance to this study. What is more important is whether or not pensioners believe the information in the ratings and respond in their migration decisions.

One might be concerned that the SSPR star ratings may be associated with the quality of other local government services for which ratings exist and were published prior the SSPR ratings, such as information contained in the 'school league tables' ratings. I find almost no correlation between the SSPR ratings received in 2002 and school quality variables, such as the average GCSE points per pupil, and the percentage of schools put into special measures. Therefore, it is likely that the 2002 SSPR ratings presented new information to the public that they did not previously have. Section 4.3.4 also provides some evidence that the star ratings that areas received do not appear to be associated with other area characteristics which may influence migration decisions.

4.3.3 Pensioners

Local authorities have some discretion in not only the level of service they provide but also the eligibility criteria.[12] Differing eligibility criteria may mean that, for example, some elderly individuals that are eligible for certain services in one local authority may not be eligible in another. In order to facilitate a consistent comparison across local authorities, instead of looking at the *service users* as defined by the local authorities, I focus on *potential users* by considering the population that are state pension claimants from the national government, of which eligibility is not dependent on receipt of local social services.[13] The state pension age during the time period of this study (1999-2006) was 60 for women and 65 for men. Take up of the state government benefits for the elderly is near universal, and should not be affected by a release of information on local social services in any way.[14] The focus of this study is how pensioners are distributed across the country, according to the quality of the social services in their area. The main dependent

[12]The Fair Access to Care Services (FACS) guidelines were established by the UK Government in 2003 as a common framework for determining individuals' eligibility for social care services. However, local authorities continued to have some discretion in determining eligibility (Fernández and Snell, 2012).

[13]For this purpose of this study, I define *pensioners* as the population over state pension age who were claiming at least one of the key benefits: attendance allowance, disability living allowance, incapacity benefit, pension credit, state pension, and severe disablement allowance. The main national government benefit for the elderly in England is basic state pension, which is payable from the state retirement age. To be eligible for the full state pension individuals need to have paid national insurance contributions for 90% of working lifetime. Individuals with insufficient national insurance contributions are still able to receive a proportion of the state pension. The main national benefit for older disabled people is attendance allowance, which is for disabled individuals that require care. For an indepth overview of the UK benefits system during the time period this study focuses on see Emmerson and Leicester (2002) or Leicester and Shaw (2003).

[14]Emmerson and Leicester (2002) report that there were 10,963,000 basic state pension claimants in March 2001. However, the April 2001 Census reports that the UK resident population of men aged 65+ and women aged 60+ was only 10,810,878 (Office for National Statistics, 2001, Table P1). Part of this discrepancy can be explained by pensioners living abroad.

variable in this study is therefore the number of pensioners living in a local authority (as a percentage of all pensioners in the country).

4.3.4 Data and Descriptive Statistics

The analysis in this study is based on data from 148 English and 22 Welsh local authorities from several different sources between the years 1999 and 2006. Descriptive statistics (2001 means) and data sources of all the main variables used in the analysis are reported in Table 4.1. Full details of the data sources and can be found in the Data Appendix in section 4A.1.

The statistics in Table 4.1 are separated according to the area's 2002 SSPR rating, and show that councils separated by star ratings are similar across other mean control variables, such as the average weekly income. The '3' star councils have the highest house prices, however this is mainly driven by a couple councils located in London, and '2' star councils have the lowest average house prices. Even though house prices are highest in the '3' star councils, they have the lowest local taxes (council tax). Therefore the difference in the star ratings do not appear to have been driven by local taxes or other area characteristics.

Figure 4.3a displays the geographic variations in the SSPR star ratings received in 2002. There are a couple of things worth noting in the figure. First, there is no overall clear geographic pattern in the star ratings, especially concerning the lowest and highest ratings, where these councils are spread throughout the country. Secondly, comparing Figures 4.3a and 4.3b, which display the geographic variations in spending, there is no visible connection between the amount spent and the ratings received. Figures 4.3c and 4.3d display the geographic variations in how pensioners are distributed across the country and the percentage of each area's population that are pensioners. There does not appear to be any clear connection between the number of pensioners in an area and the star rating it received. Figure 4.3e displays the average house price, which are highest around London and the south of England. Figure 4.3f displays the net migration rate of the population aged over 60, which is the number of people who move into the area net of those who move out (as a percentage of that area's population). The figure shows that those aged over 60 tend to migrate away from London and the surrounding area and towards the North East and South West coasts of England. Importantly, the 2001 net migration for each area appears to have no connection to the 2002 SSPR ratings they received.

4.4 Identification Strategy

The main estimates in this paper come from a difference-in-differences (DD) approach. Given that in 2002 the public were informed about the quality of social services in each area, those who

	Overall	By 2002 SSPR Rating				Source
		0	1	2	3	
Outcome variables						
Average Percentage of Pensioners	0.676	0.679	0.613	0.796	0.575	D1
Total Percentage of Pensioners	100	6.790	49.653	39.004	4.600	D1
Number of Pensioners	60,392	60,680	54,746	71,141	51,363	D1
Population Aged Over 60	69,112	69,380	62,293	81,150	64,088	P1
Total Population	334,076	339,580	309,264	377,122	314,762	P1
Migration (Pop. Aged Over 60)						
In-Migration Rate (%)	1.555	1.311	1.570	1.588	1.501	P1
Out-Migration Rate (%)	1.848	1.807	1.971	1.605	2.141	P1
Net Migration Rate (%)	-0.293	-0.496	-0.401	-0.017	-0.640	P1
Area Characteristics						
Average House Price (2015 £)	144,990	139,710	149,905	122,490	239,642	L1
Average Weekly Income (2015 £)	967	934	993	924	1,013	A1
Political Control of LA						
Labour Dummy	0.426	0.400	0.432	0.449	0.250	E1
Conservative Dummy	0.230	0.200	0.185	0.245	0.625	E1
Lib. Dem. Dummy	0.047	0	0.074	0.020	0	E1
Other Party Dummy	0.007	0	0.012	0	0	E1
No Overall Control Dummy	0.291	0.400	0.296	0.286	0.125	E1
Average Council Tax	1,017	1,029	1,028	1,018	891	L2
Gross PSS Spending on the Elderly	192	198	190	190	212	L2
Gross PSS Spending on Other	255	267	265	225	313	L2
Number of Local Authorities	148	10	81	49	8	

Table 4.1. 2001 Mean Local Authority Level Characteristics by 2002 SSPR Rating

Notes: The in-migration rate is calculated as the number of individuals that migrated into an area by the end of a given year, divided by the size of the population in that area at the beginning of the year. The out-migration rate is calculated similarly for those who migrate out of an area. The net migration rate is the in-migration rate minus the out-migration rate. Political control of the LA is the party that has a majority of the elected seats in the council. Gross PSS spending is per capita (whole population).

Sources:

A1: Annual Survey of Hours and Earnings

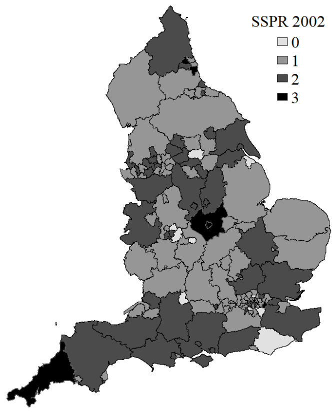
D1: From 5% samples of the DWP administrative data on the population over state pension age who were claiming at least one of the key benefits: Attendance Allowance, Disability Living Allowance, Incapacity Benefit, Pension Credit, State Pension, and Severe Disablement Allowance.

P1: Patient register data (1999-2008) from National Health Service Central Register (NHSCR).

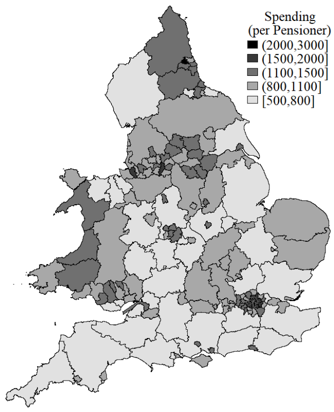
L1: Land Registry Price Paid data.

L2: Local Government Comparative Statistics (LGCS) from the Chartered Institute of Public Finance and Accountancy (CIPFA).

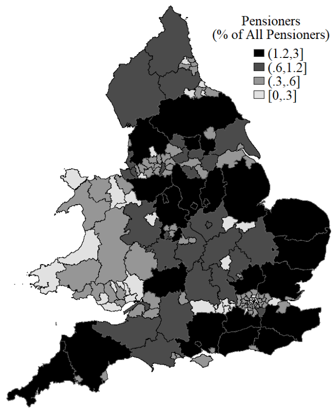
E1: Local council election results (1999-2006) are available on the BBC website.



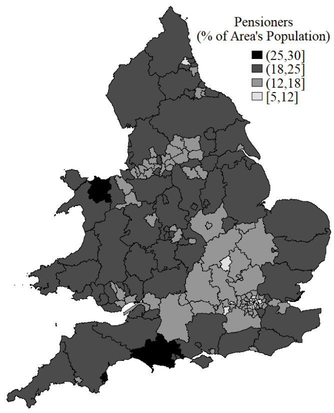
(a) 2002 SSPR Rating



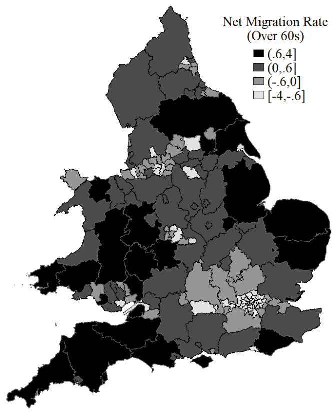
(b) PSS Spending (per Pensioner)



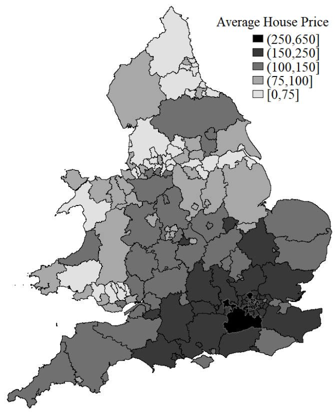
(c) Distribution of Pensioners Across Areas



(d) Pensioners (% Area's Population)



(e) Net Migration (%, Over 60s)



(f) Average House Price (000s)

Figure 4.3. Geographic Variations in 2001

Notes: This figure displays the 2001 geographic variations by local authorities (LAs) in England and Wales.

would benefit from having better social services in their area, specifically pensioners, may react by moving to that area after mid-2002. I treat this "information shock" as a quasi-experiment, and employ the DD estimator in the following way. I define the 2002 public release of the SSPR as a binary treatment variable, switching on from 2002 onwards. The DD model is based on comparing the percentage of pensioners living in areas with different quality of social services (based on their 2002 SSPR star rating), before and after the public release of information. Treating the year 2002 as a cut-off point in the empirical analysis rests upon two assertions. Firstly, before 2002, very few people knew the distribution of the quality of social services across the country. Second, the choice of using 2002 as a critical point in time is founded on the information in the publicly released SSPR ratings being widely spread.

The DD model can be written as follows:

$$Y_{it} = \beta_0 + \beta_1 R_i + \beta_2 P_t + \boldsymbol{\beta_{DD}(R_i \times P_t)} + \gamma^{'} X_{it} + u_i + \tau_t + \epsilon_{it} \tag{4.1}$$

where Y_{it} is the outcome of interest for area i in year t, R_i is the 2002 SSPR rating from 0 to 3, P_t is a dummy which takes value 1 after SSPR was introduced (2002 onwards), X_{it} denotes area characteristics, u_i is area fixed effects and τ_t is year fixed effects. The coefficient of interest is β_{DD}.

A potential problem with the DD estimate is that we do not observe the counterfactual; the areas with the highest star ratings may have seen this increase in the number of pensioners even without the publicly released ratings as some individuals may have had the information without the release. To address this, I attempt to find a control group. Welsh councils can be used to address the counterfactual question of what would have been the path of the distribution of the pensioner population across English LAs from 2002 onwards if the SSPR ratings would not have been publicly announced for the following reasons. First, Welsh and English LAs have the same structure and functions. Second, while the Welsh LAs have lower population densities than their English equivalents, the mean values of various control variables and quality measures are consistent across English and Welsh LAs, as can be seen in the appendix section 4A.3. Third, and crucially, while SSPR ratings were produced and publicly-released in England, no such ratings were produced for Welsh councils.

In the appendix section 4A.2, I predict the SSPR ratings that LAs in Wales would have received using accounting and performance measures in an ordered logit regression. A comparison between the predicted ratings (for the English LAs) and actual ratings can be seen in the appendix Figure 4A.1. Reassuringly, only a few councils have predicted ratings different than their actual rating, and only ever by one star.

Using the Welsh LAs as a control group for the English LAs, the empirical approach is to estimate the impact of the introduction of SSPR in 2002 on the outcome Y_{it} through difference-in-difference-in-differences (DDD) estimation. The DDD model can be written as follows:

$$\begin{gathered} Y_{it} = \beta_0 + \beta_1 P_t + \beta_2 E_i + \beta_3 \hat{R}_i \\ +\beta_4(P_t \times E_i) + \beta_5(P_t \times \hat{R}_i) \\ +\beta_6(E_i \times \hat{R}_i) + \boldsymbol{\beta_{DDD}(P_t \times E_i \times \hat{R}_i)} \\ +\gamma' X_{it} + u_i + \tau_t + \epsilon_{it} \end{gathered} \tag{4.2}$$

where E_i is a dummy which takes value 1 if the area is in England (the treatment group) and $\hat{R}_i$ is the predicted 2002 SSPR rating from 0 to 3. The coefficient of interest is β_{DDD}.

4.4.1 Assessing the Identification Strategy

The fundamental identifying assumption underlying the DD (and DDD) method is the assumption that time effects or trends are the same in the absence of the treatment (the SSPR ratings information release). In other words, the variable of interest should follow the same time path in each rating group in the absence of the treatment, conditional on other characteristics. This is not directly testable. However, here I provide some evidence that the common trends assumption cannot be rejected.

Visual inspections of the pre-reform years for all the main outcome variables as displayed in Figure 4A.2 in the appendix show that, with only few exceptions, areas did not appear to have different trends before the reform according to what rating they received.

A more formal test that areas followed similar trends before the treatment can be conducted by running, for the pre-treatment period from 1999 to 2001, the regression

$$Y_{it} = \tau_t + \theta_t(\tau_t \times R_i) + \gamma' X_{it} + u_i + \tau_t + \epsilon_{it} \tag{4.3}$$

where θ_t is the parameter of interest. Given that the social service ratings began in 2002, the null hypothesis that the variable of interest follows the same time path is simply $H_0 : \theta_{99} = \theta_{00} = \theta_{01} = 0$. Table 4A.3 in the appendix displays the p-values related to that null hypothesis for all of the main outcome variables considered in this paper. In almost every case, the hypothesis that they did follow a common time path cannot be rejected. In the same table, results are displayed for a test of common trends in the DDD framework. Again, in almost every case the hypothesis that they did follow a common time path cannot be rejected.

Dependent Variable: Percentage of Pensioners in Area				
	DD (Difference-in-Differences)		DDD (Triple Differences)	
	(1)	(2)	(3)	(4)
Information Effect of Star Rating				
Difference Term	0.0078**	0.0090***	0.0073**	0.0067*
	(0.0035)	(0.0032)	(0.0032)	(0.0035)
All controls		✓		✓
Number of LAs	148	148	170	170
Mean of Dep. Var.	0.676	0.676	0.588	0.588
Observations	1,184	1,184	1,360	1,360

Notes: The DD estimates in this table come from regression equation (4.1) and the DDD estimates from regression equation (4.2). The dependent variable is the number of pensioners living in a local authority (LA) as a percentage of all pensioners in the country. All regressions include LA fixed effects and year effects. Regressions in columns (1) and (2) are based on 8 years (1999-2006), for 148 English LAs. Columns (3) and (4) also include 22 Welsh LAs. Control variables are those listed under Area Characteristics in Table 4.1. Standard errors are clustered at the LA level.
* $p < 0.1$, ** < 0.05, *** $p < 0.01$.

Table 4.2. Main Results: Distribution of Pensioners Across Areas

4.5 Main Empirical Results

This section presents the main results from the analysis. I start by considering how the release of the SSPR ratings in 2002 affected the distribution of pensioners across areas, in terms of the numbers of pensioners (as a percentage of all pensioners), in areas that received different ratings. I provide robustness checks for these results. I then consider how the release affected different subgroups and population migration.

4.5.1 Distribution of Pensioners Across Areas

The main results use the distribution of pensioners across areas as the dependent variable in the DD and DDD models of equations (4.1) and (4.2).[15] The estimates reported in Table 4.2 show that the release of star ratings in 2002 did have an impact on the distribution of pensioners across areas. The DD estimates in column (2) suggest that a publicly-released one star rating increase is associated with a 0.009 percentage points (i.e. around 1.3%) increase in the percentage of all pensioners living in that area relative to others. The DDD estimates in columns (3) and (4), where Wales has been used as a control group, support the DD estimates with the point

[15] Precisely, the dependent variable is the number of pensioners that live in that area as a percentage of all pensioners in the country. This is calculated by dividing the number of pensioners in each area by the sum of all pensioners in the country and multiplying by 100. For the DD this only includes pensioners located in England, for the DDD specification it includes all pensioners located in both England and Wales.

Information Effect of Star Rating	DD Estimate	Std. Error	DDD Estimate	Std. Error
Base Specification	0.0090***	(0.0032)	0.0067*	(0.0035)
Including Trends				
LA-Specific Pre-reform Linear Trend	0.0090***	(0.0032)	0.0067*	(0.0035)
Excluding London LAs	0.0125***	(0.0044)	0.0106**	(0.0048)
Alternative variance estimators				
Bootstrapped Standard Errors	0.0090***	(0.0033)	0.0067*	(0.0036)
Collapsed for Two Periods	0.0059**	(0.0030)	0.0068*	(0.0036)
Alternative functional forms				
Log Transformation	0.0069*	(0.0040)	0.0062	(0.0070)
Non-Linear Estimation	0.0051***	(0.0010)	0.0062***	(0.0009)

Notes: The DD estimates in this table come from regression equation (4.1) and the DDD estimates from regression equation (4.2). The dependent variable is the number of pensioners living in a local authority (LA) as a percentage of all pensioners in the country. The Base Specification is column (2) in Table 4.2. All regressions include LA fixed effects, year effects and full controls unless otherwise stated. Control variables are those listed under Area Characteristics in Table 4.1. The number of observations for the specification excluding London LAs is 928 for the DD estimate, and 1,104 for the DDD estimate. The Non-Linear column is estimated using a Pooled Bernoulli quasi-MLE. Standard errors are clustered at the LA level.
* $p < 0.1$, ** < 0.05, *** $p < 0.01$.

Table 4.3. Further Robustness and Alternative Methodological Approaches

estimates remaining very similar in size, with a one star increase leading to a 0.008 percentage point (i.e. 1.4%) increase.

4.5.2 Robustness

While the DDD estimates provide some confidence in the robustness of the DD estimates, in this section I perform additional robustness checks, in particular with respect to some alternative methodological choices. I address inference and functional form dependence, and also perform a placebo test. The main robustness results can be seen in Table 4.3.

Additional controls The evidence presented in section 4.4.1 suggests that the introduction of the SSPR star ratings in 2002 is unrelated to most baseline area characteristics. Nevertheless, to examine whether the estimates are biased because of differential trends, I explicitly allow for differential LA-specific time trends. The LA-specific time trend is conducted by adding linear time trends extrapolated to the sample period based on pre-2002 data for each area. The estimates remain almost the same. Next, I check that the estimates are robust to excluding all

LAs located in the capital city (London) and the results again remain very similar, but become larger in magnitude.

Inference To ensure some outliers are not driving the statistical significance I bootstrap the standard errors. When this is done the standard errors remain essentially the same.

Bertand, Duflo, and Mullainathan (2004) point out that even with clustered standard errors, there can be downward bias in the standard errors, leading to false rejection of the null hypothesis of no treatment effect. To deal with this, I follow their recommended procedure of collapsing the time dimension to before and after the treatment, and re-estimating the model. This procedure produces very similar results to that of the baseline estimates in the paper.

Functional Form Dependence I check that the results are not dependent on the functional form by also estimating the model using a log transformation of dependent variable. I also estimate a non-linear version of the model using the methodology proposed by Papke and Wooldridge (2008) to tackle the possibility of non-linearity in case of fractional dependent variable. As the dependent variable lies between 0 and 1, I estimate a non-linear model as follows

$$Y_{it} = \Phi[\beta_{dd}(R_i \times P_t) + \gamma^{'} X_{it} + u_i + \tau_t + \epsilon_{it}] + v_{it} \tag{4.4}$$

using a pooled Bernoulli quasi-MLE. Standard errors are clustered at the LA level, allowing for serial correlation in the v_{it}. Both cases support the baseline estimates as being robust.

I also estimate an alternative highly parameterized model, similar to the specification in Figlio and Lucas (2004), which makes use of the subsequent SSPR ratings that were released after 2002. The results of this alternative specification can be seen in the appendix section 4A.5. This specification also supports the idea that the release of SSPR ratings affected the composition of the population, and led to an increase in pensioners in the areas that got the highest ratings relative to others.

Placebo In order to assess to what extent the method is sensitive to picking up effects that are unrelated to the phenomenon in question, I shift the outcome measure relative to the information release, to act as a placebo test. This test is constrained by the limited number of observations in the pre-release period, however the public release of information should have no effect on the outcome in the years before the release. Shifting the outcome reduces the size of the effect and decreases the precision of the estimate, however the effect is still detected as future years still pick up the effect. In every case, shifting the outcome variable backwards (or forwards) reduces

Information Effect of Star Rating	DD Estimate	Std. Error	DDD Estimate	Std. Error
Pensioners	0.0090***	(0.0032)	0.0067*	(0.0035)
Benefit Claimants of Working Age	0.0038	(0.0033)	0.0052	(0.0036)
Disability Related	0.0030	(0.0045)	0.0059	(0.0049)
Income Related	-0.0031	(0.0047)	-0.0030	(0.0044)
Unemployment Benefits (JSA)	-0.0352***	(0.0132)	-0.0336***	(0.0127)
House prices (log)	0.0170	(0.0109)	0.0038	(0.0133)

Notes: The DD estimates in this table come from regression equation (4.1) and the DDD estimates from regression equation (4.2). All regressions include local authority (LA) fixed effects, year effects and full controls unless otherwise stated. Control variables are those listed under *Area Characteristics* in Table 4.1. Standard errors are clustered at the LA level.
* $p < 0.1$, ** < 0.05, *** $p < 0.01$.

Table 4.4. Distribution of Benefit Claimants Across Areas

the size of the estimate which indicates that 2002 is the correct treatment year. This can be seen depicted in the appendix Figure 4A.3.

4.5.3 Distribution of Other Potential Service Users Across Areas

There are reasons to expect that the release of the SSPR ratings may affect other subgroups of the population who may also be users of social services, such as the low income or the disabled. In this section I compare the effects on the pensioner population with other national government benefit claimants.

Panel (a) of Table 4.4 shows that the release of the SSPR ratings had a stronger effect on pensioners than on working age benefit recipients. The DD point estimate for the working age benefit recipients is smaller at 0.004 percentage points and is not statistically significant. When the working age benefit recipients are separated into their main area of benefits (income related or disability related) it can be seen that most of the overall point estimate comes from the disabled, and the point estimate for those on income related benefits is actually negative. However, neither are statistically significant. This may point to the fact that those on disability related benefits are less mobile than elderly benefit claimants.

As expected, and as an additional robustness check, the SSPR ratings release appears to have no positive effect on the number of unemployment benefit (i.e. jobseeker's allowance) claimants in the area, with the point estimate being negative. This is reassuring as there is little reason to think that those claiming unemployment benefits would be users of local social services.

4.5.4 Migration

One common criticism of studies looking at welfare migration but focusing on the stock or composition of the population instead of migrants in or out of an area, is that of endogenous participation. Instead of picking up a welfare migration effect any estimates may instead pick up an effect of more individuals in an area starting to claim benefits (and not previous claimants moving into the area). This is particularly likely when a certain area introduces more generous benefits. This should not be a problem in this study, as the main outcome that is the percentage of the population claiming *national* state pension benefits, which is not dependent on the area they live in and whether or not they make use of local services. This segment of the population is the one that is most likely to benefit from local government social services, however the pensioners may or may not be eligible for these services depending on the area. Nevertheless, to ensure endogenous participation is not biasing the estimates, I check if the introduction of SSPR in 2002 had an impact on the actual migration into the area.

Indirect evidence on the migration response can be seen in the house price response. Given this information release should have made areas with higher star ratings more desirable we might see this increased demand for an area reflected in a change in house prices. The DD estimates in Table 4.4 shows that a one increase in publicly-released star rating was related to a 1.7 % percent increase in house prices, however the estimates are not statistically significant.

The next set subsection considers estimates from migration data for the population aged 60 and older. These movers cannot be subdivided into whether or not they were state pension claimants, but the estimates will still shed light on the overall effect.

4.5.4.1 In-Migration, Out-Migration and Net Migration I define the in-migration rate as the number of individuals that migrated into an area by the end of a given year, divided by the size of the population in the LA at the beginning of the year. Similarly, the out-migration rate is the number of individuals that migrated out of an area by the end of a given year, divided by the size of the population in the LA at the beginning of the year. The net migration rate is simply the in-migration rate minus the out-migration rate. The estimates that follow only focus on the migration rates of the over 60 population.

Panel (a) of Table 4.4 shows that the release of the SSPR ratings appears not to have had much, if any, of an effect on the in-migration rate of the area, with the DD point estimates being negative and imprecise. However, there appears to have been a sizable effect on the out-migration rate of the area with a -0.037 percentage point decrease in the over 60 population out-migration rate for

Migration	Panel (a): Overall Migration Rates by LA (%)					Panel (b): Migration Flow Rates Between LAs (%)				
Information Effect of Star Rating	Mean	DD	SE	DDD	SE	Mean	DD	SE	DDD	SE
In-Migration Rate (Aged 60+)	1.3706	-0.0105	(0.0147)	0.0403	(0.0396)	0.0085	-0.0001	(0.0001)	0.0001	(0.0001)
Aged 60-74	1.3421	-0.0183	(0.0182)	0.0370	(0.0445)	0.0083	-0.0000	(0.0001)	0.0001	(0.0001)
Aged 75+	1.4302	0.0004	(0.0181)	0.0488	(0.0451)	0.0091	-0.0002	(0.0001)	0.0001	(0.0002)
Out-Migration Rate (Aged 60+)	1.6642	-0.0369**	(0.0174)	-0.0203	(0.0303)	0.0102	-0.0003***	(0.0001)	-0.0003***	(0.0001)
Aged 60-74	1.6394	-0.0485**	(0.0212)	-0.0220	(0.0187)	0.0095	-0.0004***	(0.0001)	-0.0003***	(0.0001)
Aged 75+	1.7119	-0.0220	(0.0186)	0.0004	(0.0350)	0.0107	-0.0002*	(0.0001)	-0.0003	(0.0002)
Net Migration Rate (Aged 60+)	-0.2936	0.0264	(0.0243)	0.0274	(0.0243)	-0.0016	0.0002*	(0.0001)	0.0004**	(0.0001)
Aged 60-74	-0.2973	0.0302	(0.0283)	0.0318	(0.0282)	-0.0016	0.0003**	(0.0001)	0.0004**	(0.0002)
Aged 75+	-0.2817	0.0224	(0.0274)	0.0484	(0.0511)	-0.0016	0.0000	(0.0001)	0.0003	(0.0003)
Fixed Effects		Year, LA		Year, LA			Year, LA-pair		Year, LA-pair	
Number of LAs		148		170			148		170	
Number of LA-Pairs		–		–			10,878		14,365	
Observations		1,184		1,360			174,048		229,840	

Notes: In panel (a), the DD estimates in this table come from regression equation (4.1) and the DDD estimates from regression equation (4.2). In panel (b) the DD estimates come from a modified version of equation (4.1), which is given in equation (4.5). The DDD estimates in panel (b) come from the same modification for regression equation (4.2). The in-migration rate is calculated as the number of individuals that migrated into an area by the end of a given year, divided by the size of the population in that area at the beginning of the year. The out-migration rate is calculated similarly for those who migrate out of an area. The net migration rate is the in-migration rate minus the out-migration rate. In panel (a) the dependent variables are the overall migration rates for each LA, whereas in panel (b) the dependent variables are the the migration rates between LAs. All regressions include include local authority (LA) or LA-pair fixed effects, year effects and full controls unless otherwise stated. Control variables are those listed under *Area Characteristics* in Table 4.1. Regressions are based on 8 years (1999-2006). DD columns include 148 English LAs, DDD columns also include 22 Welsh LAs. Standard errors in panel (a) are clustered at the LA level, and in panel (b) are clustered at LA-pair level.
* $p < 0.1$, ** < 0.05, *** $p < 0.01$.

Table 4.5. Effects on Migration

a one star increase in SSPR rating. This suggests that people are less likely to leave an area once they find out that the social services are good. Most of this affect appears to be coming from those aged 60-74, which may indicate that either those over 75 are less mobile or are less likely to respond to an information release.

The DD estimate shows that the net migration rate of an area as rose by 0.026 percentage points for a public release of a one increase in star rating, however the estimates are not statistically significant. This estimate implies that the SSPR ratings release resulted in a smaller increase in the size of the elderly population in areas that performed well than what was found in the main estimates in Table 4.2. However, the estimates are not directly comparable as Table 4.2 focuses on state pension claimants, whereas Table 4.4 is focusing on the whole population aged over 60.

4.5.4.2 Migration Between Areas While the results in Panel (a) of Table 4.4 appear to show that the outflow of the over 60s was affected by the information release, some of the results are imprecise and there does not appear to have been any effect on in-migration. The response may be weak due to some frictions between areas, such the distance between them. To investigate this possibility, I examine the flows between LA-pairs. In particular, equation (4.1) is adjusted by including an LA-pair fixed effects so that the equation becomes

$$Y_{ijt} = \beta_0 + \beta_1 \widetilde{R_{ij}} + \beta_2 P_t + \widetilde{\boldsymbol{\beta}}_{\boldsymbol{DD}} (\widetilde{\boldsymbol{R_{ij}}} \times \boldsymbol{P_t}) + \gamma' \widetilde{X_{ijt}} + f_{ij} + \tau_t + \epsilon_{it} \tag{4.5}$$

where Y_{ijt} is the migration rate between area i and j (divided by the population size of area i), f_{ij} is an LA-pair fixed effect (with $f_{ij} = f_{ji}$), $\widetilde{R_{ij}}$ is the difference in star ratings between area i and j, and $\widetilde{X_{ijt}}$ is the difference of the control variables between area i and j. The coefficient of interest in this specification is is $\widetilde{\beta}_{DD}$. The interpretation of this estimate is how the migration rate between area i and j responds to the information being released regarding a one increase in the *difference* of star ratings across areas. Similar to before, equation 4.2 can be amended using Wales as a control group to calculate a DDD estimate, $\widetilde{\beta}_{DDD}$.

The results of this specification can be seen in Panel (b) of Table 4.4, and they line up with the estimates in Panel (a). In particular, a public release of a relative difference of one star in the SSPR ratings is associated with a reduction in out-migration of 0.0003 percentage points, which corresponds to a 3 % decrease in the outflow of pensioners from the area that received the higher rating to the area with the lower rating. Similarly, the release of SSPR is associated an increase in net migration. Again, it appears to be those aged 60-74 that are driving the estimates, with most of the point estimates for those aged over 75 being smaller and less precise.

4.6 Search Model with Nested Learning

So far, I have presented evidence that information does play a role in explaining the lack of welfare induced migration, by showing that the release of the SSPR ratings in 2002 affected pensioner migration. However, with the empirical evidence alone it is impossible to quantify the extent of the role that information plays, or to consider the impact of other factors such as moving costs. In this section, I present a search model with nested learning.

The model I present draws from the literature on equilibrium search unemployment, but instead of searching for jobs, pensioners are searching for the area that provides the best social services. The model, while on a different application, draws in particular from Moscarini (2005) and Papageorgiou (2014) who nest a job matching model into an equilibrium search unemployment framework. The model also draws from Pessino (1991), who analyses a sequential migration model with imperfect information.

In my model, pensioners search for the areas which provide the best social services. Social services is an experience good which is measured with noise. Pensioners have beliefs over the unknown quality of social services in every area and gradually learn the unknown quality of their current area's social services using Bayesian updating based on the observed services received. Based on their beliefs about the quality of the services in the area they are in (their "match"), and their beliefs over the services in other areas, each pensioner decides when and where to move.[16] The set-up is most similar to Papageorgiou (2014), who builds a job search model with three occupations and workers who learn their productivities in each job over time. In my model, beliefs are not reset upon moving to a new match, but are kept (and updated) until an individual dies. This is similar to Papageorgiou (2014) and Eeckhout and Weng (2010), but different than most of the rest of the literature (e.g. Jovanovic, 1979; Moscarini, 2005; Decreuse and Tarasonis, 2016; Li and Weng, 2017). Therefore, in the model, if a pensioner leaves an area due to the belief that the social services quality match is poor, they will keep that same belief about that area unless they return and update their beliefs or they receive outside information (i.e. an information shock).

4.6.1 Population of Social Service Users

The economy is entirely populated by risk-neutral pensioners, who search for the areas which provide the best social services. Pensioners receive utility from two sources: state pension payments from central government (which are the same for all pensioners regardless of area) and local social services (which have individual-specific match qualities) provided by an area's local

[16] Alternatively, this problem could have been set up as a reputation model (see Jolivet et al., 2016), with the councils being the seller and the pensioners being buyers.

council. The individual-specific true match quality for each council is unknown. While the true qualities are unknown, pensioners have beliefs about the quality of *every* council (not just the one they are currently in). Pensioners can choose to move for a fixed council-specific moving cost. There is a probability of death, γ. Each pensioner who dies is replaced immediately by a new pensioner, keeping the population constant. The model is in discrete time and the timing of events is as follow. Consider a pensioner beginning the period in a council j:

1. They move to a (random) new council with an exogenous probability δ.
   (a) If they do not move exogenously, they will assess their beliefs and decide if they want to move. If they choose to move, they will pick the council that they believe will offer them the highest quality. They will then move to that council, paying the council-specific fixed moving cost.
2. They receive a state pension payment from the central government, which is the same regardless of which council they are in.
3. They experience their local social services, which is specific to the pensioner and the council they are in. They then update their belief about the quality of their current council based on that experience. The beliefs about the councils they are not in remain unchanged.
4. With probability γ, they die.

If the pensioner does not die, the steps above are repeated.

4.6.2 Local Councils and Social Service Quality

In order to match with my empirical investigation earlier in the paper, in the model there are four different mean area council qualities: $\mu_0 \leqslant \mu_1 \leqslant \mu_2 \leqslant \mu_3$ which correspond to councils with '0', '1', '2' and '3' SSPR stars respectively. Assume, for now, that there are only four local councils, one of each quality type. Then for local council $j = 1, .., 4$, at birth into the model, pensioner i draws their unobserved individual-specific true qualities from the distribution:

$$q_{i,j} \sim N\left(\mu_j, \tau^2\right) \tag{4.6}$$

where τ determines the dispersion of individual qualities around the mean for each council. The individual-specific true qualities are time invariant.[17]

[17] While area amenity values (other than social services quality) are not explicitly incorporated in the model, it is possible to view other area amenity preference match qualities as being included in $q_{i,j}$.

The total per period utility received by pensioner i who is currently in council j is:

$$y_{i,j,t} = y^c + y^s_{i,j,t} \tag{4.7}$$

where $y^s_{i,j,t}$ is the observed local social services from council j at time t, and y^c is state pension payments from central government. The observed social services is an experience good, which is a function of the individual-specific true match quality but also has noise. Specifically, each period t the pensioner's observed social services is drawn from the distribution:

$$y^s_{i,j,t} \sim N\left(q_{i,j}, \sigma^2\right) \tag{4.8}$$

where σ determines the dispersion of the noise in the observed social services, which affects the variation in their social services experiences and will influence the speed at which pensioners learn.

4.6.3 Beliefs

To start, each pensioner i has their own initial belief about the quality of each council. For councils $j = 1, .., 4$, at $t = 0$,

$$p_{i,j,0} \sim N\left(\mu_j, \kappa^2\right) \tag{4.9}$$

where κ determines the dispersion of initial beliefs around the mean true qualities of the councils.

After observing their received social services each period, pensioners update their beliefs about the quality using Bayes' rule. Specifically, for the council j that the pensioner is currently living in, beliefs are updated according to the following equations:

$$p_{i,j,t_j+1} = \frac{y^s_{i,j,t_j}(\hat{\tau}_{i,j,t_j})^2 + \sigma^2 p_{i,j,t_j}}{(\hat{\tau}_{i,j,t_j})^2 + \sigma^2} \tag{4.10}$$

$$\hat{\tau}_{i,j,t_j+1} = \hat{\tau}_{i,j,t_j}\sqrt{\frac{\sigma^2}{(\hat{\tau}_{i,j,t_j})^2 + \sigma^2}} \tag{4.11}$$

where t_j is the total length of time that a pensioner has lived in their current council j.[18] Initially, at $t_j = 0$, for each pensioner and area, $\hat{\tau}_{i,j,0} = \tau$, but this value will get smaller the longer they live in a particular area and the more social services experiences they have. Over time pensioners will place less weight on new shocks to their observed social services versus their current beliefs when updating their beliefs. Note that each pensioner only updates the beliefs of the council they are living in, but has a memory of their current beliefs for every council. Therefore, $\hat{\tau}_{i,j,t_j}$ determines the accuracy of pensioner i's current belief in council j.

4.6.4 Value Functions

Let $\tilde{p}$ denote the vector of beliefs for each pensioner, and let $C(\tilde{p})$ denote their value function (excluding any moving costs) when they have the option of moving to any area, for ease of notation I suppress individual and time subscripts. Then,

$$\begin{aligned} C(\tilde{p}) = \quad & max\{y_1(\tilde{p}) + \beta(1-\delta)(1-\gamma)\mathbb{E}V_1(\tilde{p}) + \beta\delta(1-\gamma)\mathbb{E}\bar{V}(\tilde{p})\,, \\ & y_2(\tilde{p}) + \beta(1-\delta)(1-\gamma)\mathbb{E}V_2(\tilde{p}) + \beta\delta(1-\gamma)\mathbb{E}\bar{V}(\tilde{p})\,, \\ & y_3(\tilde{p}) + \beta(1-\delta)(1-\gamma)\mathbb{E}V_3(\tilde{p}) + \beta\delta(1-\gamma)\mathbb{E}\bar{V}(\tilde{p})\,, \\ & y_4(\tilde{p}) + \beta(1-\delta)(1-\gamma)\mathbb{E}V_4(\tilde{p}) + \beta\delta(1-\gamma)\mathbb{E}\bar{V}(\tilde{p})\,\} \end{aligned} \tag{4.12}$$

where $\bar{V}(\tilde{p})$ denotes the weighted average value function a pensioner would expect if they receive an exogenous move to a random council. Then, for example, the value function of a pensioner currently living in a council with '0' stars is

$$\begin{aligned} V_0(\tilde{p}) = \quad & \mathbb{I}\left(argmaxC(\tilde{p}) = 0\right) C(\tilde{p}) \\ & +(1 - \mathbb{I}\left(argmaxC(\tilde{p}) = 0\right))[C(\tilde{p}) - k_{argmaxC(\tilde{p})} \\ & +\beta(1-\delta)(1-\gamma)\mathbb{E}V_0(\tilde{p}) + \beta\delta\mathbb{E}\bar{V}(\tilde{p})] \end{aligned} \tag{4.13}$$

The solution to this problem (i.e. which is the preferred council, and whether or not to move) is derived numerically.

4.6.5 Estimating the Model

I estimate the model by matching features of the model to the data from my earlier empirical investigation. To start, I simulate individual histories for a large number of pensioners. Time

[18] For each pensioner, as there are 4 councils, $t = t_1 + t_2 + t_3 + t_4$. If the pensioner has only ever lived in council j, then $t_j = t$.

is discrete, where each period is a year. Whereas in the derivations above I assume there are only four councils, in the estimation there are six councils, however their remains only four mean council qualities.[19] There is one of each '0' and '3' star councils, and two of each '1' and '2' star councils. All pensioners start in a random council with draws of their unobserved individual-specific true qualities and their initial beliefs of the qualities for every council. The timing of events for each pensioner is as described in section 4.6.1. Each pensioner can move at most once during a period. If a pensioner dies, they are immediately replaced at the start of the next period by a new pensioner in a random council. I simulate this for a long time period to ensure a steady state is reached, and then introduce an "information shock" to mimic the release of the SSPR ratings as in the empirical investigation.

I introduce an "information shock" as follows. When the shock hits in a certain period, pensioners gain an additional social services experience for every council (not just the one that they are currently in). The additional experience does not update beliefs in the ordinary way according to equation (4.10). Instead, the experience brings beliefs, $p_{i,j,t}$, for every council j, closer to the true mean values of those councils, μ_j, by a percentage Δ_p. Specifically, at the time of the shock:

$$p_{i,j,t} = p_{i,j,t-1} + \Delta_p \left(\mu_j - p_{i,j,t-1}\right) \tag{4.14}$$

Pensioners that are born into the model after the shock has taken place have this experience immediately after birth. It is important to note that this shock does not necessarily bring beliefs closer to a pensioner's true values. This is due to the shock bringing beliefs closer to the true mean quality for each council, μ_j, not the true individual-specific values, $q_{i,j}$. This reflects the fact that the release of the SSPR ratings only gave councils a broad star rating, instead of giving pensioners personalized information. Sections 4.6.5.4 and 4.6.5.5 give a greater discussion on the form that the information shock takes, and consider the effects of an alternative shock which instead provides personalized information.

I estimate the model parameters using indirect inference (Smith, 1993; Gourieroux et al., 1993). Suppose that the parameters of the auxiliary model satisfy:

$$\bar{m} = m(X_N) = \arg\min_m Q(X_N, m)$$

where X_N is the observed data with N observations, and

[19] The reason that the model is estimated with only six councils, and not more, is due to the fact that pensioners have beliefs about the quality of *every* council. Therefore, each pensioner has the same number of state variables as there are councils. Increasing the number of councils has a very large computational cost, which is discussed more in the appendix section 4A.6. Kennan and Walker (2011) drastically reduce the size of the state space in their job search model by reducing the information set to only include wages seen in locations recently visited by the individual. In my application, however, this technique is not appropriate as the information shock is going to convey information about every area.

$$m^s(\theta) \equiv m(Y_N^s(\theta)) = \arg\min_m Q(Y_N^s(\theta), m)$$

where Y_N^s is the data generated from the s-th simulation of the model given parameter vector θ. Then the indirect inference estimator of θ, $\hat{\theta}$, satisfies:

$$\hat{\theta} = \arg\min_\theta \left(\bar{m} - \frac{1}{NS} \sum_{s=1}^{NS} m^s(\theta) \right)' W_N \left(\bar{m} - \frac{1}{NS} \sum_{s=1}^{NS} m^s(\theta) \right) \tag{4.15}$$

where W_N is an arbitrary positive definite weighting matrix, and NS is the number of simulations. More details on the computation can be found in appendix section 4A.6.

4.6.5.1 Functional Form Assumptions and Parameters All of the parameters of the model can be seen in Table 4.6. Payments from central government are normalized to 1. The discount rate, β, is set to 97%. The death rate, γ, is set at 5.3%, which gives an average life-expectancy just under 19 years.[20] The cost of moving to a '0' star council is estimated in the model, but the relative cost of moving to the other councils is based on the relative cost of house prices and average local authority rent for council housing in the area.[21] As expected, the '3' star council has the most expensive moving cost, however the differences in cost between the other councils is small. The last parameters that are pre-set are the exogenous move probabilities, which determine the probabilities that individuals are born in a certain council or that they move there if hit with an exogenous shock. These probabilities are set to match the number of councils in the data with that specific rating (dividing by the total number of councils). This distribution can be seen at the bottom of Table 4.1. The ten parameters that are estimated in the model are described in the panel (b) of Table 4.6.

4.6.5.2 Data Moments and Identification Given there are ten parameters to be estimated, I use ten moments in an exactly-identified estimation procedure. The moments that I will match are the distribution of the population across councils, the percentage of the population within each council each period who are migrating in, the percentage of moves that are to councils of

[20] In 2005, life expectancy at age 65 for men in the England was 17.4 years and for women was 20.1 years (Office for National Statistics, 2013).

[21] In the house price data, while the '3' star councils are the most expensive as expected, surprisingly, the '2' councils are the least expensive, however the cost differences between '0', '1' and '2' star councils are small. Often the pensioners that would benefit most from local social services will not own their own houses. To deal with this fact, I use data on the local authority average weekly rents for council housing. I give relative house prices and relative local authority average weekly rents equal weight when calculating the relative cost of moving found in Table 4.6.

Description	Parameter	Value
Panel (a): Pre-set Parameters		
Central Government Payment (State Pension)	y^c	1.000
Discount Rate	β	0.970
Per Period Death Rate	γ	0.053
Relative Cost of Migrating to '1' Star Councils	k_1/k_0	1.005
Relative Cost of Migrating to '2' Star Councils	k_2/k_0	1.010
Relative Cost of Migrating to '3' Star Councils	k_3/k_0	1.110
Prob. that Exogenous Move is to a '0' Star Council	$prob_0$	0.040
Prob. that Exogenous Move is to a '1' Star Council	$prob_1$	0.492
Prob. that Exogenous Move is to a '2' Star Council	$prob_2$	0.415
Prob. that Exogenous Move is to a '3' Star Council	$prob_3$	0.054
Panel (b): Parameter Estimates		
Mean Quality of '0' Star Councils	μ_0	0.280 (0.001)
Mean Quality of '1' Star Councils	μ_1	0.317 (0.010)
Mean Quality of '2' Star Councils	μ_2	0.318 (0.036)
Mean Quality of '3' Star Councils	μ_3	0.318 (0.045)
Exogenous Migration Rate	δ	0.011 (0.070)
Cost of Migrating to '0' Star Councils	k_0	3.100 (0.119)
Dispersion of Initial Beliefs	κ	1.100 (0.199)
Dispersion of Individual Qualities	τ	0.450 (0.094)
Dispersion of Social Service Noise	σ	2.000 (0.323)
Shock Parameter: *Change to p after information shocks*	Δ_p	0.580 (0.289)

Notes: The cost of moving to '1', '2', and '3' star councils are pre-set as a relative cost of moving to a '0' star council (which is an estimated parameter). When a pensioner is born into the model, or faces an exogenous move, the parameter $prob_j$ is pre-set to determine the probability that the pensioner will move to a 'j' star council. The shock parameter estimate, Δ_p, is determined by equation (4.14). Parameter estimates are from Indirect Inference estimation. The estimation procedure chooses a vector of parameter values to minimize the distance between features observed in the data and those generated by simulated data from the model, as displayed in Table 4.7. Standard errors in parenthesis are derived from numerical derivatives as described in appendix section 4A.6.1.

Table 4.6. Model Parameter Values, Pre-set (a) and Estimates (b)

Moment	Data	Model
Panel (a): Targeted		
% Population in Councils with '0' Stars	6.84	6.19
% Population in Councils with '1' Stars	49.64	51.03
% Population in Councils with '2' Stars	38.93	37.05
% Population in Councils with '3' Stars	4.59	5.72
In Migration Rate to Councils with '0' Stars	1.25	2.03
In Migration Rate to Councils with '1' Stars	1.50	1.77
In Migration Rate to Councils with '2' Stars	1.44	1.85
In Migration Rate to Councils with '3' Stars	1.40	1.82
% Migration to Council with Same Stars	39.84	35.09
Coefficient from DD regression (β_{dd})	0.11	0.11
Panel (b): Untargeted		
Out Migration Rate (Overall)	1.75	1.83
Out Migration Rate in Councils with '0' Stars	1.73	1.59
Out Migration Rate in Councils with '1' Stars	1.90	1.82
Out Migration Rate in Councils with '2' Stars	1.46	1.90
Out Migration Rate in Councils with '3' Stars	2.00	1.86

Notes: This table displays the targeted and untargeted moments from the model compared with that from the data.

Table 4.7. Data and Model Simulated Moments

the same quality, and the DD regression coefficient from the regression of the distribution of pensioners across areas on the star rating.[22]

With the amount of moments in this model, a rigorous identification argument is difficult. Even still, the moments that have been chosen to be matched are informative about the parameters in model that are being estimated. Firstly, the distribution of pensioners across the councils and the percentage of the population within each council who migrate in is informative of mean quality of each council, the dispersion of individual qualities, the dispersion of initial beliefs, the dispersion of the noise in the social services experience, and the cost of moving. The percentage of moves that are to councils of the same quality is informative for the exogenous migration probability (as most moves to councils of the same quality are likely to be exogenous). Finally, the DD regression coefficient is informative for the change to the shock parameter caused by an information shock.

4.6.5.3 Estimation of the Model Table 4.7 presents the observed and simulated moments from the model. The model does well in terms of matching most of the targeted moments in panel

[22] This is an adjusted version of the DD estimate found in Table 4.2. Given that the main baseline estimate is based on 148 areas, whereas in the model there are much fewer areas, I combine all councils of the same star rating and perform the same regression with only four councils. The resulting DD estimate has a coefficient of 0.110, with a standard error of 0.050.

(a), and untargeted moments in panel (b). The moment that the model has the most trouble matching is the percentage of movers that migrate to a council of the same quality. The reason for this is that the model is still quite restrictive with only six areas, including only one of each '0' and '3' star councils. In the data almost half of all moves are to councils of the same quality. In the model, only exogenous moves or those living in '1' or '2' star councils have the ability to move to councils of the same quality. Even if *all* moves were exogenous, given the exogenous move probabilities listed in Table 4.6 from the data, the model still will never be able to perfectly match that moment.

Panel (b) of Table 4.6 displays the parameter estimates from the simulated model. The estimation reveals some interesting results.

The estimated parameter for the exogenous migration rate is 1.1%, which given that the overall migration rate in the model is 1.8%, implies that around 60% of moves are exogenous and cannot be explained by the model. This is not surprising given that pensioners migrate for a variety of reasons other than social services. As expected, the dispersion of initial beliefs is large, and much larger than the dispersion of individual qualities.

The dispersion of the noise in the observed social services parameter is large, meaning that learning is slow. In fact, given the size of the parameter estimate it would take an average pensioner living in an around area 40 years for their beliefs to get within 25% of their true quality. Given that the average life expectancy in the model is around 19 years, most pensioners' beliefs are far away from their true qualities. The large estimate for the dispersion of noise also means that pensioners can have both large positive and negative social services experiences. Large negative shocks contribute to the overall migration rate as after a bad shock pensioners will revise their beliefs downwards and be more likely to migrate.

The utility and cost parameters also reveal interesting results. Recall that in the model central government payments have been normalized to 1. For my sample years, the standard state pension from UK government is £5,400 a year (in 2015 £s). Therefore, for social service users in the model, the mean utility from a '0' star council is £1,512. The mean utility of '1' star councils is higher at £1,710. The mean quality of '2' and '3' star councils are both slightly higher again at £1,715. The cost of moving to a '0' star council is £16,740, which is a substantial cost but may seem small when compared with estimates found in other migration models, such as Kennan and Walker (2011). However, when you consider than £16,740 is around twice the average yearly utility in the model these estimates are not that different than the estimates in Kennan and Walker (2011). It is also important to remember that compared to other migration models, this paper focuses on a very different population, and the migration costs were found using different sources of variation. The high moving costs are a result of the weak relationship in my data between migration and area quality (according to their SSPR star ratings). Although in the model part of that observed weak relationship can be accounted for by pensioners having incorrect beliefs. Therefore, a model with perfect information or more accurate beliefs would estimate even higher moving costs.

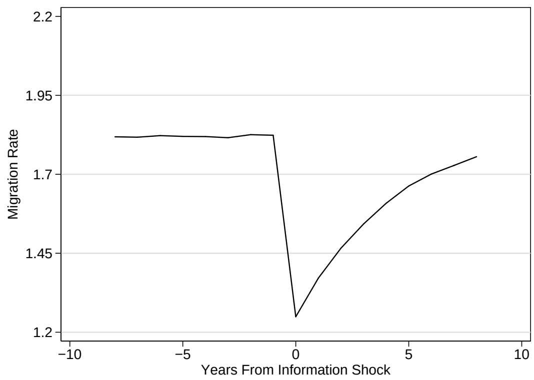
4.6.5.4 Introducing an Information Shock

As can be seen at the bottom of Table 4.6, for the model to match the DD regression from the data, pensioners' beliefs, across all areas, have to get closer to the mean qualities by 58%. Perhaps surprisingly, this information shock *decreases* gross migration at the shock and for a few periods afterwards, which can be seen in the Figure 4.4a. This is because the shock brings pensioners' beliefs closer to the mean for each area and therefore individuals with extreme beliefs (either good or bad) for certain councils revise those beliefs closer to the mean and are less likely to migrate. This is consistent with the estimates found in Panel (b) of Table 4.4, where the point estimates for out-migration are negative after the SSPR information shock. Figure 4.4a shows that after the initial decrease in migration, it gradually returns to the pre-shock level, as pensioners continue to have new social services experiences and update their beliefs accordingly.

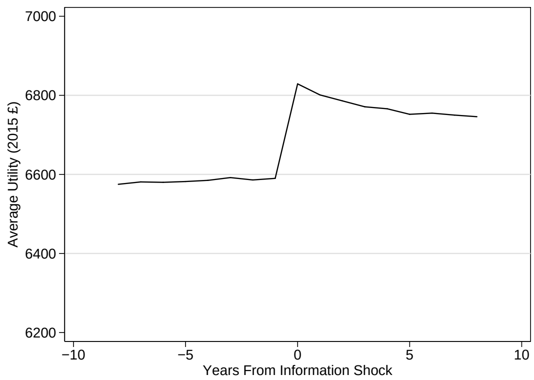
This information shock is not good for everyone, as it brings some pensioners beliefs further away from the truth. For example, consider a pensioner living in a council in which they have a bad match, and are considering migrating to a council in which they would have a good match. After this shock, their beliefs about the council they are currently in (with a bad match) are updated (upwards) closer to the mean for that area, and the council they were considering migrating to is updated (downwards) closer to the mean for that area. The pensioner is now less likely to migrate, and will stay in the bad match council for longer.

However, as can be seen in Figure 4.4b, overall the average per period utility rises after the shock by about £240 per year. Note that the shock itself does not have any direct impact on utility. The only way that the shock affects utility is indirectly by impacting migration decisions. A large part of the increase in per period utility, at least initially, is driven by less pensioners migrating and therefore not incurring the large moving costs. Although even as gross migration returns to the pre-shock levels, the increase to utility persists into the future. This means that on average the information shock was beneficial.

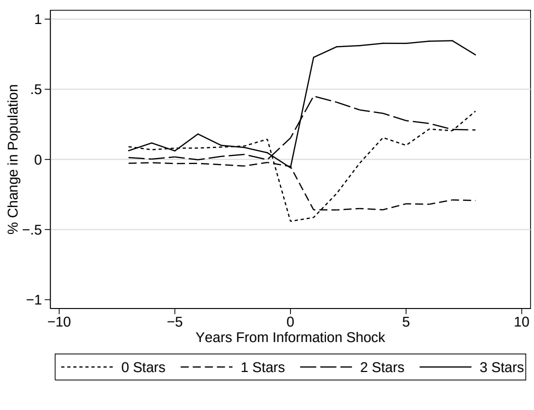
While there was a reduction in the gross migration rate, the more relevant statistic to consider when looking across council types is net migration, which will affect the population size. Figure 4.4c displays the yearly percentage change in the population sizes of the different types of councils after the information shock. As can be seen in the figure, at the time of the shock, '0' and '1' star councils see a drop in their population size, and '2' and '3' star councils see a rise. This



(a) Migration Rate

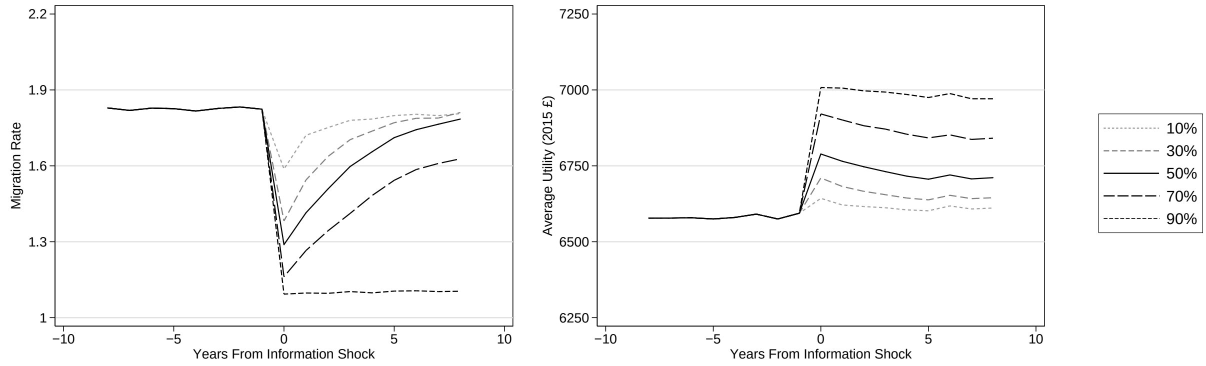


(b) Average Utility

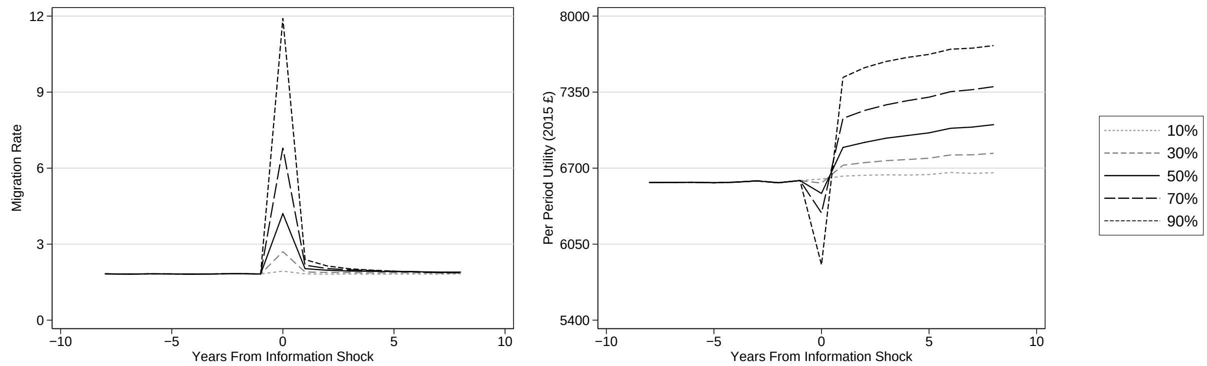


(c) Percentage Change in Population

Figure 4.4. Response to Information Shock in Model
Notes: This figure displays the results from the simulated model using the parameter values listed in Table 4.6.



(a) Change in Size of Information Shock



(b) Change in Form of Information Shock (Personalized Information)

Figure 4.5. Response to Changes in Size and Form of Information Shock

Notes: This figure displays the results from the simulated model using the parameter values listed in Table 4.6. Sub-figure (a), displays the simulated response when the size of the information shock parameter, Δ_p, takes different values. Sub-figure (b) changes the form of the information shock to be a personalized shock as described in footnote 23, and displays the simulated response when the personalized information shock parameter, $\Delta_{\bar{p}}$, takes different values.

shows that the information shock led to pensioners making better migration decisions, with the best councils seeing an increase in net migration and the worst councils seeing a decrease. The increase in net migration for '2' and '3' star councils and the decreases for '1' star council persist into the future, however the decrease to the '0' star council appears to be only temporary.

4.6.5.5 Changing Size and Form of Information Shock

This section considers the conterfactuals of changing the size of the information shock and the form of the shock.

To match the DD regression from the data, pensioners' beliefs have to get closer to the mean qualities by 58%, Figure 4.5a displays how migration and utility are affected if instead the information shock brings beliefs closer to the mean qualities by $10\% - 90\%$. As can be seen in the figure, all of the shocks reduce migration. As would be expected, bringing beliefs closer to the mean by only 10% has a very small effect on migration, and almost no effect on utility. In contrast, bringing beliefs closer to the mean by 90% has a large effect on migration and results in there being almost no endogenous migration in the model (which persists even 8 years later). This also results in increases in average period utility as a lot fewer pensioners are paying the large moving costs.

This information shock has been modeled to mimic the SSPR ratings, by bringing beliefs closer to the true mean quality for each council, not the true individual-specific values. Figure 4.5b displays shocks of different sizes that instead bring beliefs closer to the *true* qualities.[23] This mimics an information shock that gives individuals personalized information about their match qualities. As can be seen in Figure 4.5b, this type of information shock leads to different effects on migration and utility than the previous, not personalized, information shock. In each case, giving pensioners the personalized information shock results in a large increase in migration at the time of the shock, but the effect is temporary as pensioners immediately relocate to their preferred areas at the time of the shock and then remain there. This results in a temporary decrease in per period utility (as pensioners pay the large moving costs) but then average utility increases over time as more pensioners are in councils with which they have a good match.

4.7 Conclusion

While there is a substantial literature on welfare-induced internal migration, most studies find either no effect or a very modest one. This paper seeks to establish the role of information as part of the explanation for the lack of welfare-based migration. I assess whether the low observed rate of welfare migrants is due to individuals not knowing the quality of welfare programs in their area.

I focus on pensioners in England, where social services provision (which can be viewed as a benefit) is decentralized to local level and displays a wide variation in quality. Using a policy called the Social Services Performance Review (SSPR), which was introduced in 2002, where the national government gave a publicly-released rating of each local authority's social services on a scale from zero to three stars based on a series of accounting and performance measures, I analysis the distribution of pensioners across areas before and after the information shock occurred. The hypothesis is that that if information does play a role then we would expect

[23] Specifically, at the time of the shock, equation (4.14) is instead $p_{i,j,t} = p_{i,j,t-1} + \Delta_{\tilde{p}}\left(q_{i,j} - p_{i,j,t-1}\right)$.

that areas that receive a higher star rating would see a increase in pensioners compared to the areas which performed poorly in the ratings, after this information is made public. Using a difference-in-differences approach I find that a one increase in the publicly-released star rating led to a 0.01 percentage points increase in the percentage of all pensioners that live in the area relative to others, which corresponds to a 1.3 percent increase. I interpret this increase as being caused by migration and provide evidence from migration data. The migration estimates suggest that the release of the SSPR does not appear to have had much of an effect on in-migration, but did impact out-migration, with pensioners less likely to leave areas that performed well in the ratings.

I use the results from my empirical investigation to motivate a search model with nested learning. The model suggests that there is a lot of noise in the learning process, and that it takes pensioners a long time to learn the true quality of social services in their area. To generate a response of the same magnitude as that observed in the data in response to the SSPR release, the information shock would have had to bring pensioners' beliefs about the quality of each area closer to the true mean quality of those areas by 58%. Overall, I find that the information release resulted in a temporary reduction in total gross migration as pensioners with extreme beliefs revised those beliefs closer to the mean. The best councils saw increases in net migration whereas the worst councils saw decreases in net migration. On average, this had a positive effect on utility, which persisted into the future. The form that the information shock takes, and whether it provides personalized information or not, is shown to be important, with personalized information releases potentially offering much greater gains in utility.

This paper shows that information can offer a partial explanation for why there is a lack of internal welfare-migration. The findings suggest that information releases - especially regarding the quality of local services or benefits - can be met with migratory responses and lead to greater differences in population compositions across areas.

Bibliography

Audit Commission (1995). *Local Authority Performance Indicators: Vol. 1, Education, Social Services and Total Expenditure.*

BBC (2002). *Social Services Ratings.* `http://news.bbc.co.uk/hi/english/static/in_depth/health/2002/social_services_ratings/default.stm` [Accessed: 2017/11/20].

Brookshire, D. S., Thayer, M. A., Tschirhart, J., and Schulze, W. D. (1985). A test of the expected utility model: evidence from earthquake risks. *Journal of political economy*, 93(2):369–389.

Brueckner, J. K. (2000). Welfare reform and the race to the bottom: Theory and evidence. *Southern Economic Journal*, pages 505–525.

Carrillo, P., Cellini, S. R., and Green, R. K. (2013). School quality and information disclosure: Evidence from the housing market. *Economic Inquiry*, 51(3):1809–1828.

Chen, Y. and Rosenthal, S. S. (2008). Local amenities and life-cycle migration: Do people move for jobs or fun? *Journal of Urban Economics*, 64(3):519–537.

Conway, K. S. and Houtenville, A. J. (1998). Do the elderly vote with their feet?'. *Public Choice*, 97(4):663–685.

Conway, K. S. and Houtenville, A. J. (2001). Elderly migration and state fiscal policy: evidence from the 1990 census migration flows. *National Tax Journal*, pages 103–123.

Decreuse, B. and Tarasonis, L. (2016). Statistical discrimination in a search equilibrium model: Racial wage and employment disparities in the us.

Dennett, A. R. (2010). *Understanding internal migration in Britain at the start of the 21st century*. University of Leeds.

Department for Education (2018). *Compare school and college performance*. `https://www.compare-school-performance.service.gov.uk/` [Accessed: 2018/05/19].

Eeckhout, J. and Weng, X. (2010). Assortative learning. *WPU Pennsylvania*.

Emmerson, C. and Leicester, A. (2002). *A survey of the UK benefit system*. Institute for Fiscal Studies.

Fernandez, J.-L. and Forder, J. (2015). Local variability in long-term care services: Local autonomy, exogenous influences and policy spillovers. *Health economics*, 24(S1):146–157.

Fernández, J.-L. and Snell, T. (2012). Survey of fair access to care services:(facs) assessment criteria among local authorities in england.

Figlio, D. N. and Lucas, M. E. (2004). What's in a grade? school report cards and the housing market. *The American Economic Review*, 94(3):591–604.

Fiva, J. H. and Kirkebøen, L. J. (2011). Information shocks and the dynamics of the housing market. *The Scandinavian Journal of Economics*, 113(3):525–552.

Gelbach, J. B. (2004). Migration, the life cycle, and state benefits: How low is the bottom? *Journal of Political Economy*, 112(5):1091–1130.

Gourieroux, C., Monfort, A., and Renault, E. (1993). Indirect inference. *Journal of applied econometrics*, 8(S1).

Graves, P. E. and Knapp, T. A. (1988). Mobility behavior of the elderly. *Journal of Urban Economics*, 24(1):1–8.

Haisken-DeNew, J. P., Hasan, S., Jha, N., and Sinning, M. (2016). Unawareness and selective disclosure: The effect of school quality information on property prices.

Imberman, S. A. and Lovenheim, M. F. (2016). Does the market value value-added? evidence from housing prices after a public release of school and teacher value-added. *Journal of Urban Economics*, 91:104–121.

Jolivet, G., Jullien, B., and Postel-Vinay, F. (2016). Reputation and prices on the e-market: Evidence from a major french platform. *International Journal of Industrial Organization*, 45:59–75.

Jovanovic, B. (1979). Job matching and the theory of turnover. *Journal of political economy*, 87(5, Part 1):972–990.

Kennan, J. and Walker, J. R. (2010). Wages, welfare benefits and migration. *Journal of Econometrics*, 156(1):229–238.

Kennan, J. and Walker, J. R. (2011). The effect of expected income on individual migration decisions. *Econometrica*, 79(1):211–251.

Leicester, A. and Shaw, J. (2003). *A survey of the UK benefit system*. Institute for Fiscal Studies.

Li, F. and Weng, X. (2017). Efficient learning and job turnover in the labor market. *International Economic Review*, 58(3):727–750.

Lockwood, B. and Porcelli, F. (2013). Incentive schemes for local government: Theory and evidence from comprehensive performance assessment in england. *American Economic Journal: Economic Policy*, 5(3):254–286.

McKinnish, T. (2007). Welfare-induced migration at state borders: New evidence from micro-data. *Journal of public Economics*, 91(3):437–450.

Meyer, B. D. et al. (1998). *Do the poor move to receive higher welfare benefits?* Institute for Policy Research, Northwestern University.

Moscarini, G. (2005). Job matching and the wage distribution. *Econometrica*, 73(2):481–516.

Moscone, F., Knapp, M., and Tosetti, E. (2007). Mental health expenditure in england: a spatial panel approach. *Journal of Health Economics*, 26(4):842–864.

NHS Executive (1995). *The NHS Performance Guide 1994-1995*.

Office for National Statistics (2001). *Census 2001: First results on population for England and Wales*.

Office for National Statistics (2013). *Life Expectancy at Birth and at Age 65 for Local Areas in England and Wales: 2009-11*.

Papageorgiou, T. (2014). Learning your comparative advantages. *Review of Economic Studies*, 81(3):1263–1295.

Papke, L. E. and Wooldridge, J. M. (2008). Panel data methods for fractional response variables with an application to test pass rates. *Journal of Econometrics*, 145(1):121–133.

Pessino, C. (1991). Sequential migration theory and evidence from peru. *Journal of Development Economics*, 36(1):55–87.

Revelli, F. (2006). Performance rating and yardstick competition in social service provision. *Journal of Public Economics*, 90(3):459–475.

Schwartz, A. L. and Sommers, B. D. (2014). Moving for medicaid? recent eligibility expansions did not induce migration from other states. *Health Affairs*, 33(1):88–94.

Smith, A. A. (1993). Estimating nonlinear time-series models using simulated vector autoregressions. *Journal of Applied Econometrics*, 8(S1).

Social Services Inspectorate (2002). *Performance Ratings for Social Services in England 2001-2002*.

Wiswall, M. and Zafar, B. (2015). How do college students respond to public information about earnings? *Journal of Human Capital*, 9(2):117–169.

4A Appendix

4A.1 Data Appendix

Data on the SSPR ratings come from the Department of Health's Social Services Inspectorate website.[24]

Local authority spending data and area characteristics come from the yearly Local Government Comparative Statistics (LGCS) available on the website of the Chartered Institute of Public Finance and Accountancy (CIPFA).

Data on state pension and other benefit claimants comes from the Department of Work and Pensions (DWP). A count of people claiming at least one key DWP working-age benefit by local authority is taken from the Work and Pensions Longitudinal Study (WPLS). Data on the pension-age client group comes from 5% samples of the DWP administrative data on the population over state pension age who were claiming at least one of the key benefits: attendance allowance, disability living allowance, incapacity benefit, pension credit, state pension, and severe disablement allowance.

Migration numbers into and out of local authorities come from Patient register data (1999-2008). This data is from National Health Service Central Register (NHSCR) and contains number of people moving from a GP in one area to another. It has both the inflow and outflow by broad age bands across all the areas in England and Wales. Dennett (2010) shows data is of high quality and the migration numbers match well with the 2001 Census.

Various control variables are linked to the data by local authority, this includes house price data are from the Land Registry, and the average weekly full time wage comes from the Annual Survey of Hours and Earnings.

4A.2 Constructing a Control Group (Predicting 2002 SSPR Ratings for Wales)

The 2002 SSPR Assessment included evidence from inspections and reviews, monitoring and performance indicators (Social Services Inspectorate, 2002). In order to ensure that the performance indicators had sufficient weight in the star rating system, and to provide an additional consistency check to ensure that councils were treated in the same way, a subset of 11 performance indicators were defined as the Key Performance Indicators, most of which exist for both England and

[24]The Social Services Inspectorate was replaced by the Commission for Social Care Inspection in 2004 which was subsequently replace by the Care Quality Commission in 2009. Their website can still be accessed through the UK's National Archives website.

Determinants of SSPR Ratings 2002		
Variables	Ordinal Logit	
Older people helped to live at home (% over 65s)	0.113**	(0.040)
Adults and older people receiving a statement of needs (% population)	0.123***	(0.054)
Number of Day Care Clients		
Elderly	0.006*	(0.003)
Physical Disability	-0.022*	(0.011)
Learning Disability	0.020**	(0.009)
Mental Disability	-0.021**	(0.010)
Accounting cost controls (Gross Total Cost)	✓	
Pseudo R^2	0.65	
Observations	148	

Notes: The sample includes all English Local Authorities (LAs) in 2002. The dependent variable is the 2002 SSPR star rating for each English LA. The accounting cost controls include the gross total cost in each service area related to PSS. Coefficients are displayed as odds ratios.
* $p < 0.1$, ** < 0.05, *** $p < 0.01$.

Table 4A.1. Predicting SSPR for Wales

Wales. In practice, a lot of weight was put on the assessors and not all of the same performance indicators are available for Wales.

To get around this I construct the control sample using a ordinal logit regression model to predict the 2002 SSPR ratings based on a wide range of personal social services accounting level data and comparable performance indicators (that exist for both England and Wales) for each local authority. I also include dummies for the accounting area of each local authority (to reflect that costs vary by area, as set out by the SSPR guidelines) and set Wales to be in the same cost group as the Midlands. The results from the ordinal logit model can be seen in Table 4A.1. Figure 4A.1 displays the geographic variation in the predicted social service ratings.

4A.3 Testing the Identification Strategy

In this subsection I present some comparison statistics for England and Wales and test the identification strategy. Table 4A.2 displays summary statistics for all local authorities in England and Wales, separated by their predicted SSPR. Figure 4A.2 allows for visual inspections of the pre-reform years for all the main outcome variables. Table 4A.3 tests the common trends

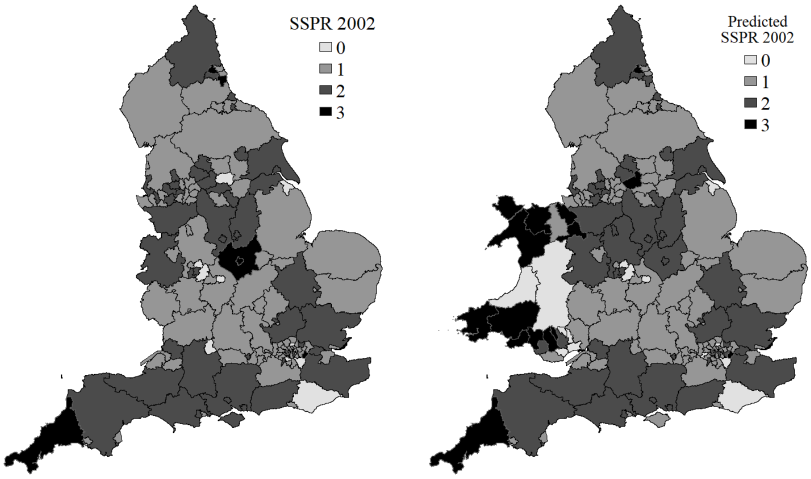


Figure 4A.1. Geographic Variation: 2002 Actual and Predicted SSPR Ratings
Notes: This figure displays the actual 2002 SSPR star ratings in England (left) versus the predicted 2002 SSPR star ratings for England and Wales (right).

assumption as described in section 4.4. Regressions testing for common trends in the DDD framework use the following equation,

$$Y_{it} = \tau_t + \theta_{ddd,t}(\tau_t \times R_i \times E_i) + \theta_{1,t}(R_i \times E_i) + \theta_{2,t}(\tau_t \times E_i) + \theta_{3,t}(\tau_t \times R_i) + \gamma' X_{it} + u_i + \tau_t + \epsilon_{it} \quad (16)$$

4A.4 Additional Figures

4A.5 Alternate Empirical Specification

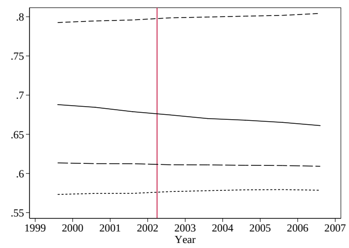
I estimate an alternative highly parameterized model, similar to the specification in Figlio and Lucas (2004). The model is specified as

$$Y_{it} = \tau_t + \lambda G_{it} + \gamma' X_{it} + u_i + \tau_t + \epsilon_{it} \quad (17)$$

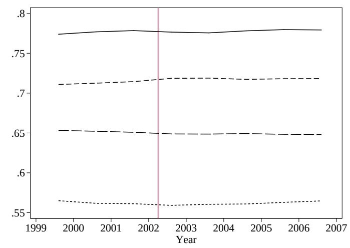
	Predicted 2002 SSPR Rating							
	England				Wales			
	0	1	2	3	0	1	2	3
Outcome variables								
Avg. Percentage of Pensioners	0.692	0.552	0.791	0.473	0.263	0.217	0.281	0.296
Total Percentage of Pensioners	5.536	46.368	38.759	3.311	1.578	0.651	0.562	3.256
Pensioners	65,812	52,495	75,249	44,957	25,000	20,633	26,700	28,127
Population Aged Over 60	74,750	60,110	85,286	57,471	28,833	23,433	31,350	32,109
Total Population	365,663	297,604	398,106	287,443	135,067	99,133	149,200	136,673
Migration (Pop. Aged Over 60)								
In-Migration Rate (%)	1.256	1.441	1.406	1.411	1.502	2.316	1.259	1.824
Out-Migration Rate (%)	1.910	1.811	1.443	2.165	1.441	1.805	0.910	1.398
Net Migration Rate (%)	-0.654	-0.370	-0.036	-0.754	0.062	0.511	0.349	0.426
Area Characteristics								
Average House Price	146,431	146,823	125,225	259,697	87,858	102,664	78,394	73,768
Average Weekly Income	946	986	928	1041	831	832	837	826
Political Control of LA								
Labour Dummy	0.375	0.429	0.469	0.143	0.667	0	1	0.636
Conservative Dummy	0.250	0.179	0.265	0.571	0	0.333	0	0
Lib. Dem. Dummy	0	0.071	0.020	0	0	0	0	0
Other Party Dummy	0	0.012	0	0	0.333	0.333	0	0.273
No Overall Control	0.375	0.310	0.245	0.286	0	0.333	0	0.091
Average Council Tax	1067	1022	1020	884	770	877	792	808
Gross PSS Spending (per capita)								
Elderly	205	189	191	215	194	160	206	219
Other	296	260	227	336	249	207	271	211
Number of Local Authorities	8	84	49	7	6	3	2	11

Table 4A.2. 2001 Mean Local Authority-Level Characteristics by Predicted 2002 SSPR Rating, England and Wales

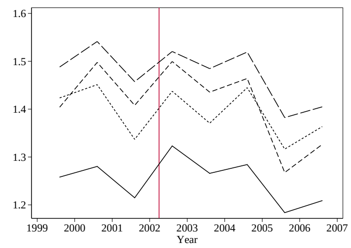
Notes: This table displays the same statistics as listed in Table 4.1, but local authorities are separated by predicted 2002 SSPR star ratings.



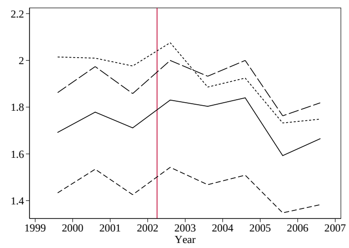
(a) % of Pensioners in Area



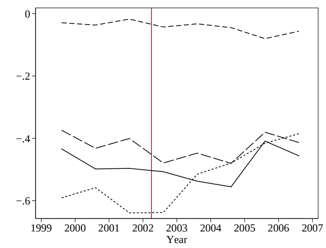
(b) % WA Benefit Claimants in Area



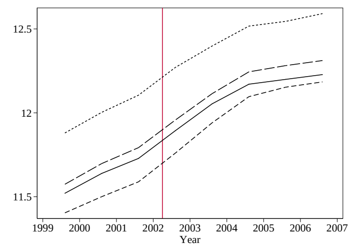
(c) In-Migration Rate (%)



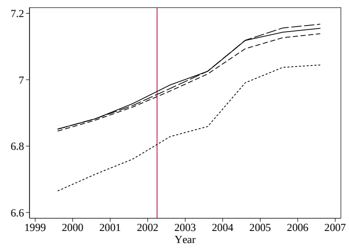
(d) Out-Migration Rate (%)



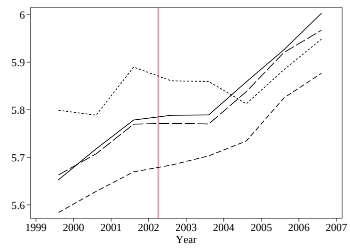
(e) Net Migration Rate (%)



(f) House Prices (log)



(g) Average Council Tax Bill (log)



(h) PSS Spending (per capita, log)

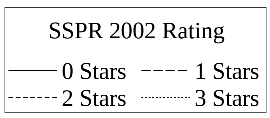


Figure 4A.2. Mean Trends, by 2002 SSPR rating

Notes: This figure displays the mean trends over time of all the main outcome variables for local authorities separated by 2002 SSPR rating. Subfigure (a) displays the number of pensioners living in a local authority (LA) as a percentage of all pensioners in the country. Subfigure (b) displays number of Working-Age (WA) benefit claimants living in a LA as a percentage of all the WA benefit claimants in the country. The in-migration rate in subfigure (c) is calculated as the number of individuals that migrated into an area by the end of a given year, divided by the size of the population in that area at the beginning of the year. The out-migration rate in subfigure (d) is calculated similarly for those who migrate out of an area. The net migration rate in subfigure (e) is the in-migration rate minus the out-migration rate.

Variables	Testing Eq. (4.3) (1) p-value	Testing Eq. (16) (2) p-value
Percentage of Pensioners in Area	0.412	0.205
Percentage of Working Age		
Benefit Claimants in Area	0.765	0.623
Low Income	0.157	0.864
Disabled	0.865	0.602
Percentage of JSA Claimants in Area	0.634	0.682
Movers (Population Aged 60+)		
In-Migration Rate	0.659	0.736
Out-Migration Rate	0.359	0.276
Net Migration Rate	0.262	0.327
Movers (Population Aged 60-74)		
In-Migration Rate	0.770	0.765
Out-Migration Rate	0.238	0.175
Net Migration Rate	0.295	0.168
Movers (Population Aged 75+)		
In-Migration Rate	0.453	0.552
Out-Migration Rate	0.373	0.397
Net Migration Rate	0.207	0.261
House prices (Log)	0.480	0.922

Notes: Probability of rejecting the null hypothesis of similar time path between local authorities (LAs) with different SSPR Ratings in 2002 in the pre-treatment period when the null is true. P-values related to the null hypothesis $H_0 : \theta_{99} = \theta_{00} = \theta_{01} = 0$ from a regression of equation 4.3.

Table 4A.3. P-values related to the null hypothesis $H_0 : \theta_{99} = \theta_{00} = \theta_{01} = 0$

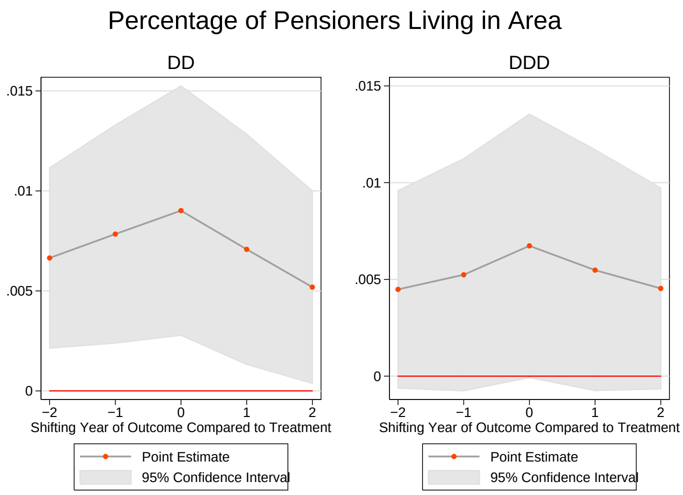


Figure 4A.3. Placebo Test

Notes: This figure displays the point estimates from regression equation (4.1) and the DDD estimates from regression equation (4.2). The dependent variable in the regression equations is the number of pensioners living in a local authority (LA) as a percentage of all pensioners in the country. The dependent variable is shifted by a number of years relative to the the treatment variable (the 2002 SPPR information release), as displayed in the x-axis, to act as a placebo test.

where Y_{it} it is the outcome variable of interest, τ_t is the set of year dummies, G_{it} a series of dummy variables reflecting the assignment of each particular star rating to an area, which take on a value of zero prior to 2002, and a value of one from 2002 onwards if the area received that particular star rating in that year. This differs from the main specification in the paper which only made use of the 2002 rating. Welsh councils have a value for zero star ratings for all time periods.

Table 4A.4 shows that the estimated effect of an area receiving a '3' star rating is associated with a 0.1 percentage points increase in the percentage of pensioners in those areas, relative to '0' star areas. As can be seen in the table, a rating of '2' stars, is estimated to be associated with about the same effect as '3' star areas relative to '0' star areas. Where the estimated effect of '1' stars versus '0' stars associated with a statistically insignificant 0.07 percentage point increase in the population of pensioners. This may point to the fact that people do not view '2' or '3' stars as that different, as is also found in the model estimates in section 4.6.5.3.

4A.6 Computational Details

The stationary equilibrium of the quantitative model is found using value function iteration on equation (4.12), with moving costs included, to find the relevant policy functions. The model is estimated with six areas, and I set the number of points in the support of the belief distribution to six. As each pensioner has beliefs about every area, this means the relevant state space contains

	% of Pensioners		In-Migration Rate (%)		Out-Migration Rate (%)		Net Migration Rate (%)	
Years included	Up to 2003	Up to 2005	Up to 2003	Up to 2005	Up to 2003	Up to 2005	Up to 2003	Up to 2005
Estimated effect of 3 stars versus 0 stars	0.0118*	0.0124*	-0.0279	-0.0302	-0.1120***	-0.0925**	0.0840	0.0623
	(0.0070)	(0.0069)	(0.0436)	(0.0290)	(0.0382)	(0.0428)	(0.0601)	(0.0531)
Estimated effect of 2 stars versus 0 stars	0.0126**	0.0117**	-0.0352	-0.0537**	-0.0523*	-0.0514*	0.0171	-0.0023
	(0.0053)	(0.0055)	(0.0356)	(0.0249)	(0.0306)	(0.0288)	(0.0431)	(0.0395)
Estimated effect of 1 stars versus 0 stars	0.0078	0.0072	-0.0557*	-0.0603**	-0.0341	-0.0207	-0.0216	-0.0396
	(0.0050)	(0.0050)	(0.0333)	(0.0240)	(0.0303)	(0.0288)	(0.0416)	(0.0415)
Estimated effect of 3 stars versus 2 stars	-0.0008	0.0007	0.0073	0.0235	-0.0596*	-0.0411	0.0669	0.0646
	(0.0056)	(0.0045)	(0.0336)	(0.0187)	(0.0319)	(0.0327)	(0.0522)	(0.0399)
Estimated effect of 2 stars versus 1 stars	0.0048*	0.0045*	0.0205	0.0065	-0.0182	-0.0307	0.0387	0.0373
	(0.0027)	(0.0024)	(0.0197)	(0.0155)	(0.0222)	(0.0220)	(0.0308)	(0.0266)
Estimated effect of 3 stars	0.0022	0.00273	-0.0449	0.0245	0.0706*	0.107**	-0.115*	-0.0825
	(0.0053)	(0.0048)	(0.0490)	(0.0501)	(0.0392)	(0.0440)	(0.0648)	(0.0568)
Local Authorities	170	170	170	170	170	170	170	170
Observations	850	1,212	850	1,212	850	1,212	850	1,212

Notes: The estimates in this table are from regression equation (17). The first dependent variable is the number of pensioners living in a local authority (LA) as a percentage of all pensioners in the country. The in-migration rate is calculated as the number of individuals that migrated into an area by the end of a given year, divided by the size of the population in that area at the beginning of the year. The out-migration rate is calculated similarly for those who migrate out of an area. The net migration rate is the in-migration rate minus the out-migration rate. All regressions include LA fixed effects, year effects, and full controls. Regressions are for 170 English and Welsh LAs. Control variables are those listed under *Area Characteristics* in Table 4.1. Standard errors are clustered at the LA level.
* $p < 0.1$, ** < 0.05, *** $p < 0.01$.

Table 4A.4. Alternative specification

46,656 elements for each area, and each of the six transition matrices contain $46,656^2$ elements. Increasing the number of points in the support of the belief distribution beyond six increases computational burden significantly and has very little effect on the estimates. Note that the discretization is used in order to calculate transition matrices and policy functions, but for the actual simulation there are no restrictions on beliefs (i.e. they do not lie on a grid). Therefore the discretization does not affect the calculation of per period utility, which is only influenced by decisions to migrate.

I nest the algorithm for solving the stationary equilibrium of the model into the following indirect inference estimation algorithm:

1. Choose an initial value of the parameter vector θ, θ_0, and set the objective function equal to infinity.

2. Solve for equilibrium of model using value function iteration.

3. Create $NS = 100$ data sets with my areas over 40 years with the information shock occurring at year 30. For each data set, compute the vector of moments $m^i(\theta)$

4. Compute $\left(\bar{m} - \frac{1}{NS}\sum_{i=1}^{NS} m^i(\theta)\right)' \left(\bar{m} - \frac{1}{NS}\sum_{i=1}^{NS} m^i(\theta)\right)$ and update θ_{i+1} based on θ_i and the value for the objective obtained;

5. Repeat steps 2 through 4 as many times as possible.

I use a simulated annealing algorithm for the minimization of the objective and start with a number of initial guesses to ensure that the global minimum is attained.

4A.6.1 Estimating Standard Errors The weighting matrix, W_N, in the indirect inference estimator, $\hat{\theta}$ (equation (4.15)) is the inverse of the variance-covariance matrix of the empirical moments, Ω. This is obtained by taking the variance-covariance matrices of the time average variables and regression coefficient from the DD regression, and constructing a block diagonal matrix. Then for a fixed NS, as the sample size tends to infinity, the indirect inference estimator is asymptotically normal with variance $var(\hat{\theta}) = \left(1 + \frac{1}{NS}\right) \left[\left(\frac{\partial m(\theta)}{\partial \theta}\right)' \Omega^{-1} \left(\frac{\partial m(\theta)}{\partial \theta}\right)\right]^{-1}$.