ABSTRACT

Title of Dissertation: Essays In Public Economics

John Iselin
Doctor of Philosophy, 2025

Dissertation Directed by: Professor Melissa Kearney
Department of Economics

This dissertation investigates behavioral responses to taxation across three key domains: labor supply, migration, and tax evasion. Chapter 1 introduces the three papers that make up this dissertation, and places them in the context of the field of public economics. Chapter 2 examines the labor supply effects of the California Earned Income Tax Credit (CalEITC), a state-level supplement to the federal EITC targeted at very-low-income workers with children. Using a triple-difference approach and American Community Survey data, I find that the CalEITC led to a shift from full-time to part-time work among single mothers, particularly those in larger households, consistent with a discrete model of labor supply. Chapter 3, coauthored with Karen Conway and Jonathan Rork, evaluates how state tax preferences for the elderly affect interstate migration. Leveraging tax return data and variation in state policy over two decades, we find that while overall income taxes slightly affect elderly migration, elderly-specific tax breaks do not, suggesting limited behavioral response to these targeted incentives. Chapter 4 coauthored with Daniel Reck, explores how assumptions about income misreporting affect estimates of top income inequality.

We propose a framework for expressing the different sets of assumptions recent authors in this space have made, which allows us to separate economic and methodological drivers of the differences between their approaches. We also review the current state of the economic literature, concluding that relying on audit data to measure misreporting likely understate inequality growth by failing to account for the rise of pass-through income among the very-wealthy.

Essays in Public Economics

by

John Iselin

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2025

Advisory Committee:
Professor Melissa Kearney, Chair/Advisor
Professor Katharine Abraham
Professor Judith K. Hellerstein
Professor Daniel Reck
Professor Neslihan Uler

© Copyright by
John Isaac Iselin
2025

Dedication

To Daniel, who already has my heart.

And to my mother, father, and sister, who made me and hold me to this day.

Acknowledgments

First and foremost, I would like to thank the professors at the University of Maryland who helped me through the last five years of study that led to this dissertation. I could not have written these papers without Melissa Kearney, my advisor, who provided both guidance and support throughout this process - thank you. You always seemed to know exactly what push I wanted or care I needed. To Daniel Reck, I am so glad you arrived when you did at Maryland; working with you – first as a research assistant and then as a co-author – has taught me so much about what collaboration in academia can accomplish. To Katharine Abraham, your wise words and advice always seem to land exactly where they were needed. And to Judith Hellerstein, you taught me so much about how to teach, mentor, and think in a way that represents the field of economics at its best; thank you. I would also like to thank Professor Neslihan Uler for taking the time to join my committee; I appreciate the time and care you took considering my work. In addition, there are numerous other professors at Maryland who have contributed to my growth as a scholar; in particular, I would thank Jessica Goldberg, Ethan Kaplan, Guido Kuersteiner, Luminita Stevens, Nolan Pope, John Shea, and Sergio Urzua for their advice, support, and encouragement over the last half-decade.

I would also like to thank a number of former professors and mentors of mine, without whom I could not have made it even to the starting line of this PhD. To begin with, I want to thank Kimberly Clausing, Noelwah Netusil, and Jonathan Rork at Reed College, who set me on

the path that led to this dissertation, and support me even to this day. I would also like to thank my former colleagues from the Tax Policy Center, especially Leonard Burman and Kim Rueben, who showed me how to put the tools of economics to use making the world a better place. I would also like to thank my professors at the Goldman School of Public Policy at the University of California at Berkeley – particularly Hilary Hoynes and Jesse Rothstein – who encouraged me to push myself to become a creator of research, not just a consumer. There have also been many who have provided me with mentorship and advice over the years, including Karen Conway at the University of New Hampshire, Martha Gimbel at the Budget Lab at Yale, Lucas Goodman at the Office of Tax Analysis, and Judy Scott-Clayton at Teachers College, Columbia University, to whom I am deeply grateful.

I could not have completed this PhD without support from those around me, both inside and outside the Maryland community. In fact, I feel myself very lucky that I am supported by four separate communities: my colleagues and friends at UMD, my former classmates from UC Berkeley, my former junior staff companions from the Council of Economic Advisors, and my wider community in DC and beyond. Within UMD, I am grateful for the support of Cameron Conrad, Taylor Landon, Lea Randell, and Maranna Yoder, who have always supported and inspired me. And a special, very dear, thanks goes to Erica Ryan, without whom this PhD would have felt like a much darker, much more lonely place. I am so happy that so many of my peers from Berkeley have stayed in my life: May, Tara, Emlyn (and Jax!), Olivia, and so many others, you remind me every day that policy is more than just papers, it is real and living and important. I was lucky enough to be able to take a year away from my PhD to serve as as a Staff Economics with President Biden's Council of Economic Advisers, and while there met a group of amazing, smart, devoted, and funny economists (even though some of them went on to be lawyers). I

learned so much from them that went into these papers, and they continue to enrich my life to this day. Thank you in particular to the other junior staff at CEA; to Ryan, to the other denizens of 483 (Anna, Sarah, and Natalie), to all of you, can't wait to see you again soon. Finally, I want to thank my amazing community outside of economics, both in DC and elsewhere. There are too many to list you all, but I want to call out a few in particular: Lydia, Liv, Nathaniel, Eric, Ben, William, Michael, Ryan, Andrew, Nick, Bryan, Noah, Malachi, Marcus, Madeline, Anthony, and so many others. Thank you for keeping me grounded, for celebrating and commiserating with me, and for being yourselves.

Table of Contents

Dedication ii

Acknowledgments iii

Table of Contents vi

List of Tables ix

List of Figures xi

Chapter 1: Introduction 1

Chapter 2: The Labor Supply Effects of the California Earned Income Tax Credit 6
- 2.1 Introduction . . . 7
- 2.2 Background . . . 13
- 2.3 Conceptual Framework . . . 17
- 2.4 Empirical Strategy and Data . . . 20
  - 2.4.1 Empirical Strategy . . . 20
  - 2.4.2 Data . . . 25
- 2.5 Effect on Employment . . . 30
  - 2.5.1 Trends . . . 30
  - 2.5.2 Main Effects . . . 32
  - 2.5.3 Robustness Checks . . . 37
  - 2.5.4 Decomposition by prior employment . . . 46
  - 2.5.5 Heterogeneity . . . 50
- 2.6 Effect of the CalEITC on Earned Income . . . 57
- 2.7 Elasticities . . . 63
- 2.8 Marginal Value of Public Funds . . . 66
- 2.9 Conclusion . . . 70

Chapter 3: Elderly State Income Tax Incentives and Migration: Evidence from Administrative Tax Data 74
- 3.1 Introduction . . . 75
- 3.2 Background . . . 77
  - 3.2.1 What do we know about inter-state (Near-) Elderly Migration? . . . 79
  - 3.2.2 State Income Tax Policies Targeting the Elderly . . . 81
- 3.3 Data . . . 84

3.3.1 Interstate Elderly Migration in the 2000s . . . 86
3.3.2 State Income Tax Benefits for the Elderly . . . 92
3.3.3 Are changes in Tax Policies and Migration Patterns Correlated? . . . 98
3.4 PPML Estimates . . . 100
3.4.1 Empirical Strategy . . . 100
3.4.2 Results . . . 104
3.4.3 Heterogeneity By Age . . . 110
3.5 Case Studies of Isolated Policy Changes . . . 112
3.5.1 Empirical Strategy . . . 113
3.5.2 Georgia . . . 115
3.5.3 Iowa . . . 116
3.6 Cost of Elderly-Specific Policies . . . 119
3.7 Conclusion . . . 120

Chapter 4: Comment on "Income Inequality in the United States: Using Tax Data to Measure Long-Term Trends" by Auten and Splinter 124
4.1 Introduction . . . 125
4.2 Framework . . . 126
4.3 Conceptual Argument . . . 131
4.4 From Fiscal to National Income . . . 135

Appendix A: Chapter 2 Appendices 138
A.1 Additional Tables and Figures . . . 138
A.2 Data . . . 170
A.2.1 Sample construction . . . 170
A.2.2 Imputation of qualifying children . . . 171
A.2.3 Predicted Earnings . . . 172
A.3 Earni . . . 173
A.4 Elasticity Calculations . . . 178
A.5 Inference . . . 181

Appendix B: Chapter 3 Appendices 189
B.1 Additional Tables and Figures . . . 189
B.2 Dataset Construction . . . 200
B.2.1 Flow-level data . . . 202
B.2.2 Income profiles . . . 203
B.2.3 Individual-level sample . . . 205
B.2.4 Cost dataset . . . 205
B.2.5 Construction of controls . . . 206

Appendix C: Chapter 4 Appendices 210
C.1 Additional Tables and Figures . . . 210
C.2 Additional Details on Auten and Splinter's Methodology . . . 215
C.3 Treatment of negative incomes and changes to the Tax System due to 1980s Reforms 222
C.3.1 Substitution Across Forms . . . 222

C.3.2 Partnership Loss Shelters in the 1980s 223
C.3.3 Sensitivity Analysis for Assumption 2.3 225

Bibliography 230

List of Tables

2.1 Descriptive statistics, single women aged 20-50 via the ACS . . . 29
2.2 Triple-difference estimates of the effect of the CalEITC on annual employment. . 34
2.3 Triple-difference estimates of the effect of the CalEITC on weekly employment via the CPS . . . 47
2.4 Triple-difference estimates of the effect of the CalEITC on annual employment, by count of qualifying children . . . 53
2.5 Triple-difference estimates of the effect of the CalEITC on annual employment, by count of adults in household . . . 54
2.6 Triple-difference estimates of the effect of the CalEITC on annual employment, by age of youngest qualifying child . . . 56
2.7 Triple-difference estimates, effect of the CalEITC on annual earnings . . . 59

3.1 Elderly preferences in state individual income tax systems . . . 83
3.2 Net in-migration rates by state, year, and age-groups. . . . 89
3.3 R2 values from gravity model. . . . 90
3.4 Top 30 per-capita elderly migration flows . . . 93
3.5 Effect of after-tax rates on inter-state migration by age, via Poisson Pseudo Maximum Likelihood . . . 107
3.6 Effect of elderly-specific after-tax rates on inter-state migration by age, via Poisson Pseudo Maximum Likelihood . . . 109
3.7 Effect of elderly-specific after-tax rates on inter-state migration across the elderly, via Poisson Pseudo Maximum Likelihood . . . 111
3.8 Elderly Benefit Cost Estimate for 2019 . . . 121

4.1 Estimates of the top 1 percent share, varying R^c_{1t} and E^c_{1t} . . . 136

A.1 State-level population and sample summary statistics . . . 158
A.2 Triple-difference balance test . . . 159
A.3 Triple-difference estimates, annual verses weekly measures of employment . . . 160
A.4 Triple-difference estimates of the effect of the CalEITC on annual employment, 2012-2019 . . . 161
A.5 Synthetic Difference-in-Difference Estimates, Weighted . . . 162
A.6 Triple-difference estimates of the effect of the CalEITC on annual employment, simulated instrument approach . . . 163

A.7 Placebo test using Longitudinal CPS, conditioning on prior-year employment . . 164
A.8 Triple-difference estimates of the effect of the CalEITC on weekly employment via the CPS, 2012 through 2017, including the first two quarters of 2015 165
A.9 Triple-difference estimates of the effect of the CalEITC on annual employment, by race and Hispanic origin 166
A.10 Triple-difference estimates of the effect of the CalEITC on annual employment, by disability status 167
A.11 Triple-difference estimates, effect of the CalEITC on annual earnings 168
A.12 Hours spent on different activities, by employment status, based on data from Bastian and Lochner (2022) 169
A.13 Triple-difference estimates of the effect of the CalEITC on employment, conditional on reporting wage income 175
A.14 Triple-difference estimates of the effect of the CalEITC on self-employment . . . 176
A.15 Triple-difference estimates of the effect of the CalEITC on annual employment, low-education sample, with different inference procedures 188

B.1 Sample composition by age and AGI 192
B.2 Effect of after-tax rates on inter-state migration by age without controls, via Poisson Pseudo Maximum Likelihood 193
B.3 Effect of after-tax rates on inter-state migration by age, 2 year leads, via Poisson Pseudo Maximum Likelihood 194
B.4 Effect of elderly and non-elderly after-tax rates on inter-state migration by age without controls, via Poisson Pseudo Maximum Likelihood 195
B.5 Effect of after-tax rates on inter-state migration by age, 2 year leads, via Poisson Pseudo Maximum Likelihood 196
B.6 Effect of elderly-specific after-tax rates on inter-state migration across the elderly, 2-year leads, via Poisson Pseudo Maximum Likelihood 197
B.7 Effect of after-tax rates on inter-state migration by among Elderly, dropping no-IIT states, via Poisson Pseudo Maximum Likelihood 198
B.8 Effect of after-tax rates on inter-state migration by among Elderly, using F1040 Zip Codes, via Poisson Pseudo Maximum Likelihood 199
B.9 Elderly Benefit Cost Estimate for 2019 201

C.1 Estimates of the top 1 percent share, varying R^c_{1t} and E^c_{1t}, 1962-2014 210

List of Figures

2.1 Federal and California EITC benefits schedule and budget constraint, TY 2016. 15
2.2 Histograms of workers . . . 16
2.3 Budget constraint for parent with 2 qualifying children (QC) in 2016 . . . 19
2.4 Full-time and part-time employment in the ACS. . . . 31
2.5 Event-study estimates of the effect of the CalEITC on annual employment. . . . 35
2.6 Effect of the CalEITC on employment, by annual weeks of work . . . 38
2.7 Specification curves over sample and controls . . . 41
2.8 Specification curves over control pool . . . 44
2.9 Effect of the CalEITC on employment, by prior-year employment using the longitudinal CPS. . . . 49
2.10 Changes in the earnings distribution over time, California verses control states 60
2.11 Distribution of Marginal Value of Public Funds estimates . . . 69
2.12 Implied fiscal spillovers of the CalEITC under different assumptions . . . 71

3.1 Inter-state migration by age and and AGI: 2000-2022 . . . 87
3.2 Average elderly tax benefit across AGI brackets for MFJ returns. . . . 95
3.3 Distribution of elderly tax benefit across AGI brackets for MFJ returns. . . . 97
3.4 Relationship between net in-migration and elderly tax benefits . . . 99
3.5 Georgia elderly tax benefits and net in-migration . . . 117
3.6 Iowa elderly tax benefits and net in-migration . . . 118

4.1 Assumptions about Misreporting Allocated to the Top 1 Percent and Their Effects 130
4.2 Dynamics of the Composition of Fiscal Income . . . 133

A.1 Federal and California EITC benefits schedule for 2015 and 2017. . . . 139
A.2 Benefits schedules for the federal EITC, CTC, and CalEITC for TY 2016 . . . 140
A.3 State-level trends in unemployment (2006-2019) . . . 141
A.4 Binding state minimum wages in control pool, 2010 through 2017 . . . 142
A.5 Benefits schedules for a single filer with 2 QC: 2018 and 2019 with TCJA CTC and Young Child Tax Credit . . . 143
A.6 Triple-difference estimate of the Effect of the CalEITC on the After-tax rate . . . 144
A.7 Annual full-time and part-time employment, via the CPS . . . 145
A.8 Quarterly full- and part-time employment in California via the CPS . . . 146
A.9 Event-study estimates of the effect of the CalEITC on annual employment . . . 147

A.10 Event-study estimates of the effect of the CalEITC on annual employment, 2012-2019 . . . 148
A.11 Effect of the CalEITC on Employment, by hours worked per week . . . 149
A.12 Per capita social safety net expenditures and caseloads . . . 150
A.13 Event-study estimates of the effect of the CalEITC on annual employment, using count of own children. . . . 151
A.14 Event-study estimates of the effect of the CalEITC on annual employment, by count of qualifying children . . . 152
A.15 Event-study estimates of the effect of the CalEITC on annual employment, by number of adults in the household . . . 153
A.16 Event-study estimates of the effect of the CalEITC on annual employment, by Race and Hispanic Origin . . . 154
A.17 Event-study estimates of the effect of the CalEITC on annual employment, by disability status . . . 155
A.18 Changes in the earnings distribution over time, alternative California distributions . . . 156
A.19 Effect of the CalEITC on employment by earnings bins. . . . 157
A.20 Event-study estimates of the effect of the CalEITC on employment restricted to wage workers . . . 177
A.21 Event-study estimates of the effect of the CalEITC on annual self-employment. 178

B.1 Average elderly tax benefit across AGI brackets for single returns. . . . 190
B.2 Estimated Effect of Elderly HoH on Property Tax Liabilities in Georgia . . . 191
B.3 Elderly Property Tax Differentials By Year and Method . . . 209

C.1 Evaluating Assumption 1: $R_{1t}^R = R_{1t}^c$ . . . 211
C.2 Reported non-wage income shares of total fiscal income . . . 212
C.3 Effect of misreporting on the top one percent share of fiscal income and pre-tax National Income . . . 213
C.4 Misreporting share of Income by Type, Assumption, and Place in Distribution 214
C.5 Sensitively to treatment of negatives incomes . . . 229

Chapter 1: Introduction

Public economics examines how government policies writ large impact the general welfare, turning the tools and frameworks of economics to the task of understanding and evaluating the ways governments interact with the economy. This dissertation focuses on one mechanism of government in particular – taxation – and considers how the revenue raising objectives of governments interacts with potential intended or unintended behavioral responses to these policies. In general, tax policies are designed to raise revenue for a range of government services while minimizing, with some notable exceptions, distortions of individual, household, or firm decision making and preserving or enhancing equity. The three papers below each consider a different potential distortion created by tax policy – namely changes in labor supply, migration, and tax evasion – and each attempting to contributes to the field's understanding of individual responses to policy changes and the measurement of such responses.

In Chapter 2, I study the effect of the California Earned Income Tax Credit on the labor supply decision of single mothers. There is a robust literature on the optimal taxation of lower-income individuals and families that has focused on how to balance the benefits of transfers to lower-income individuals with distortions of individual behavior that these transfers can cause (Mirrlees, 1971, Saez, 2002, Saez et al., 2012). The exact nature of individual-level reactions – notably whether individuals enter or exit the labor force or marginally adjust their hours of work –

has large effects on the overall welfare implications of transfers that operate through the tax code. In the United States, tax-based social safety net programs targeted at families with children, like the Earned Income Tax Credit (EITC) or the Child Tax Credit (CTC), are designed to encourage work, distorting labor supply in ways economists consider to be socially optimal. With few exceptions (for example, during the COVID-19 pandemic) these programs provided little benefit to families without any earned income, then phased in benefits with earnings up to a fixed level. The majority of the literature has found that this structure induces single mothers in particular to enter the labor-market (Eissa and Liebman, 1996, Meyer and Rosenbaum, 2001, Hoynes and Patel, 2018, Bastian and Jones, 2021, Schanzenbach and Strain, 2021). In this context, there is a concern that efforts to provide larger transfers to taxpayers lower in the income distribution could reduce labor supply, decreasing overall welfare. However, while there has been several recent papers examining the elimination of earnings tests (see Enriquez et al. (2023), Pac and Berger (2024), Ananat et al. (2024), Vivalt et al. (2024), Balakrishnan et al. (2024), Goldin et al. (2024)) less is known about how workers respond to efforts to adjust the phase-in rates of these programs, separate from changes to the overall level of benefit generosity. My first chapter studies the implementation, in 2015, of the California Earned Income Tax Credit (CalEITC), the state supplement to the federal tax credit, which provided a targeted transfer to very-low-income workers with children. I use its differential effect across women with and without children as quasi-experimental variation to understand the labor supply response of workers to changes in incentives at the very bottom of the earnings distribution. Using a triple-difference methodology, I finds a significant shift from full-time to part-time work, alongside an overall increase in part-time employment. This shift is consistent with labor supply models where workers select from discrete employment options based on financial incentives, highlighting how policy changes influence labor market

outcomes. Though there is some caution about extrapolating from earnings-measured derived from ACS data to any estimate of welfare, the overall results imply that altering the phase-in rate of work-based tax credits can cause single mothers, particularly in large, multi-adult households, to reduce their labor supply while not dropping out of the labor force altogether.

In Chapter 3, I, along with Professors Karen Conway and Jonathan Rork, study the effect of state-specific elderly tax preferences on inter-state migration. One common consideration among policymakers weighing changes to tax policy is whether higher (lower) taxes will induce residents to leave (stay) in the relevant area and repel (attract) new residents. These effects can change the overall level of revenue gained or lost due to a tax policies, impacting welfare via required changes to expenditures or taxes to make up the difference. Economic research has attempted to measure the responsiveness of individuals to changes in tax rates, with a considerable focus on particular groups like star scientists and inventors (Akcigit et al., 2016, Moretti and Wilson, 2017), athletes Kleven et al. (2013), and the very wealthy Young and Varner (2011), Kleven et al. (2014), Agrawal and Foremny (2019), Moretti and Wilson (2023). One group that has become particularly attractive to US policymakers are the elderly, due to their growing share of the population, their relative affluence Conway and Rork (2016), and the fact that they consume a bundle of government services (namely health care and social insurance) that is paid for primarily by the federal government (as opposed to younger residents, whose children consume educational services). This has led state governments to design a range of state individual income tax policies that lower taxes on elderly residents relative to that of their non-elderly counterparts. We use U.S. tax data and two decades of tax policy changes to study whether these policies induce elderly individuals to move between states (or stay in a given state). We find that overall individual income taxes do induce small changes in inter-state migration for the elderly, but elderly-specific

policies do not. It is unclear whether this is due to a lack of information about these policies, or if these policies provide small benefits (relative to total income) to those higher-income taxpayers who appear most responsive to tax changes.

In Chapter 4, I, along with Professor Daniel Reck, consider how the measurement of misreported income (or income that should have been reported on a tax return but was not) impacts the dynamics of income inequality over the last 60 years. There is a spirited debate over the degree to which the top 1 percent of the income distribution has increased since the 1980s, with Piketty et al. (2018) and Auten and Splinter (2024) finding diverging answers. Both sets of authors use the same underlying data – US tax records – and the same definition of income – National Income as defined by the U.S. Bureau of Economic Analysis – but each makes a different set of assumptions regarding how to allocate sources of income that are not observed on tax returns across the income distribution. One of the most important of these additions is misreported income, which makes up 3 and 7 percent of national income annually. Auten and Splinter (2024) use information from random audits to discipline their distributional estimates of misreporting, finding minor increases in the share of income held by the top one percent since the 1980s, while Piketty et al. (2018) rely on the observed distribution of reported income and find a more substantial increase in inequality. We propose a framework for expressing the different sets of assumptions made by both sets of authors. Doing so allows us to separate complex but ultimately minor points regarding their methodologies from the more important effects of underlying economic assumptions made by both sets of authors. We then evaluate the existing literature, finding that the methodology employed by Auten and Splinter (2024) does not account for the growing importance of pass-through entities since the 1980s. Given that misreporting from these entities is likely not captured by traditional audits, and that these corporate forms have been growing in importance at

the top of the income distribution, any methodology based on audit data will likely underestimate inequality. This work captures the importance the estimation of a particular undesirable form of behavior response – tax avoidance or evasion – on the estimation of income inequality in the United States.

Chapter 2: The Labor Supply Effects of the California Earned Income Tax Credit

Abstract[1]

I study the labor supply effects of the implementation of the California Earned Income Tax Credit (CalEITC) in 2015, which broke with the structure of the federal EITC by providing large transfers to very-low-income workers. Using a triple-difference approach, I find a 3 percentage-point increase in part-time employment a matching decline in full-time work among single mothers. This result, coupled with a observed decline in earned income, implies that the CalEITC induced a shift from full-time to part-time work. This response is consistent with a model of labor supply in which workers select between a discrete set of employment options at different points in the earnings distribution. This response is concentrated among women living in large, multi-adult households, pointing to an important role for intra-household bargaining in allowing women to respond to the credit. Accounting for behavioral effects and fiscal spillovers from the labor supply response on federal and state taxes, I estimate that the Marginal Value of Public Funds of the CalEITC is between 0.6 and 0.9.

[1]I would like to thank Melissa Kearney, Katharine Abraham, Judith Hellerstein, and Daniel Reck for their advising and support. I would also like to thank Sebastian Galiani, Jessica Goldberg, Ethan Kaplan, Guido Kuersteiner, Luminita Stevens, Nolan Pope, Sergio Urzua, and other seminar participants at the University of Maryland, as well as seminar participants at the 2024 National Tax Association Annual Conference, the Congressional Budget Office, and the Penn-Wharton Budget Model. Finally, I am grateful to Daniel Arias and Erica Ryan for their many insights.

2.1 Introduction

There is a large and growing literature demonstrating that providing resources to low-income households with children can both reduce child poverty in the short run and improve children's long-term outcomes across a range of measures (National Academies of Science, 2019, Page, 2024). Since the 1990s, two major tax-based safety net programs—the Earned Income Tax Credit (EITC) and, more recently, the Child Tax Credit (CTC)—have aimed to support low-income families with children while minimizing potential distortions to labor supply. To achieve this, benefits from these programs phase in with earned income, creating an incentive for non-working individuals to enter the labor market. As a result, non-workers and very low-income families often receive little to no support from these programs. Economists and policymakers have debated redesigning these programs to address this gap, proposing reforms ranging from unconditional cash transfers to eliminating earnings tests for low-income workers and adjusting parameters of existing tax credits. Understanding how these changes might affect labor supply is critical to assessing to estimating their effect on welfare. While the labor supply incentives of the federal EITC have been widely studied, less is known about how individuals respond to changes in the phase-in structure of these credits—a key lever in many proposed reforms.

In this paper, I provide new evidence on labor supply responses at the very bottom of the earnings distribution, using quasi-experimental variation from the introduction of the California Earned Income Tax Credit (CalEITC). Passed in 2015, the CalEITC was designed to deliver larger benefits to very-low-income, primarily part-time workers with children—those who would have received only modest benefits from the federal EITC (Rueben et al., 2017). The CalEITC nearly doubled the phase-in rate of the combined federal and state EITCs, enabling workers to

reach the maximum refundable credit at much lower earnings levels than under the federal credit alone, as shown in Figure 2.1a. This direct targeting of very-low-income workers has the potential to substantially improve household welfare: in 2015, a single parent with two children earning $6,925 could receive a combined EITC of $5,220—almost doubling their pre-tax income—with about half of the benefit coming from the CalEITC.[2] However, the policy may have altered the labor supply decisions of impacted households, and this effect both theoretically ambiguous and essential to any welfare analysis.

To assess these effects, I build on the empirical literature on EITC-induced labor supply by using a triple-difference approach. I compare employment outcomes for single women with and without children, inside and outside California, before and after the CalEITC's implementation. This strategy addresses a key limitation of traditional difference-in-differences models comparing parents and non-parents, which can be biased if these groups are subject to different economic shocks or policy trends (Kleven, 2024). By differencing out the trend between parents and non-parents in other states, I mitigate this concern, assuming that the parent–non-parent difference in California would have followed a similar path to the control states in the absence of the policy. I focus on the population of single women because this group is both the most exposed to the CalEITC given their distribution of earnings and because they have been shown to be the most responsive to labor supply incentives induced by tax changes by prior research.[3]

Using American Community Survey (ACS) data from 2012 to 2017, I find no change in overall employment following the introduction of the CalEITC. However, part-time work in-

[2] A single parent with two children in 2015, earning $6,925, would be eligible for a federal EITC of $2,862 and a CalEITC of $2,358, totaling $5,220.

[3] I considered the labor supply effects of the CalEITC on single men and married couples, but the results, particularly the pre-treatment employment trends, are noisy enough to rule out any determination regarding the effect of the policy on these populations.

creases by about 3 percentage points, while full-time work falls by a similar amount.[4] These offsetting changes suggest that the increase in part-time work is driven by individuals who would otherwise have worked full-time. I find suggestive evidence from longitudinal Current Population Survey (CPS) data that supports this conclusion. The results are robust to a variety of specifications and alternative approaches, including simulated instruments and Synthetic Difference-in-Differences (SDID) (Arkhangelsky et al., 2021). Using ACS data on weeks worked, I find no evidence that single mothers adjusted the number of weeks they worked per year, suggesting that responses occurred along the hours margin rather than the weeks margin. Importantly, these results reflect the short-term response to of women to the CalEITC, as the short duration of the post-treatment period prevents me from extending the results to the medium and long-term.

To explore mechanisms behind the shift from full-time to part-time work, I examine heterogeneity across household and demographic characteristics. The largest shift occurs among women with one or two qualifying children and those living in multi-adult households. These findings imply that household composition affects a woman's ability to adjust labor supply: more children may reduce flexibility, while additional adults may increase it. This highlights the potential importance of of intra-household bargaining over labor and home production when considering the labor supply responses of individuals to tax changes (Chiappori, 1992, Blundell et al., 2005). Notably, the labor supply response is not limited to mothers of very young children, suggesting that the shift to part-time work is not driven solely by parents facing childcare responsibilities.

Both the welfare effects of this transition and the appropriate model of labor supply to use

[4] As detailed in Section 2.4.2, I classify part-time and full-time work based on weekly hours: individuals working fewer than 35 hours are part-time, while those working 35 or more are full-time.

to evaluate it depend largely on the counterfactual income a full-time worker would have earned absent the CalEITC. Conceptually, the shift from full- to part-time work could reflect either a marginal reduction in hours (from the federal EITC kink-point to the new CalEITC kink-point) or a larger, discrete shift in labor supply—consistent with models in which individuals choose among a limited set of work options (Kosonen and Matikka, 2023). I present several pieces of suggestive evidence favoring the latter explanation, including a large reduction in average income and a shift downward in the distribution of earned income. However, I also caution that the ACS earnings data are insufficient to isolate precisely where in the earnings distribution a mother would have been employed but for the CalEITC. I highlight this as a direction for future research, alongside the need for better longitudinal earnings data that can track the transitions across the income distribution.

Using these results, I calculate a labor supply elasticity that captures the degree to which single mothers alter their labor supply in response to the CalEITC discrete jumps in earnings; a "mobility" elasticity (Saez, 2002, Kosonen and Matikka, 2023, Kreiner et al., 2015). Given the uncertainty regarding the counter-factual income of a full-time worker, I construct a range of estimates, based on several possible alternatives. I find that the mobility elasticity of between 1.0 and 1.7; which are large in the context of the prior federal EITC reforms. They are larger, in fact, than all but those estimated based on the 1986 EITC reform reform.[5] However, the group of women who respond to the CalEITC–single mothers in multi-adult households–might be uniquely able to respond to the incentives of the CalEITC, relative to prior reforms.

Finally, following Hendren and Sprung-Keyser (2020), I estimate the Marginal Value of

[5]As noted by Kleven (2024), the 1986 reform might be uniquely prone to larger elasticity estimates as it made only small changes to the tax rate faced by single mothers. This means that any small bias in the estimation of the effect of the reform could have large effects on the overall estimated elasticity.

Public Funds (MVPF) for the CalEITC—a measure of the cost to the government of providing one dollar of benefits, accounting for behavioral responses. My estimate incorporates fiscal spillovers to federal and state income taxes but excludes long-run impacts on children and other safety net programs due to data limitations (Bastian and Jones, 2021, Bastian and Michelmore, 2018). I estimate the MVPF of the CalEITC to be between 0.6 and 0.9—well below the MVPF of 3.2 to 4.2 estimated by Bastian and Jones (2021) for the federal EITC. This gap is driven primarily by reduced earnings from the shift to part-time work and interactions with the federal tax and transfer system. These results highlight that the net fiscal cost of state-level programs can depend critically on how they interact with federal policies.

This paper presents one of the first evaluations of the CalEITC, a major anti-poverty initiative in the nation's largest state. Previous work, such as Rueben et al. (2017), has offered theoretical and descriptive analysis of the program, and Neumark and Li (2024) find no change in overall employment—a result they attribute to the CalEITC's narrow targeting and California's high minimum wage. My paper is the first to examine a broader range of labor supply responses and to provide an MVPF estimate for the program. It also connects to research on CalEITC take-up (Iselin et al., 2023) and on the California Young Child Tax Credit, which evolved from the CalEITC (Goldin et al., 2024, Unrath, 2024).

I also contribute to the emerging literature on discrete labor supply responses to tax changes. Most prior work has relied on continuous labor supply models, often incorporating optimization frictions, and has typically focused on estimating elasticities near the intensive-margin kink (Saez, 2010, Chetty et al., 2011, Chetty, 2012, Kleven and Schultz, 2014). However, as Kosonen and Matikka (2023) notes, discrete labor supply models can account for behavioral responses away from marginal incentives. My findings provide suggestive evidence that workers above the

CalEITC kink adjusted their behavior in a manner consistent with such models.

Finally, this paper adds to the literature on labor supply responses—especially among women—to changes in EITC design. It complements earlier work documenting extensive-margin increases in labor force participation in response to EITC expansions (Eissa and Liebman, 1996, Meyer and Rosenbaum, 2001, Hoynes and Patel, 2018, Bastian and Jones, 2021, Schanzenbach and Strain, 2021, Bastian, 2020), and builds on more recent evidence of muted participation effects (Kleven, 2024). Intensive-margin effects remain elusive, particularly outside the self-employed, where some observed responses are likely driven by reporting behavior rather than real changes (Garin et al., 2021). This paper also relates to recent research on earnings-benefit decoupling, such as the fully refundable 2021 CTC (Enriquez et al., 2023, Pac and Berger, 2024, Ananat et al., 2024) and California's elimination of the earnings test for the YCTC in 2022 (Goldin et al., 2024, Unrath, 2024), both of which show little impact on employment. Related experimental evidence on unconditional cash transfers has found moderate labor supply reductions (Vivalt et al., 2024, Balakrishnan et al., 2024). These findings are highly relevant, as they suggest alternative strategies to modifying phase-in parameters for reaching low-income families.

The remainder of the paper is structured as follows. Section 2.2 describes the structure of the federal and California EITCs and other relevant policies. Section 2.3 outlines two conceptual labor supply models. Section 2.4 presents the data and empirical strategy. Section 2.5 shows the main employment results. Section 2.6 examines the effect of the CalEITC on earnings. Section 2.7 calculates labor supply elasticities implied by my labor supply and earnings results. Section 2.8 estimates the MVPF of the CalEITC. Finally, Section 2.9 concludes.

2.2 Background

The federal EITC is currently one of the largest means-tested cash transfer program in the United States, providing $64 billion in refundable tax credits to 31 million low-income workers in 2022 (IRS, 2023). The federal EITC operates as a refundable tax credit, meaning that eligible taxpayers can receive a refund from the government even if their credit exceeds the total amount of taxes they owe. A taxpayer's EITC eligibility and benefit amount is primarily a function of their earned income, with the parameters of that function varying based on the number of children and marital status of the taxpayer. The generosity of the credit increases substantially with the number of children a tax filer can claim on their return, with childless workers eligible for only very small credit amounts. Importantly, a dependent on a tax return must meet several additional criteria—based on their age, residency, and relationship to the taxpayer—to be counted as a *qualifying child* for the purpose of EITC eligibility and benefits calculations. The federal EITC, as depicted by the solid green (or black, when the combined EITC schedule and the federal Schedule overlap) line in Figure 2.1a, consists of three regions.[6] First, in the "phase-in region" a taxpayer's EITC credit amount increases at a fixed rate per dollar earned until a specified earning level is reached. Second, in the "plateau region" the EITC credit amount holds constant until a second earning level is reached. Third, in the "phase-out region" the credit amount decreases at a fixed rate per dollar earned until it reaches $0.

In June 2015, California introduced the California Earned Income Tax Credit (CalEITC). Unlike most of the 31 other states and the District of Columbia that also have state-level EITC supplements, California's credit is not simply a fixed percent of the federal EITC. Like the federal

[6] I report the same figures for 2015 and 2017 in Figure A.1.

EITC, the generosity of CalEITC is a function of a tax unit's earnings and number of qualifying children, with very low credit amounts for tax units without kids. However, the CalEITC uses a different function to calculate benefits amounts. The credit, depicted as the dashed green line in Figure 2.1a, does not include a plateau region, and phases in and out at a much lower level of earned income. For 2015 and 2016, individuals were only eligible for the CalEITC if they had annual earnings that placed them on the phase-in region of the Federal EITC.[7] The result was the joint federal-state EITC schedule (the black line), which has a faster phase-in rate and a longer pseudo-plateau region than the federal credit, meaning taxpayers could receive very close to their maximum potential EITC payment at a lower level of annual earnings.[8]

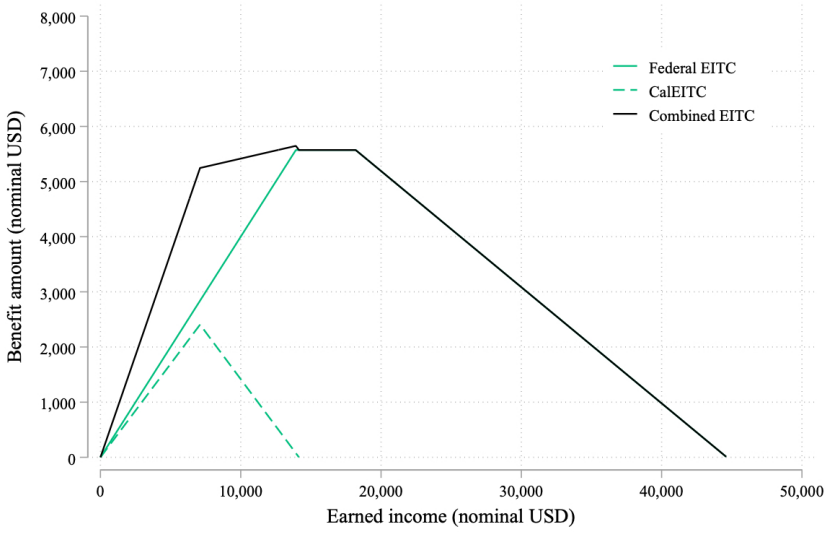
The CalEITC has undergone a number of changes since its introduction (Davis and White, 2019). First, prior to 2017, self-employment income was not counted as earnings for the purpose of calculating the CalEITC, leaving only wage and salary income, a deviation from the federal EITC. This was changed in 2017, when California amended the definition of earnings to bring it in line with the federal EITC. Second, also in 2017, tax filers earning between $10,000 and $22,300 (for a family with two QC) became eligible for the CalEITC (see Figure A.1b). However, these newly-eligible tax units could only claim very small credit amounts, leaving unchanged the fundamental focus of the CalEITC on very-low-income workers.

As shown in Figure 2.2a and noted by Rueben et al. (2017), workers eligible for the largest CalEITC benefits are likely to be working part-time. Super-imposing the federal and Califor-

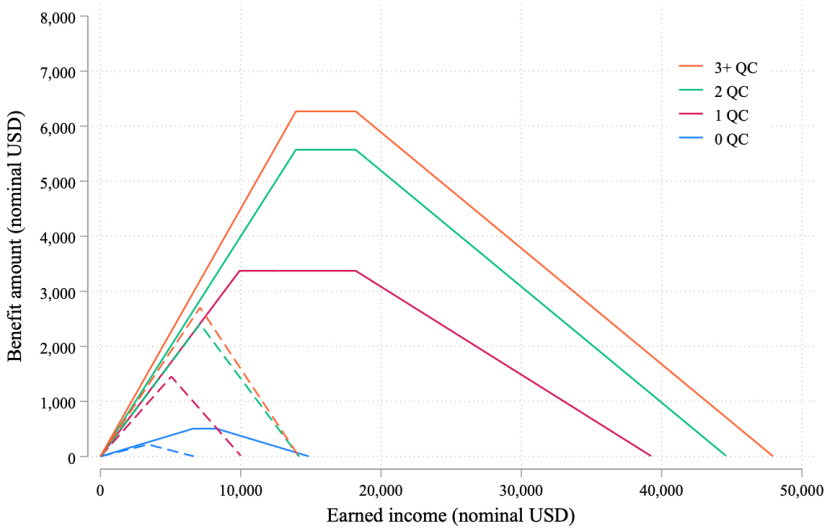
[7] An additional requirement is that a tax unit's adjusted gross income was below the maximum earnings level at which an individual could still claim any CalEITC benefits. In other words, even if a tax unit's earnings were $7,000, if they reported enough other non-earned income, they would no longer be eligible for the CalEITC. The federal EITC has the same rule.

[8] The CTC during this period (as shown in Figure A.2) provided a maximum credit of $1,000 per child. When considering the combined refundable tax credit schedule (federal plus state EITC and CTC), the effect of the CalEITC is less about an expansion of the plateau region, and rather reflects the introduction of a kink into the budget constraint for very-low levels of earnings (Chetty et al., 2011).

Figure 2.1: Federal and California EITC benefits schedule and budget constraint, TY 2016.



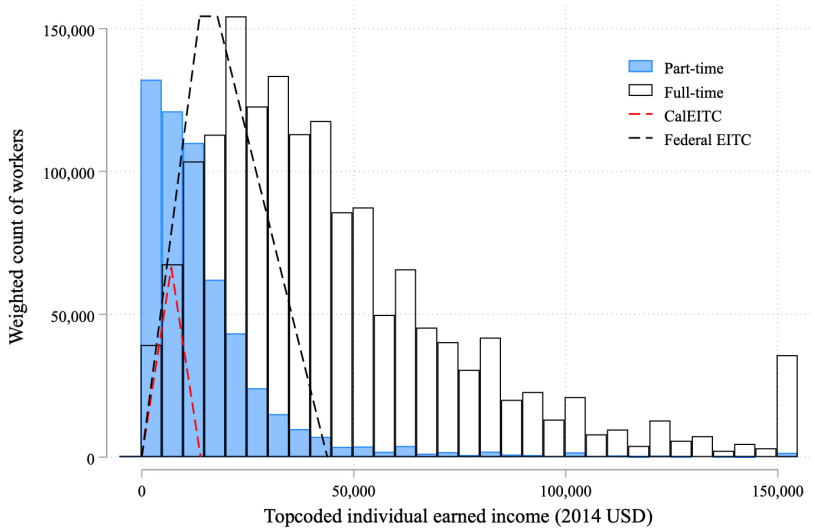
(a) EITC schedule for a parent with 2 qualifying children



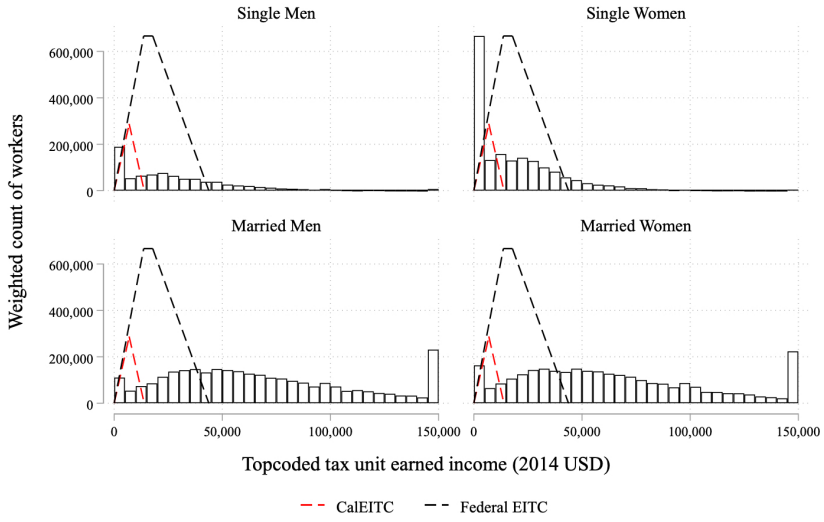
(b) EITC schedule by count of qualifying children

Note: The parameters of the federal EITC used to create this figure were obtained via the Urban-Brookings Tax Policy Center, while the parameters of the California EITC schedule were backed-out of the Earned Income Tax Credit tables from California Form FTB 2514. The benefit schedules shown assume a single filer whose adjusted gross income is solely composed of wage income. The colors across both figures reflect the number of qualifying children (blue for zero, red for one, green for two, and orange for three), with black representing the combined EITC schedule for a single parent with two qualifying children. Dashed colored lines represent the CalEITC, and solid colored lines represent the federal EITC.

Figure 2.2: Histograms of workers



(a) Single women workers in 2014, by full- and part-time status



(b) Workers with qualifying children (2012-2014), by marital status and sex

Note: These figure are derived from American Community Survey (ACS) data from 2014 for Figure 2.2a and 2012-2014 for Figure 2.2b. Both series are limited to sample of workers, aged 20-50, who are US citizens, not in the armed services, not in school, and not potential qualifying children as define in Appendix A.2.1. The first panel focuses on single women, while the second on married and single men and women with qualifying children. All histograms plot the weighted count of workers in each earned income bin, separately by full-time and part-time workers in the first panel and by marital status and sex in the second. Earned income has been top-coded at $150,000, and defined at the individual level for the first panel and at the tax-unit level in the second. Full time work is defined as working at least 35 hours a week, and part time work is defined as working under 35 hours a week. The red and black dashed lines reflects the CalEITC and federal EITC benefit amount respectively, for a single mother with 2 qualifying children as of 2015 (in the case of the CalEITC) and 2014 (in the case of the federal EITC).

nia EITC schedules for a single parent with two qualifying children, the eligibility range for the CalEITC overlaps with the area of maximum density of part-time workers. Further, as the California minimum wage increased starting in 2014, fewer full-time workers would be expected to have earnings in the CalEITC-eligible range without substantial changes in labor supply across other margins (for example, weeks worked). This was very much the intention of California policymakers, who were attempting to allocate additional resources to lower-income, part time workers who receive relatively small benefit amounts from federal tax credits. Further, as shown in Figure 2.2b, the group with the most workers earning in the CalEITC-eligible range are single mothers: while there are single men and married couples who are eligible for this credit, based on pre-treatment earnings distributions, single mothers are the most likely to benefit from this credit, absent any change in labor supply response.

2.3 Conceptual Framework

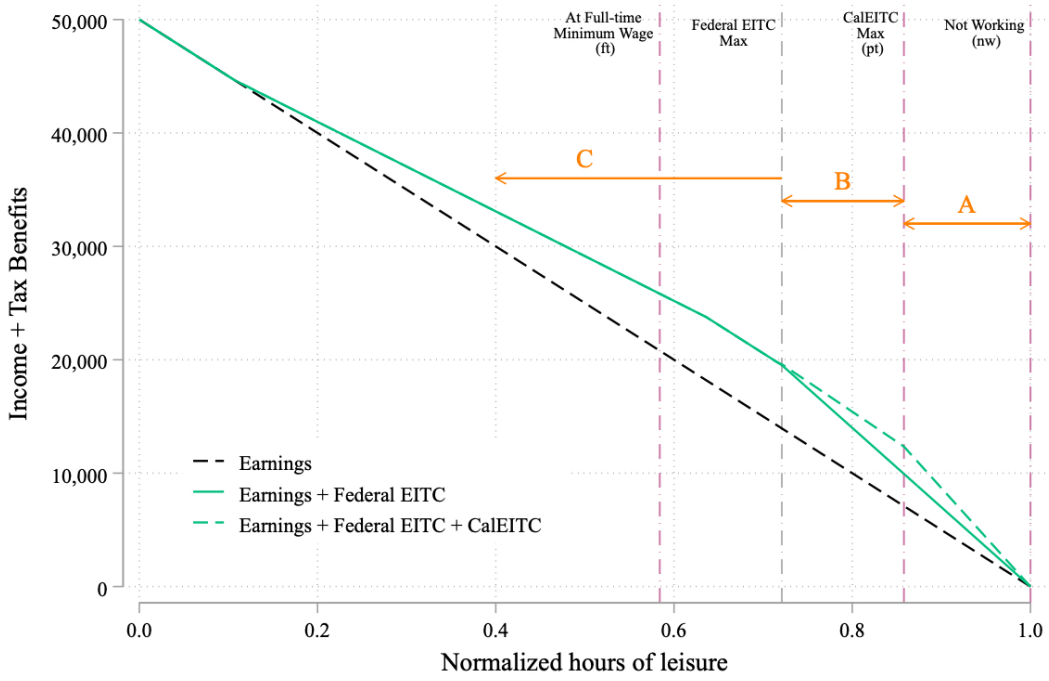
In this section, I consider the potential margins by which taxpayers might alter their labor supply. I argue that the combined federal and state EITC schedule is the most relevant for gauging the responses of California filers, as both credits have nearly identical eligibility rules and because it was exceedingly rare for a tax filers to claim the CalEITC without claiming the federal EITC as well (Iselin et al., 2023, Davis and White, 2019). As shown in Figure 2.3, the introduction of the CalEITC shifted out the budget constraint of an eligible tax filer, increasing their net-of-tax income at higher relative levels of leisure.

Following closely the structure and argument of Kosonen and Matikka (2023), consider an agent facing the budget constraint depicted in Figure 2.3, with an utility function $u(c, z)$, where c

represents consumption (or after-tax income), and z is gross income (Saez et al., 2012, Feldstein, 1999). Utility is increasing in c and decreasing in z, with the latter reflecting a number of factors, including lost utility from leisure, value from home production, and dis-utility from work. In a standard, continuous model of labor supply, individuals can freely choose $\{c, z\}$ by adjusting effort across a number of dimensions. Therefore, the budget constraint faced by a tax filer is $c = z * (1 - \tau) + R$, where $1 - \tau$ is the net-of-tax rate, and R is virtual income. Mapping this to the simple case where workers can only adjust labor supply across hours worked results in $c = z * (1 - \tau) + R = w * (1 - l) * (1 - \tau) + R$ where w is the exogenous wage, l is a normalized measure of leisure, as plotted in Figure 2.3.

Assuming an agent can freely select the optimal level of consumption and leisure to maximize their utility, I would expect one of two responses to the implementation of the CalEITC: an extensive-margin response where agents enter the labor market at some point in regions A or B as noted on Figure 2.3, or an intensive margin change where agents marginally adjust their effort. For the former, this would mean agents who prior to the reform chose not to work ($l = 1$), might be induced to increase their hours worked (reduce l), now that the return to a unit of labor has increased sufficiently. For the latter, agents who, prior to the reform, would have been located in region A, on the lower half of the federal EITC phase-in range, would face an income effect pushing them to work more (decrease l) and an income effect to work less (increase l), while agents on the second part of the phase-region B–have a theoretically unambiguous incentive to work less (increase l). Past work has generally found that agents, particularly single mothers, respond along the extensive rather than intensive margin (Schanzenbach and Strain, 2021, Kleven, 2024, Saez, 2010, Chetty et al., 2011, 2013, Bastani and Selin, 2014, Nichols and Rothstein, 2015). Importantly, given a continuous model of labor supply and reasonable assumptions regarding the

Figure 2.3: Budget constraint for parent with 2 qualifying children (QC) in 2016



Note: The parameters of the federal EITC used to create this figure were obtained via the Urban-Brookings Tax Policy Center, while the parameters of the California EITC schedule were backed-out of the Earned Income Tax Credit tables within the California Form FTB 2514. Normalized leisure represents the proportion of potential working time spent on leisure. For each choice of leisure (the x-axis), I calculate the annual pre-tax income for a potential minimum wage worker (the dashed black line), as well as the resulting after-tax income including only the federal EITC (the green line) and the federal and California EITC (the dashed green line). The right-most purple dashed line represents the total amount of normalized leisure associated with not working. The next dashed line to the left represents the level of normalized leisure that a worker would have if they had earnings that maximized their CalEITC benefit amount. The next (slightly faded) dashed line represents the level of normalized leisure that maximizes the federal benefit amount. It is faded to represent the potential inaccessibility of a given point on the earnings distribution. The left-most dashed line represents the amount of leisure of a full-time worker, working 52 weeks out of the year, at the binding state minimum wage ($10 in 2016). The regions A (between the third and fourth dashed lines from the left) represents the region where I might expect, if the substitution effect outweighs the income effect, a marginal increase in labor supply. The region B (between the second and third dashed lines) is the region where, given that the income and substitution effects work in the same direction, I might expect a decline in labor supply. Region C (everything to the left of the second dashed line) is where I would expect no response to the CalEITC in a model of continuous labor supply, but where I might expect a decline if taxpayers are choosing between a few discrete points in the earnings distribution.

utility function I would not expect agents in region C of Figure 2.3 (the whole region to the left of the lighter dashed line) to respond to the CalEITC.

Now, consider a discrete model of labor supply, in which, following Saez (2002) and Koso-

nen and Matikka (2023), agents choose between J discrete levels of earnings denoted $j = 1, ..., J$. Each location j produces utility equal to $u(c_j, z_j)$, where $c_j = z_j * T_j + R$. Define T_j as the average tax rate at that level of gross income. There are any number of reasons why a worker might have a limited set of available choices in terms of hours worked; see, for example, the case discussed by Chetty et al. (2011) where firms offer a constrained set of possible hours due to existing production technology. Assuming, as above, that agents only respond across hours worked, each discrete point j can be associated with a specific l_j. In a simple example, assume that agents can choose to not work (nw), work part-time (pt), and work full-time (ft), where $1 = l_{nw} > l_{pt} > l_{ft}$. I denote these three points as darker dashed lines on Figure 2.3. Agents then choose the level l_j to maximize their utility. Importantly, while this setup still allows for the extensive-margin response to the CalEITC seen in the continuous model, I would not expect to see any marginal adjustments in line with the intensive-margin responses discussed above. However, this model does rationalizes larger drops in labor supply from l_{ft} to l_{pt}.

2.4 Empirical Strategy and Data

2.4.1 Empirical Strategy

My main specification is a cross-sectional triple-difference design, comparing the difference between single women with and without qualifying children (difference 1), before and after the implementation of the CalEITC in 2015 (difference 2), in California and in the group of untreated states (difference 3). This estimate recovers a causal estimate of the CalEITC on employment under the assumption that without the CalEITC, the difference between the employment of single women with and without qualifying children in California would have had the same

trend as the matching difference in the pool of untreated states (Olden and Møen, 2022). This approach takes advantage of two untreated groups: those individuals without qualifying children, and individuals with qualifying children outside of California. The former has been used repeatedly by the literature, given that the federal EITC provides much smaller benefit amounts to tax units without qualifying children (Eissa and Liebman, 1996, Bastian and Jones, 2021, Bastian, 2020, Eissa and Hoynes, 2004, Hoynes, 2019, Jones and Ziliak, 2022). The same approach is appropriate in the case of the CalEITC because of the similarly small credit amounts the state credit provides to non-parents (see Figure 2.1b). This also controls for economic and demographic trends within California that have the same impacts on parents and non-parents, as well as policy changes that affect both groups equally. Comparing parents to non-parents does have potential issues; namely, that having a child exposes parents to both the direct effect of the child as well as the effects of other programs whose benefit schedules are also a function of family size (Kleven, 2024). If there are concurrent policy changes or economic shocks that differentially impacts parents and non-parents, then the simple difference-in-difference could be biased. My triple difference approach is designed—in large part—to account for this critique.[9] The threats to the validity of this strategy come from either shocks within a given state that differentially impact parents and non-parents, or nation-wide shocks with heterogeneous impacts across parental status and state (i.e. a shock that impacts mothers differently from single women without kids, but does so in different ways in different states). I discuss both in detail in Section 2.5.3.

In constructing my pool of control states, I want to exclude any states that experienced shocks that did not impact individuals in California, and include states that best reflect the coun-

[9]Concurrently with this paper, Neumark and Li (2024) also implement a similar triple-difference approach, though their pool of control states is marginally different from the ones I consider here.

terfactual trends employment in California would have experienced without the CalEITC. First, I exclude Alaska, to avoid complications stemming from Alaska's Permanent Fund Dividend (PFD), a state-level program that makes annual payments (not conditioned on work) to Alaska residents, which was highly variable during this period. Second, I exclude states that experienced changes in their own state EITC policies between 2008 and 2019.[10] These restrictions, also outlined in Table A.1, create my primary sample of 28 states, including California. I also perform several robustness checks, where I alter the pool of untreated states, discussed further in Section 2.5.3. It is worth noting that, while many states have implemented versions of the CTC in recent years, none were implemented prior to 2019, meaning there are not other major tax credits operating at the state-level that might alter the labor supply incentives of low-income residents during this period. I estimate the effects of the CalEITC on several different employment measures using Equation 2.1:

$$\begin{aligned} y_{ist} =& \alpha + \beta \text{Treated}_{ist} + X_{it}\eta + (U_{st} \times \text{QC}_{it})\theta + (\text{MW}_{st} \times \text{QC}_{it})\mu \\ &+ \delta_t + \gamma_s + \text{QC}_{it} + (\text{QC} \times \delta)_{it} + (\text{QC} \times \gamma)_{ist} + (\delta \times \gamma)_{st} + \epsilon_{it} \end{aligned} \tag{2.1}$$

In this equation, y_{ist} is the outcome variable for person i in state s at time t, QC_{it} is a fixed effect for the number of qualifying children (capped at 3) assigned to person i, δ_t is a year fixed effect, and γ_s is a state fixed effect. The final three terms are fixed-effects for the interactions between pairs of state, year, and number of qualifying children fixed effects. The coefficient of

[10]These include Colorado, Connecticut, Hawaii, Illinois, Iowa, Kansas, Louisiana, Maine, Maryland, Massachusetts, Michigan, Minnesota, Montana, New Jersey, New Mexico, Ohio, Oregon, Rhode Island, South Carolina, Vermont, and Wisconsin.

interest is β, with treatment defined as:

$$\text{Treated}_{ist} = \mathbf{1}\left[(\text{QC}_{it} > 0) \;\;\&\;\; (\text{t} \geq 2015) \;\;\&\;\; (\text{s} = \text{CA})\right] \tag{2.2}$$

In other words, an individual is treated if she have at least one qualifying child, during or after 2015, and are in California. I cluster standard errors at the state-level.[11]

The vector X_{it} is a set of demographic controls, including measures of age, education, race, Hispanic origin, the minimum age among the qualifying children, and the number of adults in the household. These are designed to control for both subtle changes in the demographic makeup of states across my sample, as shown in Table A.2, and for other policies whose benefit generosity is a function of family composition. I also control for differential effects of labor market trends across different types of families by controlling for state unemployment interacted with the number of qualifying children ($U_{st} \times \text{QC}_{it}$). Prior work by Hoynes and Patel (2018) and Schanzenbach and Strain (2021), among others, have highlighted that women with and without children may be differentially impacted by business cycle dynamics. In my setting, the recovery from the Great Recession was relatively slow in California, as shown in Figure A.3, which implies this control is all the more important. Finally, I also control for any potential differential impact of state-level minimum wage policies by interacting the binding state minimum wage with the number of qualifying children ($\text{MW}_{st} \times \text{QC}_{it}$). As shown by Figure A.4, 12 states in

[11] Recent work has shown that inference using the cluster-robust variance estimator (CRVE) is complicated in cases where there are relatively few clusters, heterogeneity across cluster size, and–importantly–only one treated cluster. In these cases, CRVE-based inference risks over rejecting the null hypothesis that this is no effect. In Appendix A.5, I review and implement a range of alternative inference procedures, noting that many of them face the opposite problem, a risk of severe under-rejection, in cases where the issues above interact. However, the specification with my full set of controls is statistically significant at the five percent level when using the approach proposed by Ferman and Pinto (2019), and at either the five or ten percent level when using wild-cluster bootstrap proposed by Cameron et al. (2008).

my sample had state minimum wages in excess of the national level, and all of those increased their minimum wage at least once between 2010 and 2017. Importantly, California began a series of hikes in 2014, right before the implementation of the CalEITC. This is only a threat to my identification strategy if there is a differential impact of minimum wages on employment outcomes across single women with and without children. The recent literature examining this question has found mixed results. Cengiz et al. (2022), using a cross-state event-study design, found no differential effect of minimum wages on single mothers. Godøy et al. (2024), using a distributional difference-in-difference approach, found that single mothers respond to minimum wage increases by increasing their full-time employment via jobs paying close to the minimum wage, while other groups have employment effects that are indistinguishable from 0. In this case, without controlling for differential effects of minimum wages across the number of qualifying children, my estimates of the effect of the CalEITC on overall or full-time employment would be biased upwards.

My empirical strategy can be extended to an event-study specification by interacting a modified version of the treatment indicator as defined in Equation (2.2) with each year, as shown in Equation (2.3), allowing me to both consider pre-treatment trends and observe the heterogeneity of the treatment effect over time.

$$\begin{aligned} y_{ist} =& \alpha + \sum_{j \neq 2014} \beta_j \left(\text{Year}_{j=t} \times \mathbf{1}\left[(\text{QC}_{it} > 0) \ \ \& \ \ (\text{s} = \text{CA}) \right] \right) \\ &+ X_{it}\eta + (U_{st} \times \text{QC}_{it})\theta + (\text{MW}_{st} \times \text{QC}_{it})\mu \\ &+ \delta_t + \gamma_s + \text{QC}_{it} + (\text{QC} \times \delta)_{it} + (\text{QC} \times \gamma)_{ist} + (\delta \times \gamma)_{st} + \epsilon_{it} \end{aligned} \tag{2.3}$$

My main specification is limited to 2012 through 2017 to avoid potential impacts the re-

covery from the Great Recession, the expansion of the CTC in 2018 as part of the Tax Cuts and Jobs Act (TCJA), and the creation of the Young Child Tax Credit (YCTC) in California in 2019, depicted in Figure A.5.

In a "first-stage" of sorts, I estimate a version of Equation (2.3) where the outcome variable is the after-tax rate (or $1 - \tau$, where τ is the average tax rate) faced by women after I assign each person in my sample the level of earned income that would have maximized their CalEITC receipt. The estimates, reported in Figure A.6, shows that, prior to the CalEITC, the difference in tax rates between low-income parents and non-parents were very similar between California and in the pool of control states. After the implementation of the CalEITC, the after-tax rate increased dramatically, reflecting a unique change in tax rates for very-low-income tax filers with children in California.

2.4.2 Data

My primary data source is the American Community Survey (ACS) from 2010 through 2017 (and occasionally through 2019), downloaded via IPUMS (Ruggles et al., 2023). The ACS is an annual random survey of approximately 1 in every 100 individuals in the U.S., designed by the US Census Bureau to provide up-to-date demographic, social, economic, and housing data annually, rather than once-a-decade as was previously available via the census long-form questionnaire. While ACS data is available starting in 2000, the design of the survey was fully implemented starting in 2006. The ACS is a monthly survey, but results from each month in the year are pooled by the US census bureau to produce publicly available annual statistics and micro data. The ACS asks questions both about employment during the last week and during the

last year. Critically, the reference period for these annual measures of employment (or income) is the 52 weeks prior to conducting the survey, not the calendar or tax year within which the survey is taken. The timing of the ACS reference period raises the concern that an individual's reported employment in the 2015 ACS represents actual employment in 2014. I address this issue in Section 2.5.1 by using monthly data via the Current Population Survey (CPS) to show that the change in employment—particularly part-time employment—began after the enactment of the CalEITC in June 2015.

My main outcome variables are indicators for whether an individual worked during the last 12 months, and whether they were working full- or part-time. I follow the Bureau of Labor Statistics (BLS) definition of full- versus part-time work, where full-time is defined as "usually" working 35 or more hours a week, and part-time as working fewer than 35 hours. ACS respondents are asked to report the usual numbers of hours they worked per week, conditional on working in a given week, providing the required information to replicate this definition. I complement this analysis with employment results by finer bins of hours worked, as well as measures of weeks worked in a given year.[12] I also separately report overall employment results, though not results broken down by full or part-time status, for a measure of employment when the reference period was the week prior to the interview.[13] The ACS also asks several questions about income that an individual earned during the 12 months prior the interview, including wages, salaries,

[12]I can further look at whether the employment effects are driven by increases in self-employment. While the ACS does not ask a specific question regarding self-employment during the prior year, I can determine if a worker was - at any point - self-employed in the last year by observing whether or not they reported any self-employment income (and in some cases I restrict focus to those reporting self-employment income in excess of $1,000 annually). These results are reported in Appendix A.3. I find an increase in self-employment during 2015-2016, likely representing either an initial misunderstanding of the CalEITC eligibility rules or an effort to supplement part-time wage work with self-employment income.

[13]The ACS does not ask a question regarding hours of work from the week prior to the interview, restricting information gleaned from this reference period.

self-employment, and investment income.

My main sample consists of single women aged 20-50 without a college degree, who meet a number of criteria. They must be US citizens, not in school, not in group quarters or in the armed services, and cannot potentially be qualifying children themselves. The logic behind each of these restrictions is discussed in Appendix A.2.1, but is intended to isolate a group of likely eligible individuals without any major confounding characteristics. I limit my sample to single women without a college degree to focus on the sample of lower-income women who are likely to be impacted by these incentives without conditioning directly on earned income (Hoynes and Patel, 2018, Schanzenbach and Strain, 2021). In order to identify which adults have qualifying children, I utilize information on household composition and demographics to identify both who is likely to be a qualifying child, and which adult in a household is likely to claim that child, a process that is described in detail in Appendix A.2.2. This abstracts from recent work that has shown some evidence of strategic claiming of children within households (Jones and O'Hara, 2016, Splinter et al., 2017). Given that this likely results in some mis-assignment of treatment status, I expect some attenuation of my treatment effect, meaning that my results should be viewed as a lower-bound of the estimated effect of the CalEITC on the labor supply response.

I present descriptive statistics for my sample in Table 2.1, across the different dimensions of the triple-difference. Importantly, California and the pool of control states differ significantly in the racial and ethnic make-up of the sample, as well as the rate of self-employment, and levels of earned income. These differences alone are not problematic per-se: rather, the concern would be if any of these differences threaten the assumptions of my triple-difference approach. In Table A.2, I report the result of a balance test, where I estimate a version of Equation (2.1) with no demographic or unemployment controls, using a range of demographic characteristics as

dependent variables. I find statistically significant effects for several demographic variables, most notably across educational attainment and the share of the population that is Hispanic versus Non-Hispanic White. While these differences are small enough to be unlikely to have a large effect on my results, I include a range of demographic controls in my main specification, and employ, as a robustness check, a synthetic difference-in-difference (SDID) approach that weights the pool of control units to account for such differential trends. My results remain substantively unchanged with the inclusion or exclusion of controls, or the use of SDID.

I also utilize monthly data from 2012 through 2017 from the monthly Current Population Survey (CPS), accessed via IPUMS (Flood et al., 2022). The CPS provides information from a range of very detailed questions on weekly employment, as well as information on household composition, geographic location, family relationships, and demographics. However, the sample size is notably smaller that that of the ACS.[14] I utilize the CPS to supplement my findings by taking advantage of the longitudinal feature of the survey.[15] Using the longitudinal CPS, I can separately identify the effect of the CalEITC on those who were employed full-time, part time, or not employed 12 months prior.

[14]For 2015, for example, I observe 12,069 single women in California that are in my main sample, verses 4,886 in the pooled monthly CPS, 1,647 in the longitudinal sample, and 621 in the CPS ASEC, which is an annual supplement to the CPS that contains additional information on income and other economic characteristics. It is also worth noting that one of the major benefits of the CPS (monthly or ASEC) for prior research has been its coverage of the pre-2006 period, allowing for research on all the major changes in federal EITC policy. However, the ACS was fully implemented for the period of interest regarding the CalEITC. Researchers studying more recent periods of EITC reforms have also used the ACS (Wilson, 2023).

[15]CPS interviews are conducted in a rotating pattern, with households interviewed for four months, not interviewed for the next eight months, then interviewed again for four months. This structure means that for four months per person (the last four months in the sample), it is possible to link an individual to their employment characteristics from one year prior, via their first interview. To adjust for attrition, I construct a set of longitudinal weights using a weight-calibration procedure referred to as iterative proportional fitting (raking), implemented via the *ipfraking* package in Stata (Newton et al., 2014). The weights are designed to allow the longitudinal sample for a given month (with the reference month being end of the 12-month period) to be representative of the CPS as a whole in that month.

Table 2.1: Descriptive statistics, single women aged 20-50 via the ACS

	CA - 2012-2014		CA - 2015-2017		Controls - 2012-2014		Controls - 2015-2017	
Variable	No QC	1+ QC	No QC	1+ QC	No QC	1+ QC	No QC	1+ QC
Age	31.9	34.4	31.4	34.6	33.1	33.8	32.4	33.8
	(9.9)	(8.4)	(9.4)	(8.2)	(10.3)	(8.2)	(10.0)	(8.0)
Adults in HH	2.9	2.2	3.0	2.3	2.5	1.9	2.5	1.9
	(1.4)	(1.2)	(1.5)	(1.3)	(1.2)	(1.0)	(1.2)	(1.0)
Hispanic	40.0	49.5	44.7	51.7	13.7	17.4	15.4	18.6
	(49.0)	(50.0)	(49.7)	(50.0)	(34.4)	(37.9)	(36.1)	(38.9)
Non-Hispanic White	38.0	27.3	33.6	25.2	58.5	47.5	56.1	45.8
	(48.5)	(44.6)	(47.2)	(43.4)	(49.3)	(49.9)	(49.6)	(49.8)
Non-Hispanic Black	10.4	14.1	10.2	13.5	23.0	30.7	23.0	30.7
	(30.5)	(34.8)	(30.3)	(34.1)	(42.1)	(46.1)	(42.1)	(46.1)
Non-Hispanic Other	11.6	9.1	11.5	9.6	4.9	4.4	5.4	4.9
	(32.0)	(28.8)	(31.9)	(29.5)	(21.5)	(20.6)	(22.6)	(21.7)
No High School	13.9	18.9	13.3	16.7	14.8	17.7	13.7	16.4
	(34.6)	(39.1)	(34.0)	(37.3)	(35.6)	(38.2)	(34.4)	(37.0)
High School Grad	41.7	38.9	43.4	40.9	47.7	44.3	48.8	44.9
	(49.3)	(48.8)	(49.6)	(49.2)	(49.9)	(49.7)	(50.0)	(49.7)
Some College	44.4	42.2	43.2	42.4	37.4	38.0	37.5	38.7
	(49.7)	(49.4)	(49.5)	(49.4)	(48.4)	(48.5)	(48.4)	(48.7)
Employed last year	72.5	71.6	75.1	75.6	72.3	74.0	74.2	76.5
	(44.6)	(45.1)	(43.2)	(43.0)	(44.7)	(43.9)	(43.8)	(42.4)
Any self-employment	4.5	3.8	4.3	4.7	3.0	3.4	3.1	3.5
income last year	(20.7)	(19.2)	(20.2)	(21.1)	(17.1)	(18.1)	(17.4)	(18.5)
Employed full time	47.9	51.0	51.3	53.5	51.2	52.8	53.2	56.2
last year	(50.0)	(50.0)	(50.0)	(49.9)	(50.0)	(49.9)	(49.9)	(49.6)
Employed part time	24.6	20.6	23.8	22.0	21.2	21.2	20.9	20.4
last year	(43.1)	(40.4)	(42.6)	(41.4)	(40.9)	(40.8)	(40.7)	(40.3)
Average hours worked	35.2	36.0	35.6	36.2	36.2	36.3	36.5	36.7
per week last year	(11.1)	(10.4)	(10.6)	(10.1)	(10.7)	(10.1)	(10.4)	(9.9)
Weeks worked last	44.3	44.3	45.0	45.3	44.4	44.0	45.0	44.6
year	(13.3)	(13.4)	(12.6)	(12.4)	(13.2)	(13.6)	(12.7)	(12.9)
Earned income (2019	27,327.6	29,703.7	28,617.8	30,405.6	24,473.3	24,082.6	24,913.9	25,145.4
USD)	(27,901.1)	(27,787.2)	(29,869.5)	(28,623.5)	(23,333.2)	(21,386.0)	(23,456.0)	(23,087.3)
Observations	21,572	14,409	22,387	13,432	104,665	81,491	105,504	77,712

Note: This table is derived from the American Community Survey (ACS) from 2012-2017, limited to the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, without a college degree, and not potential qualifying children as define in Appendix A.2.2. The sample is decomposed by state (California verses controls), period (2012-2014 and 2015-2017), and presence of qualifying children (QC). The pool of control states is limited to states without EITC policy changes between 2008 and 2019, and excludes Alaska, Hawaii, and DC. Summary statistics are weighted using individual-level survey weights. In this table, age, count of adults in the household (Adults in HH), average hours worked per week last year, weeks worked per year, and earned income are all reported as means, wile the rest of the variables are reported as percentages (ex. the mean of a variable equal to 100 if the given individual in the sample is Hispanic).

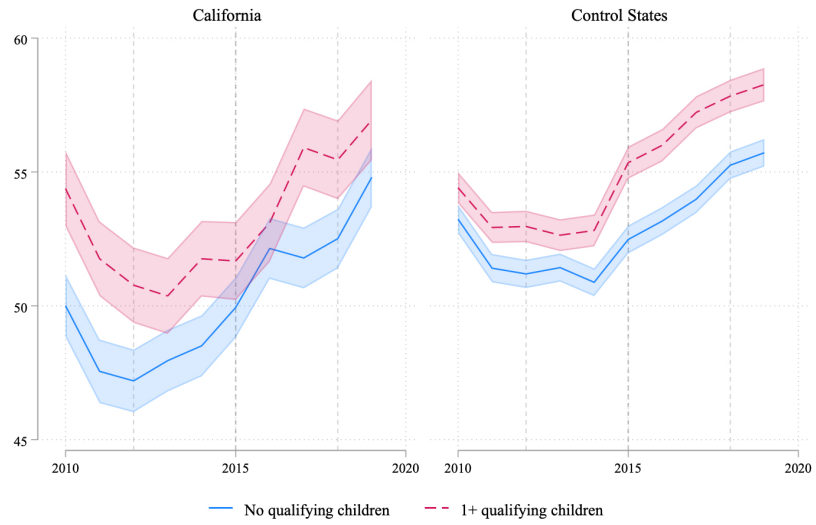
2.5 Effect on Employment

2.5.1 Trends

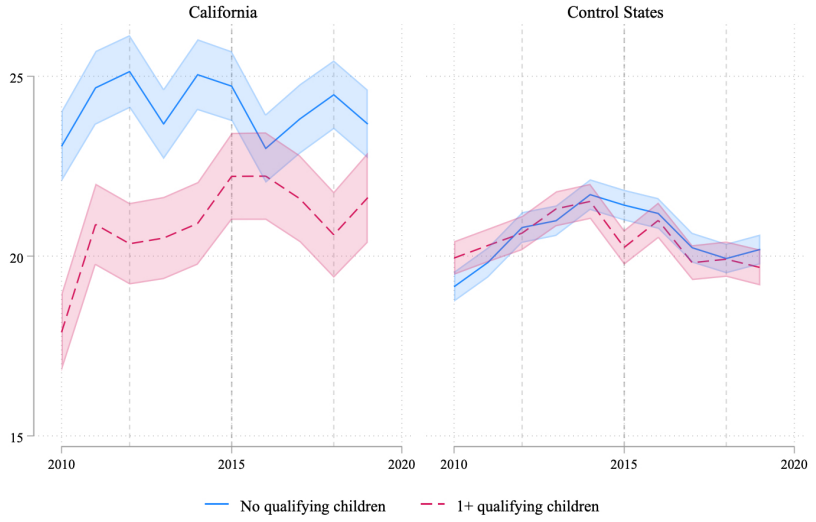
First, in Figure 2.4 I report the trends in the percent of the population of single women without a college degree employed full-time and part-time, separately for four groups: those with and without qualifying children, and in California versus in the rest of the United States. I plot the trends from 2010 through 2019, with the implementation of the CalEITC indicated by the dashed line in 2015, and the 2018 implementation of the TCJA denoted by the lighter dashed line further to the right. Across all samples, I see matching trends in full- and part-time work coming out of the Great Recession, reflecting the fact that during the pre-treatment period of my sample, the entirety of the United States was experiencing a slow recovery that featured a high-level of involuntary part-time work that had not yet returned to its pre-recession level (Kudlyak, 2019). For the purpose of my analysis, this is not problematic unless the differing trend in the recovery of single women in California with and without qualifying children would have been different from the same difference in the pool of control states without the CalEITC. For the majority of the pre-treatment period, starting in 2012, as denoted by the first light dashed line. As a robustness check (included in Figure 2.7c) I extend my pre-treatment period to 2010-2014, which does not change my results.

I can also directly observe the jump in part-time work in 2015, where part-time work among single women without qualifying children declines (as it does in the rest of the United States) while part-time work increases for single women with qualifying children. This trend holds until 2018, when the trend reverses for one year before returning to a higher level in 2019. The

Figure 2.4: Full-time and part-time employment in the ACS.



(a) Full-time employment



(b) Part-time employment

Note:Data for this figure are derived from the American Community Survey, collapsed to produced annual averages by the presence of qualifying children, using individual-survey weights. The sample is limited to single women, aged 20-50, who are US citizens, not in the armed services, not in school, without college degrees, and not potential qualifying children as define in Appendix A.2.2. Full- and part-time employment is defined based on usual hours worked per week last year, with full-time work defined as working 35 or more hours per week, and part time as working fewer than 35 hours. Grey bars indicate recessions, with dashed lines indicating the beginning (2012) and end (2018) of the sample period, as well as the beginning of treatment (2015).

effect is less obvious for full-time work: it appears that any relative decrease in full-time work for single mothers in California stems from a slower recovery, relative to the other groups in the sample. However, by 2017 the recovery among single mothers in California has re-emerged. I see similar (though nosier) trends when the same figures are generated using employment records via the monthly CPS (measuring weekly employment rather than annual employment), as shown in Figure A.7. Further, by re-creating the same figure at the quarterly, rather than annual, level, (see Figure A.8) I show that workers appear to begin to respond to the CalEITC after the implementation of the CalEITC in June 2015. Similarly (though less dramatically), the slower increase in full-time work begins after the implementation of the CalEITC, ameliorating concerns that the ACS reference period is leading to these estimates capturing some effect from prior to the implementation of the policy.

2.5.2 Main Effects

Table 2.2 presents my main results, where I estimate the effect of the CalEITC on employment using Equation (2.1), run on my preferred sample of single women without a college degree. Column 1 shows the results of the basic triple-difference without any controls, Column 2 adds in the set of demographic controls, Column 3 adds in the controls for state unemployment interacted with the number of QC a women could claim, and Column 4 adds the controls for the binding state minimum wage interacted with the count of QC. The first panel reports the results where the dependent variable is an indicator equal to 100 (for scaling purposes) if the woman was ever employed in the prior year. Across all four specifications, the effect is small, and statistically insignificant.[16] The second and third panels show the results when the dependent variables are

[16]The results are nearly identical when the dependent variable is employment last week (see Table A.3).

indicators for full- or part-time work: here the results are larger and statistically significant across all specifications. Full time work decreases by between 2.1 and 3.9 percentage points, while part-time employment increases by between 2.7 and 3.2 percentage points. Across all these models, the drop in full-time and increase in part-time work are statistically indistinguishable from each other in terms of magnitude, and imply a shift of between 45,000 and 65,000 single mothers on net from full-time to part-time work.[17]

I report the event-study estimates of the effect of the CalEITC on employment, as estimated by Equation (2.3), in Figure 2.5. These estimates include controls for demographics, unemployment, and minimum wages (i.e. are the analogous estimate to column 4 in Table 2.2).[18] The results are consistent with those shown in Table 2.2, with no effect on overall employment but a sharp drop in full-time work and a similarly sharp increase in part-time work. The response by single women on both the full-time and part-time margins begins immediately in 2015. Interestingly, the economic and statistical significance of the increase in part time work is consistent throughout the post-treatment period, while the decline in full-time work in 2017 is only about half the magnitude of the effect observed in 2016, in keeping with the overall, un-adjusted trends seen in Section 2.5.1. This speaks to the fact that while the shift from full-time to part-time work is immediate and robust, it reflects the short-term effect of the CalEITC. While I do estimate effects over the 2015-2019 period (see Table A.4 and Figure A.10), these estimates are potentially confounded by the expansion of the CTC as part of the TCJA and the implementation of the YCTC. Over this period, the magnitude of the labor supply response is smaller, implying either

[17]The increase in the size of the full-time effect moving from column 3 to 4 matches the expected direction of bias given the results of Godøy et al. (2024): if single mothers become more likely to enter the labor force relative to non-mothers given a higher minimum wage, then prior to controlling for such an effect, any confounding of the CalEITC implementation with the 2014 minimum wage would bias my estimates upwards (towards 0), as it would be capturing both a negative effect of the CalEITC and a positive effect of the minimum wage change.

[18]The full set of event-study estimates with alternative sets of controls are shown in Figure A.9.

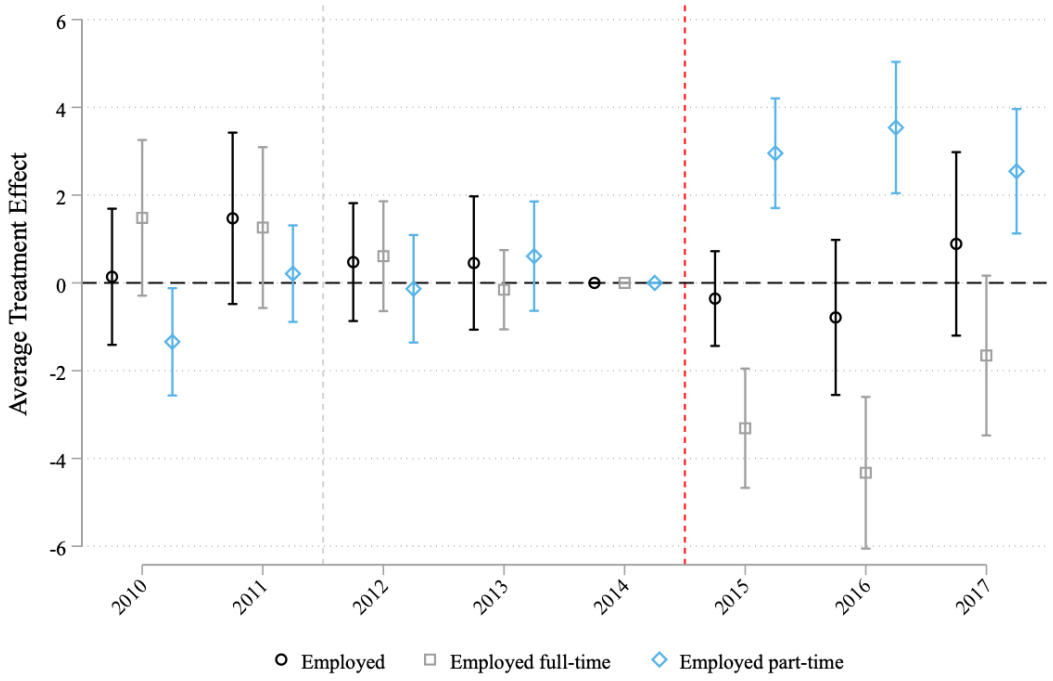
Table 2.2: Triple-difference estimates of the effect of the CalEITC on annual employment.

	(1) Basic Triple Difference	(2) Add Demographic Controls	(3) Add Unemployment Controls	(4) Add Minimum Wage Controls
Panel A: Employed in the last 12-months				
ATE	0.7*	0.3	0.1	-0.7
	(0.4)	(0.4)	(0.5)	(0.5)
Observations	441,172	441,172	441,172	441,172
Adj. R-Square	0.009	0.076	0.076	0.076
Treated group mean in pre-period	71.6	71.6	71.6	71.6
Implied employment effect	11,718	5,209	1,646	-11,737
Panel B: Employed full-time in the last 12-months				
ATE	-2.1***	-2.4***	-2.9***	-3.9***
	(0.4)	(0.4)	(0.6)	(0.6)
Observations	441,172	441,172	441,172	441,172
Adj. R-Square	0.008	0.068	0.068	0.068
Treated group mean in pre-period	51.0	51.0	51.0	51.0
Implied employment effect	-35,710	-41,262	-49,447	-66,628
Panel C: Employed part-time in the last 12-months				
ATE	2.8***	2.7***	3.0***	3.2***
	(0.4)	(0.4)	(0.6)	(0.6)
Observations	441,172	441,172	441,172	441,172
Adj. R-Square	0.003	0.028	0.028	0.028
Treated group mean in pre-period	20.6	20.6	20.6	20.6
Implied employment effect	47,428	46,472	51,093	54,890
Triple-Difference	Yes	Yes	Yes	Yes
Add Demographic Controls	No	Yes	Yes	Yes
Add Unemployment Controls	No	No	Yes	Yes
Add Minimum Wage Controls	No	No	No	Yes

Note: This table reports the results of Equation (2.1), estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, who do not have a college degree, and who are not potential qualifying children (QC) as define in Appendix A.2.2, via the American Community Survey (ACS) from 2012-2017. Panels A, B, and C vary only in the dependent variables, which are indicators for any employment, full-time employment, and part-time employment at any point in the last 12-months, respectively. Full time work is defined as working at least 35 hours a week, and part time work is defined as working under 35 hours a week. In columns 2, 3, and 4, demographic controls include indicators for 5-year age bins, race and Hispanic origin groups, educational attainment, bins for the age of the individual's youngest qualifying child, and the number of adults in an individual's household. Column 3 adds controls for state unemployment interacted with the number of QC, and Column 4 adds controls for the binding state minimum wage interacted with the number of QC. Standard errors clustered at the state level. * 0.10 ** 0.05 *** 0.01

that the shift in labor supply was a temporary phenomena, or that other changes related to the TCJA, the YCTC, or overall economic conditions mean that my triple-difference approach can no longer isolate the effect of the CalEITC.

Figure 2.5: Event-study estimates of the effect of the CalEITC on annual employment.



Note: This figure reports the results of Equation (2.3), estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, not potential qualifying children as define in Appendix A.2.2, and who do not have a college degree, via the American Community Survey from 2010-2017. The light dashed line represents the beginning of the pre-treatment period for my preferred estimates. Each of the event-studies has a different dependent variable, defined as an indicator equal to 100 (for scaling purposes) if an individual is employed (using the relevant measure) in a given year. This means the point estimates can be interpreted as the percentage-point effect of the CalEITC on employment (however defined). Full time work is defined as working at least 35 hours a week, and part time work is defined as working under 35 hours a week. Each model is run with the full set of demographic controls, as well as controls for state unemployment and state minimum wages interacted with the number of qualifying children. Standard errors are clustered at the state-level.

As discussed above, there is clear evidence of a increase in part-time work and a decrease in full-time work, at least in the initial years following the introduction of the CalEITC. I examine this transition more closely by decomposing the shift in hours-worked using finer bins of hours-worked to rule out that I am observing what is actually a very small shift in hours worked over the

threshold. In Figure A.11, I decompose the overall employment effect by bins of hours worked per week. I find evidence for reductions in employment among single women who work between 40 and 60 hours a week, and evidence for increases in employment between 20 and 34 hours. While it is difficult to determine the exact size of the net reduction in hours worked, these results leave open the possibility of very large reductions in average hours. Interestingly, the point in the hours distribution where I observe the largest increase in employment, between 30 and 34 hours per week, would result in an annual income of at least $13,500 for a worker employed for most of the year, above the level that would maximize receipt of the CalEITC.[19] This result (and a similar result with earned income discussed in Section 2.6), could be rationalized by a model where, due to adjustment or information frictions, individuals are unable to target their annual earnings appropriately to maximize their CalEITC receipt (Chetty et al., 2011). Alternatively, it could be explained by a salience effect: the CalEITC increases the salience of refundable tax credits in general, leading to some pool of tax units shifting to a level of earnings that maximizes the federal EITC and CTC. Finally, it could be a feature of noisy ACS hours data: there could be some error given the nature of self-reported earnings data or the timing of reference period. Nevertheless, there is a clear shift downwards in the number of hours worked for single women with children.

Hours of work per week is only one of several potential margins of response to the CalEITC. As discussed in Wilson (2020), many single women with a high school degree or less are less likely to make a single extensive margin decision; rather, they make several such decisions over the course of the year, entering or exiting the labor force multiple times. Therefore, I estimate the

[19] Namely, a worker employed 50 weeks a year, 30 hours a week would make $13,500 a year at the 2015 minimum wage in California.

effect of the CalEITC on employment separately by binned weeks worked per year, observable in the ACS.[20] First, as shown in Figure 2.6a, I see no change in overall employment across the distribution of weeks worked per year. In Figure 2.6b and Figure 2.6c, I observe that the entirety of the full- and part-time effects occur among those working 50 to 52 weeks a year. This implies that workers are responding to the CalEITC by adjusting the hours they work per week, rather than the weeks they work per year, consistent with a set of resource-constrained individuals, for whom going a prolonged period without work would be costly.

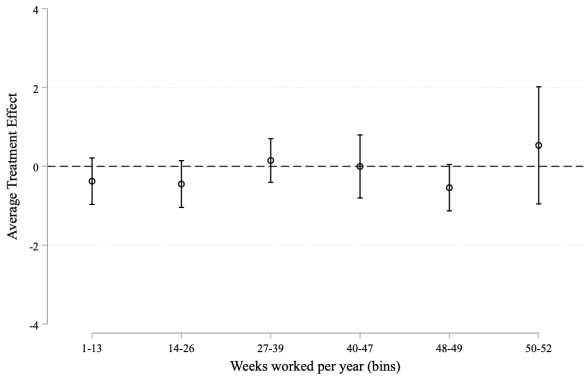
Taken together, these results show that, after the implementation of the CalEITC, full-time work fell by between 2.1 and 3.9 percentage points, while part-time work increased by between 2.7 and 3.2 percentage points. Looking over time, women responded to the policy immediately, in 2015, and the increase in part-time work persisted through 2017, while the decrease in full-time work had faded by the final year of the sample.

2.5.3 Robustness Checks

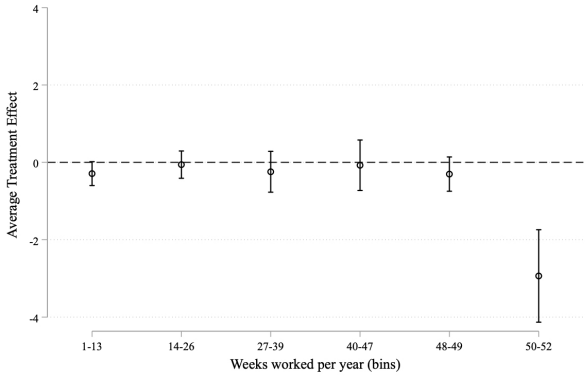
I next consider a range of robustness checks for the results presented above. As previously mentioned, the triple-differences approach requires that differences in trends across the two dimensions would have remained constant in the absence of the CalEITC. If the pool of control states does not approximate the counterfactual trend in employment differences between single women with and without qualifying children in California, this assumption may be violated. To account for this, I utilize the Synthetic Difference-in-Difference (SDID) approach proposed by Arkhangelsky et al. (2021). Building on both the difference-in-difference literature and the Syn-

[20] Specifically, I estimate Equation (2.1), where the dependent variable is equal to 1 if the women is employed (however defined) and reports working a specific number of weeks in the prior year.

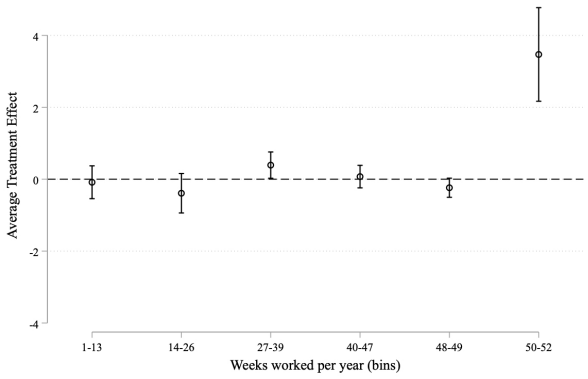
Figure 2.6: Effect of the CalEITC on employment, by annual weeks of work



(a) Employment by weeks worked per year



(b) Full-time employment by weeks worked per year



(c) Part-time employment by weeks worked per year

Note: This figure reports the results of Equation (2.1), estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, not potential qualifying children as define in Appendix A.2.2, and who do not have a college degree, via the American Community Survey from 2012-2017. The dependent variable is an indicator equal to 100 (for y-axis scaling purposes) if the individual is employed (however defined) and reports working a given number of weeks (based on ACS-reported bins of weeks worked) per year. Each model is run with the full set of demographic controls, as well as controls for state unemployment and state minimum wage interacted with the number of qualifying children. Standard errors are clustered at the state-level.

thetic Control Method, SDID uses the pool of controls units to construct a reasonable synthetic version of California that has parallel trends in the periods prior to treatment. I create a panel dataset at the county-year level and re-estimate the average treatment effect of the CalEITC on my employment outcomes.[21] As reported in Table A.5, I find between a 2.6 and 3.7 percentage point drop in full-time work and between a 2.9 and 3.6 percentage point increase in part-time work, confirming the direction and magnitude of the effects I observe in my triple-difference approach.

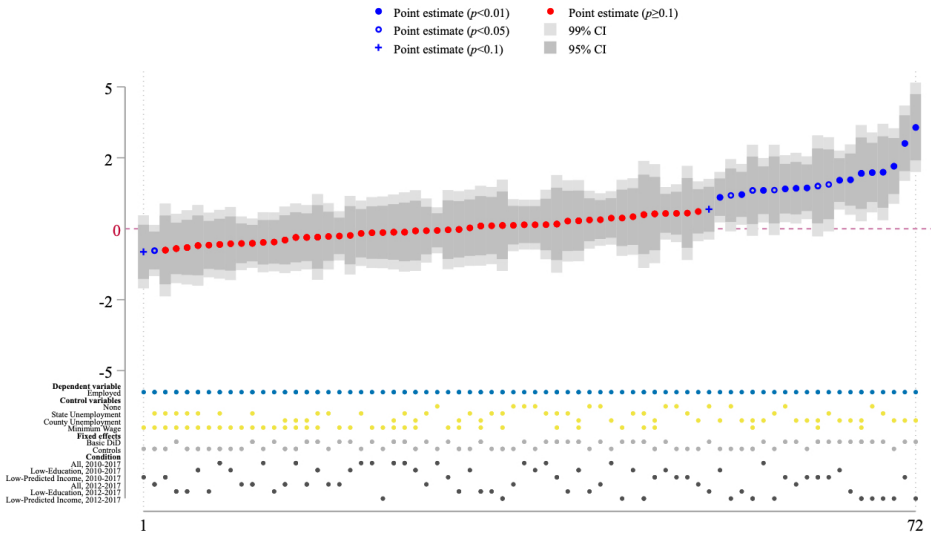
Second, I consider how choices made in the selection of my sample and specification might impact my results. I do so by interacting a range of specification and sample choices together to show whether or not my results are an outliers and assess whether there are patterns in the effect of some of these choices (Kleven, 2024). I plot the resulting range of coefficient estimates in a series of specification curves.[22] Figure 2.7c shows that the part-time results are robust to all combinations, with a median effect of approximately a 3 percentage point increase in part-time employment. Figure 2.7b shows the full-time results; for all but three estimates, I find a statistically significant negative coefficients, and in no case is the estimated effect positive. However,

[21] The current Stata and R implementation of SDID assume that treated units all have equal weights. Since I want to obtain a population-weighted average treatment effect, I implement a modified SDID procedure that runs a single SDID model for every county in California, then report the weighted average of the resulting treatment effects, which I refer to as the County-Specific approach. I also modify the bootstrap procedure discussed by Clarke et al. (2023) to construct standard errors. I consider two dependent variables: the average employment rate in a given county among single mothers, and the difference in the employment rate within a county between single mothers and non-mothers. The latter is my preferred estimate, as it approximates my triple-difference approach.

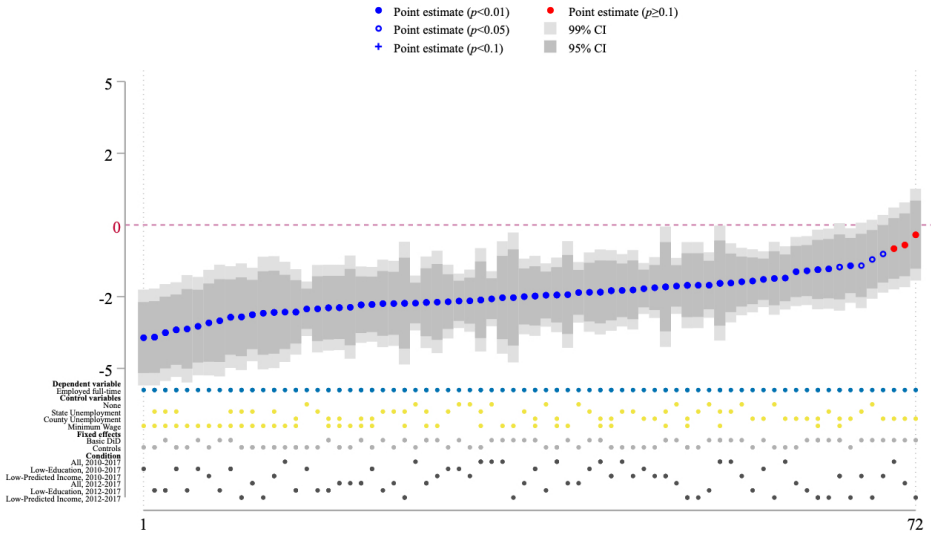
[22] Specifically, using the *specurve* package in Stata (https://github.com/mgao6767/specurve), I run all combinations of three sets of options. The first set governs which, if any, unemployment and minimum wage controls will be included, all interacted with the count of qualifying children. The options include using none, controlling only for state-level or county-level unemployment rates, controlling only for minimum wages, and controlling for both minimum wages and either state or county unemployment. I consider county-level minimum wages to attempt to account for more local variation in labor market conditions. For the second set, I toggle between including and excluding demographic controls. For the third, I vary the sample of single women (all, low-education, and low-predicted earnings) and the time period (2010 through 2017, and 2012 through 2017) used in my analysis. The restriction on predicted earnings is meant as a proxy for low-income workers likely to be in the CalEITC-eligible range (Eissa and Liebman, 1996, Kleven, 2024). For more information about predicted income, see Appendix A.2.3.

my estimate of the full-time effect is one of the largest in magnitude, and for approximately one quarter of the estimates–predominantly those where I limit the sample to women with predicted income in the bottom half of the income distribution, without controlling for demographic factors, and use county-level (rather than state-level) unemployment rates, and omit minimum wage controls–there is a statistically significant increase in overall employment, reflecting the smaller reduction in full-time work in these specifications, as shown in Figure 2.7a. I take two things from this: that there are clearly a set of specifications that produce a net increase in employment, but that these results may be biased due to the omission of important controls and the particular samples used.

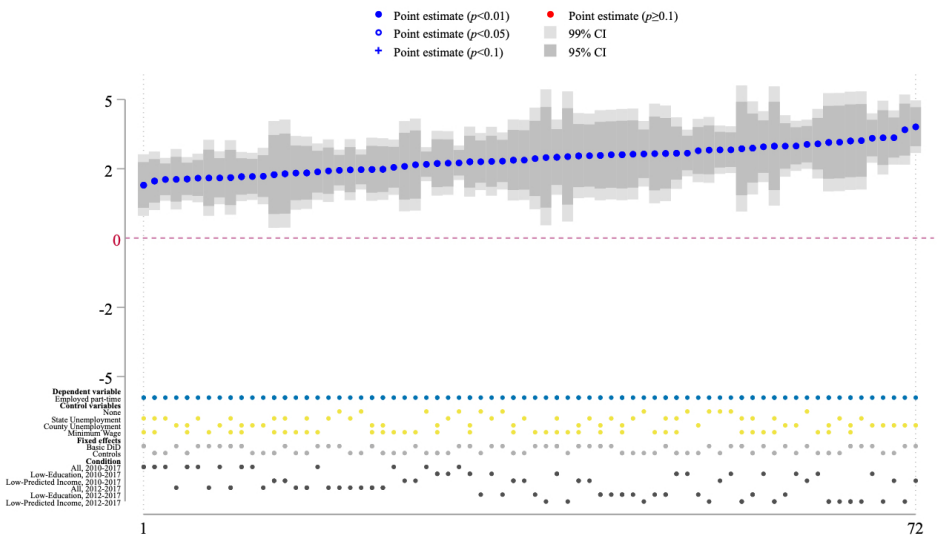
Figure 2.7: Specification curves over sample and controls



(a) Employed in the last 12 months



(b) Employed full-time in the last 12 months



(c) Employed part-time in the last 12 months

Note: These figures reports the results of Equation (2.1), estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, and not potential qualifying children as define in Appendix A.2.2. Each panel has a different dependent variable: (a) any employment in the last 12-months, (b) employed full-time in the last 12-months, and (c) employed part-time in the last 12-months (each equal to 100 if true, for y-axis scale). Each dot is the coefficient on the treated indicator, varying the sample (all single women, single women without a college degree, or single women with low predicted earnings), the inclusion or exclusion of demographic controls, the inclusion or exclusion of unemployment controls (either state or county) interacted with the number of qualifying children, and the inclusion or exclusion of controls for state binding minimum wages interacted with the count of qualifying children. Standard errors are clustered at the state-level. These plots are generated via the *specurve* Stata package.

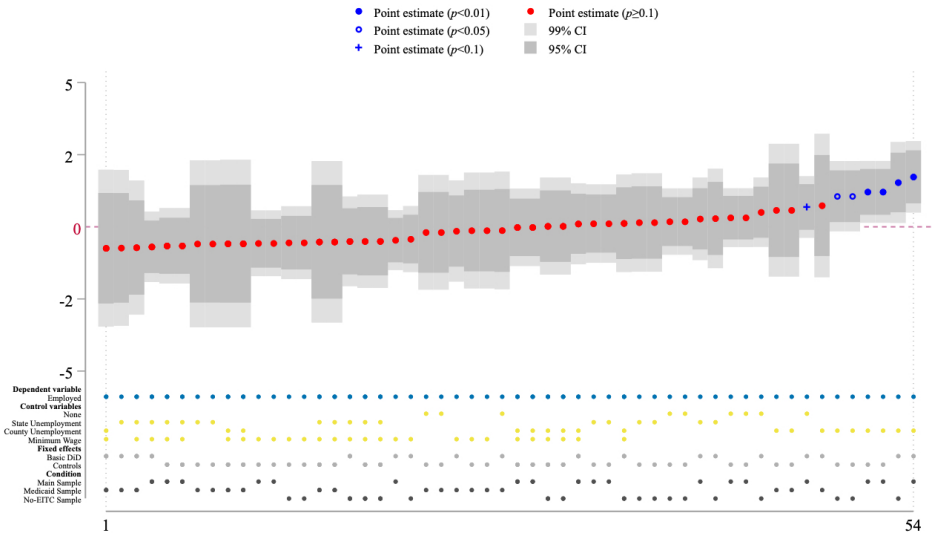
Third, a concern with this identification strategy is that there is some other state-level policy shock concurrent to the CalEITC either in California or in the pool of control states that might drive this effect. I directly examine the trends in expenditures and caseloads in other social safety net programs in Figure A.12. I observe that, with the potential exception of a slightly slower decline in Supplemental Nutrition Assistance Program (SNAP) caseloads per capita in California relative to control states, there appears to be no evidence of other changes in social safety net policies that might confound my results. The major trend is the recovery from the Great Recession for counter cyclical programs, which argues for the inclusion of the unemployment controls, interacted with the count of qualifying children, to capture any effect such programs.

Fourth, I consider how choices in my set of control states might impact my results, constructing two alternative sets of control states, both of which start with my initial pool of 27. First, following the norm in the synthetic control literature (see Abadie (2021)), I drop all states with a state EITC prior to 2020.[23] Second, I restrict the sample to the pool of states that expanded Medicaid in 2014.[24] As part of the Affordable Care Act, states were able to expand access to Medicaid to families with higher-levels of earnings. As noted by the Congressional Budget Office (CBO) in its analysis of the ACA, loosening the ties between employment and health insurance could induce some workers to reduce their participation in the labor force, either along the extensive or intensive margins (Harris and Mok, 2015). Existing empirical evidence of the effect of Medicaid expansion on labor supply is mixed, with some null findings (Duggan et al., 2019) and some evidence of a transitory negative employment effect (Peng et al., 2020). The results are presented in Figure 2.8, which mirrors the first set of specification curves, but plots variations in the pool of control states, along with demographic, unemployment, and minimum wage controls. Overall, I see markedly similar results to those in Figure 2.8: the CalEITC appears to induce an increase in part-time work of between 3 and 4 percentage points, mirrored by a marginally smaller in magnitude and more variable drop in full-time work.

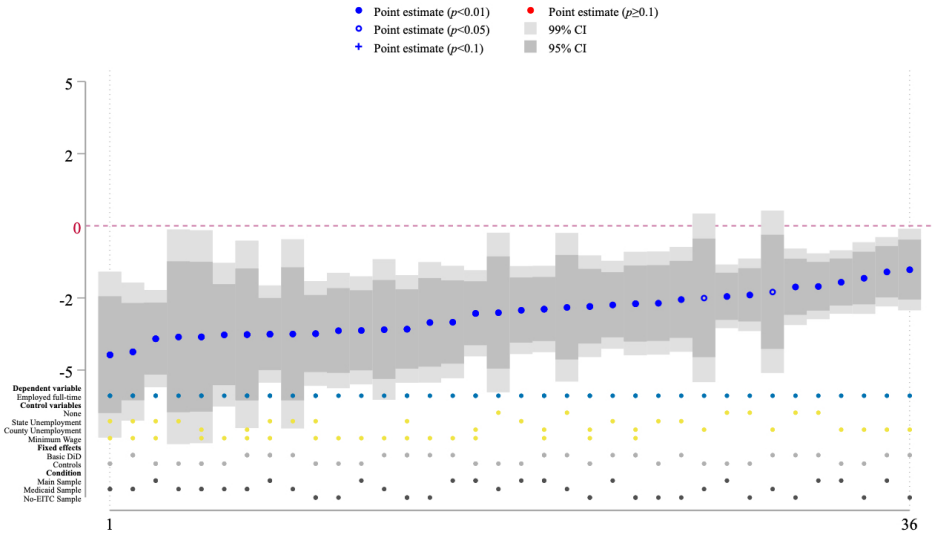
[23] Specifically, I drop Delaware, Indiana, Nebraska, New York, Oklahoma, and Virginia, leaving a set of 20 control states. This approach is similar to the main specification in Neumark and Li (2024), who uses the same pool of control states with the exception of the inclusion of Alaska.

[24] States that are dropped include Alabama, Florida, Georgia, Idaho, Indiana, Mississippi, Missouri, Nebraska, North Carolina, Oklahoma, Pennsylvania, South Dakota, Tennessee, Texas, Utah, Virginia, and Wyoming, leaving a set of 9 control states.

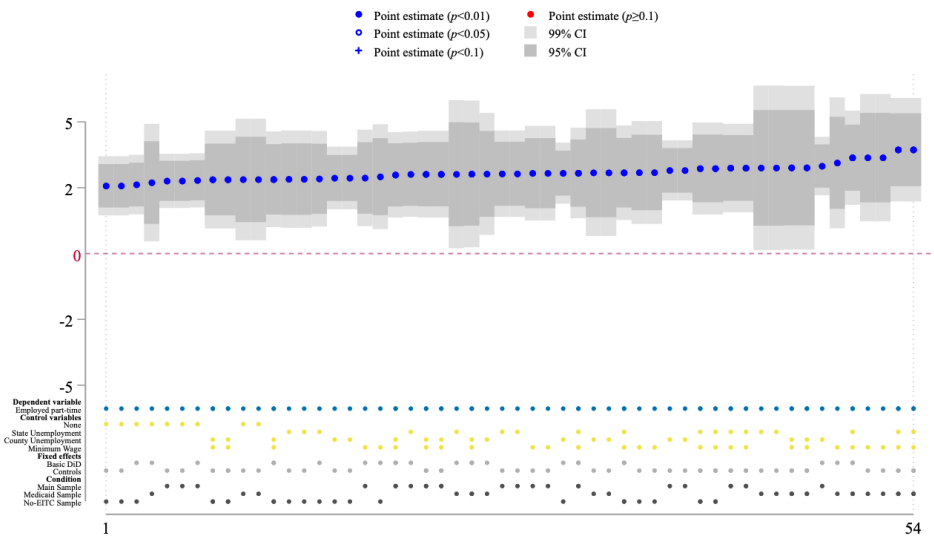
Figure 2.8: Specification curves over control pool



(a) Employed in the last 12 months



(b) Employed full-time in the last 12 months



(c) Employed part-time in the last 12 months

Note: These figures reports the results of Equation (2.1), estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, without a college degree, and not potential qualifying children as define in Appendix A.2.2. Each panel has a different dependent variable: (a) any employment in the last 12-months, (b) employed full-time in the last 12-months, and (c) employed part-time in the last 12-months (each equal to 100 if true, for y-axis scale). Each dot is the coefficient on the treated indicator, varying the control pool (no EITC changes, no EITC, with Medicaid expansion), the inclusion or exclusion of demographic controls, the inclusion or exclusion of unemployment controls (either state or county) interacted with the number of qualifying children, and the inclusion or exclusion of controls for state binding minimum wages interacted with the count of qualifying children. Standard errors are clustered at the state-level. These plots are generated via the *specurve* Stata package.

My results are also robust to using a women's self-reported count of their own children in the household to assign treatment (see Figure A.13), and to assessing the effect of the CalEITC on labor supply outcomes using a simulated instrument approach, following Bastian and Jones (2021) (see Table A.6). In Appendix A.3 I rule out that the increase in labor supply is fully driven by self-employment income, finding that approximately 75 percent of the increase in part-time work coming from workers reporting wage income. I do find that self-employment does increase after the implementation of the CalEITC, in a pattern consistent with both part-time workers supplementing their wage income with self-employment income or a temporary learning effect, where workers were unaware that the CalEITC definition of earnings deviated from that used by

the Federal EITC.

2.5.4 Decomposition by prior employment

The full-to-part time shift in employment I observe could be explained by two mechanisms: that previously full-time employees reduce their number of hours worked in response to the CalEITC, or that workers that would have entered the labor market as full-time workers absent the CalEITC choose instead to work part-time. I utilize the longitudinal nature of the monthly CPS to distinguish between these effects, estimating the effect of the CalEITC on my employment measures, conditional on an individual's employment one-year prior. Given the smaller sample size of the CPS, exacerbated by further restricting to low-education women who appear in my longitudinal sample, I view these results as suggestive, but not definitive. I use the same overall approach as with the ACS: estimating Equation (2.1) with the full set of demographic and economic controls. I take advantage of the monthly data and classify treatment as beginning after the enactment of the CalEITC in June of 2015 by dropping the first two quarters of 2015.

I perform a simple test for bias in Table A.7: if single women with and without kids are differentially likely to change their labor supply conditional on their employment status, and this changes over time, then my results conditioning on prior-year employment may be biased. I find that in no case do single mothers respond in a statistically significantly different way than non-mother after 2015: however, coefficients are larger for those employed part-time or unemployed in the prior year. These two groups are marginally smaller, and I therefore interpret these coefficients with some additional caution.

I present the results of this analysis in Table 2.3. Columns 1 and 2 show the overall results,

Table 2.3: Triple-difference estimates of the effect of the CalEITC on weekly employment via the CPS

	(1)	(2)	(3)	(4)	(5)
	Full Sample	Longitudinal Sample			
	All	All	Full-time last year	Part-time last year	Not employed last year
Panel A: Employed last week					
ATE	4.4**	5.8	-2.0	5.7	2.3
	(2.0)	(4.6)	(2.6)	(6.9)	(3.8)
Observations	215,832	62,813	27,796	11,987	23,030
Adj. R-Square	0.068	0.070	0.025	0.037	0.062
Treated group mean in pre-period	59.9	59.2	59.2	59.2	59.2
Panel B: Employed full-time last week					
ATE	0.3	1.4	-6.2*	0.7	0.7
	(2.1)	(4.4)	(3.4)	(4.1)	(3.4)
Observations	215,832	62,813	27,796	11,987	23,030
Adj. R-Square	0.059	0.061	0.029	0.018	0.031
Treated group mean in pre-period	43.0	42.9	42.9	42.9	42.9
Panel C: Employed part-time last week					
ATE	4.0***	4.5**	4.2**	5.0	1.6
	(0.9)	(2.0)	(2.0)	(5.7)	(2.5)
Observations	215,832	62,813	27,796	11,987	23,030
Adj. R-Square	0.016	0.021	0.014	0.017	0.031
Treated group mean in pre-period	16.9	16.3	16.3	16.3	16.3
Basic Triple-Difference	Yes	Yes	Yes	Yes	Yes
Add Demographic Controls	Yes	Yes	Yes	Yes	Yes
Add Unemployment Controls	Yes	Yes	Yes	Yes	Yes
Add Minimum Wage Controls	Yes	Yes	Yes	Yes	Yes

Note: This table reports the results of Equation (2.1), estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, who do not have a college degree, and who are not potential qualifying children as define in Appendix A.2.2, via the monthly Current Population Survey (CPS) from 2012-2017, omitting the first two quarters of 2015. Panel's A, B, and C vary only in the dependent variables, which are indicators for any employment, full-time employment, and part-time employment last week, respectively. Full time work is defined as working at least 35 hours a week, and part time work is defined as working under 35 hours a week. Column 1 shows the results of the model run over the full sample, while Columns 2 through 5 report the results of the full longitudinal sample, with columns 3 through 5 restricting the sample further conditional on an individual's employment status one year ago. All specifications include demographic controls, including indicators for 5-year age bins, race and Hispanic origin groups, educational attainment, bins for the age of the individual's youngest qualifying child, and the number of adults in an individual's household, as well as controls for state unemployment rates interacted with the number of qualifying children. Standard errors clustered at the state and individual level. * 0.10 ** 0.05 *** 0.01

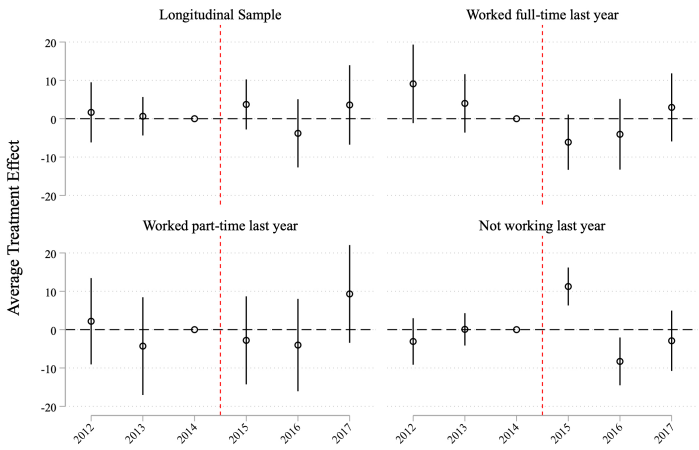
not conditioning on prior year employment, for the full monthly CPS sample and the longitudinal sample. In both, overall employment increases, driven by an increase in part-time work, though in the longitudinal sample the effect on overall employment is statistically insignificant, and there is no decrease in full-time work. Columns 3 through 5 condition on prior-year employment. Here,Isee that full-time work falls and part-time work increases among women who were employed full-time least year. This result implies that the shift I observe in the ACS data likely emerges from those women who worked full-time last year and chose to reduce their hours worked as a response to the CalEITC.

I also report the event-study estimates from these models in Figure 2.9. Each sub-figure depicts the results for each dependent variable (mirroring the last two rows in Table 2.3), within which there are four further sub-figures for each of the longitudinal samples (mirroring columns 2 through 5). For ease of presentation, I omit the event study estimates for column (1) and row (2). While there is an increase in overall part-time work, concentrated in 2015 and 2016, there is not a decrease in overall full-time work. There is some evidence that the decrease in full-time work among women who were employed full-time last year may be driven some other set of factors, as evidenced by a pre-trend for that sample. However, the presence of some shift from full-time to part-time work seems to be supported by the matching shift into part-time work among these same women. Finally, among women who were not employed last year, in 2015 there was a notable increase in full-time work followed by a decline in 2016. However, the 2015 sample is the nosiest, given I am missing six months of the sample, so should be taken with some caution.[25]

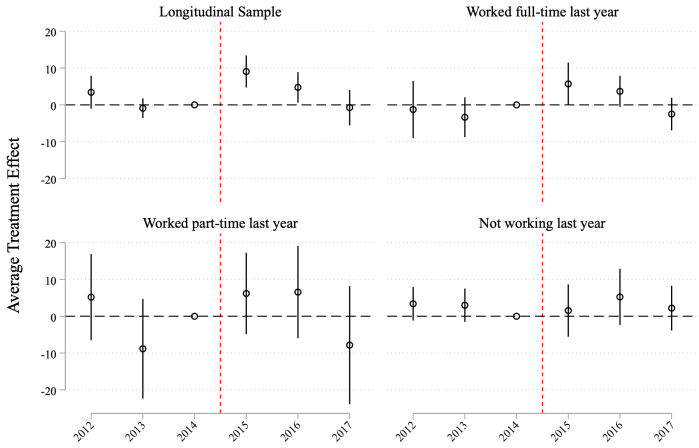
While there is clear evidence of some short-term increase in part-time work, the presence of pre-trends and more complex patterns of response when considering full-time work makes it

[25]I present the results over the full 2010 through 2017 period in Table A.8, and the results are broadly consistent.

Figure 2.9: Effect of the CalEITC on employment, by prior-year employment using the longitudinal CPS.



(a) Employed full-time last week



(b) Employed part-time last week

Note: This figure reports the results of Equation (2.3), estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, who do not have a college degree, and who are not potential qualifying children as define in Appendix A.2.2, via the monthly Current Population Survey (CPS) from 2012-2017, omitting the first two quarters of 2015. The sample is further limited to those individuals who can be matched to their own CPS record 12-months prior to the current interview, and the sample is re-weighted to be representative of the current-month CPS. The two sub-figures vary only in their dependent variables, which are indicators for full-time employment and part-time employment last week, respectively. Full time work is defined as working at least 35 hours a week, and part time work is defined as working under 35 hours a week, and both are defined relative to the week prior to the CPS interview. Within each sub-figure, the panels show the results of the event-study run over different samples: the full longitudinal sample, and the sample limited based on an individual's employment status one year ago. All specifications include demographic controls, including indicators for 5-year age bins, race and Hispanic origin groups, educational attainment, bins for the age of the individual's youngest qualifying child, and the number of adults in an individual's household, as well as controls for state unemployment rates interacted with the number of qualifying children. Standard errors clustered at the state-level for the longitudinal sample. * 0.10 ** 0.05 *** 0.01

difficult to pin down exactly the margins of response. However, the clearest picture is that there was some sort of shift between full-time and part-time work among women who were employed full-time last year. This does not rule out other margins: the increase in full-time work in 2015 among those who did not work in the prior year could imply some short-term weeks worked response I did not detect in the ACS. Overall, these results seem to imply that there was some shift from full-time to part-time work, in line with my main ACS results.

2.5.5 Heterogeneity

To explore mechanisms behind the shift from full-time to part-time work, I examine heterogeneity across a number of household and demographic characteristics. Namely, I analyze variation by number of qualifying children, household size (measured by the number of adults in an ACS household), and the age of the youngest child claimed by a mother, to better understand how family structure and caregiving responsibilities may influence responsiveness to CalEITC incentives. These factors capture potential flexibility in labor supply decisions, with some—like number of qualifying children—directly featured in prior literature (see Kleven (2024)), and others—such as the number of adults in the household—feed into considerations about intra-household bargaining and earnings vs. home production decisions (Cherchye et al., 2012). The goal with this analysis is to determine whether the reduction in hours worked is common across household characteristics or unique to certain living arrangements. I report heterogeneity results across several other characteristics, including race, Hispanic origin, and disability status, in the appendix. These supplementary analyses serve two purposes: they acknowledge that different groups may respond differently to labor market incentives, and they reflect normative concerns

that policymakers may have about potential adverse effects of tax policy reforms on historically disadvantaged populations. For each estimate of the heterogeneity of response, I re-estimate Equation (2.1), restricting the sample to the specific group of women. This means I am re-estimating the entire model, including the coefficients for my control variables, for each group.

First, I estimate the effect of the CalEITC across the number of qualifying children claimed by a single mother, comparing single women with no children to those with 1, 2, or 3 or more qualifying children.[26] As shown in Table 2.4, I find no effect of the CalEITC on the labor supply of women with three or more qualifying children. I estimate a roughly even (in magnitude) drop in full-time and increase in part-time work for women with one qualifying child, and a larger drop in full-time work for mothers with two qualifying children that exceeds the magnitude of the increase in part-time work for that sample, though I cannot reject the null that the two coefficients are equal. The event-study results for this sample are reported in Figure A.14, which show that it takes longer for women with two qualifying children to respond fully than women with one qualifying child, potentially reflecting some frictions associated with entering the labor market that rises with the number of children. Interestingly, while the effect for mothers with one qualifying child disappears by 2017, it remains for mothers with two qualifying children. These results diverge from the research on the EITC reforms in the 1990s, which found that the employment effects of the federal EITC increased with the number of qualifying children (Hoynes and Patel, 2018, Kleven, 2024).[27] The fact that I observe the a small and statistically insignificant response for mothers with three or more children implies that household composition can impact

[26] As shown in Figure A.1, CalEITC benefits increase substantially when moving from one to two qualifying children, and much more modestly when moving between 2 and 3 or more kids.

[27] Prior to the 1990s, the EITC provided the same benefit amount to families irrespective of how many qualifying children they claimed. In the 1990s, families with 2 or more children received a larger benefit increase than those with one child, and in 2008 an additional benefit was given to those families with three or more qualifying children.

labor supply responses separate from changes in the benefit amount.

Second, I assess the effect of the CalEITC by the number of adults in the woman's household. As shown in Table 2.5 I find that employment effects are concentrated among households with three or more adults (inclusive of the women in my sample). Here, part-time work increases by 5.4 percentage points, and full-time work decreases by 8.0 percentage points. This is mirrored in the event-study results in Figure A.15, where the response is clearest among large households. Among households with one or two adults, the effects on full- and part-time work are smaller in magnitude, not always statistically significant. The concentration of the effects in larger households is indicative of a role for household-level decision-making in the response to the CalEITC, with implications for the large literature on household labor supply decisions (Chiappori, 1992, Blundell et al., 2005, Cherchye et al., 2012). One straight-forward story is that it is less costly for single mothers in larger households to reduce their labor supply given the presence of other earners (or non-earned income) in the household. There may also be more complex stories about trading off household childcare or other household responsibilities between members of the household in the presence of the CalEITC.

Third, I assess the effect of the CalEITC differentially across the age of a women's youngest qualifying child. If single women are shifting from full- to part-time work due primarily to childcare-related home-production, I would expect to see the effects of the CalEITC concentrated among women with a child under the age of 6. However, as I show in Table 2.6, the employment response is not unique to women with young children. Rather, there is evidence of such a transition both among mothers with young (0-5-year-olds) and older (13+) children. This implies that the observed shift in labor supply is not only a story regarding childcare costs, but rather that childcare could be one of several types of household responsibilities that are traded off

Table 2.4: Triple-difference estimates of the effect of the CalEITC on annual employment, by count of qualifying children

	(1) Full sample	(2) 1 QC	(3) 2 QC	(4) 3+ QC
Panel A: Employed in the last 12-months				
ATE	-0.7	0.3	-2.1*	-0.7
	(0.5)	(0.7)	(1.1)	(1.8)
Observations	441,172	344,996	312,670	291,762
Adj. R-Square	0.076	0.080	0.082	0.083
Treated group mean in pre-period	71.6	73.9	72.4	64.9
Implied employment effect	-11,737	2,264	-11,528	-2,371
Panel B: Employed full-time in the last 12-months				
ATE	-3.9***	-3.8***	-5.9***	-1.2
	(0.6)	(1.1)	(1.6)	(1.2)
Observations	441,172	344,996	312,670	291,762
Adj. R-Square	0.068	0.070	0.069	0.067
Treated group mean in pre-period	51.0	51.9	53.0	45.8
Implied employment effect	-66,628	-30,677	-31,745	-4,038
Panel C: Employed part-time in the last 12-months				
ATE	3.2***	4.0***	3.7***	0.5
	(0.6)	(0.8)	(1.0)	(1.6)
Observations	441,172	344,996	312,670	291,762
Adj. R-Square	0.028	0.035	0.033	0.035
Treated group mean in pre-period	20.6	22.0	19.3	19.1
Implied employment effect	54,890	32,941	20,217	1,668
Basic Triple-Difference	Yes	Yes	Yes	Yes
Add Demographic Controls	Yes	Yes	Yes	Yes
Add Unemployment Controls	Yes	Yes	Yes	Yes
Add Minimum Wage Controls	Yes	Yes	Yes	Yes

Note: This table reports the results of Equation (2.1), estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, who do not have a college degree, and who are not potential qualifying children as define in Appendix A.2.2, via the American Community Survey (ACS) from 2012-2017. Panels A, B, and C vary only in the dependent variables, which are indicators for any employment, full-time employment, and part-time employment at any point in the last 12-months, respectively. Column 1 reports the results from the full sample, while columns 2 through 4 restrict the sample to include those with no qualifying children and the number reported (1, 2, or 3+). All models include demographic, state unemployment, and minimum wage controls. Standard errors clustered at the state level. * 0.10 ** 0.05 *** 0.01

Table 2.5: Triple-difference estimates of the effect of the CalEITC on annual employment, by count of adults in household

	(1) Full sample	(2) 1 adult (self)	(3) 2 adults	(4) 3+ adults
Panel A: Employed in the last 12-months				
ATE	-0.7	0.8	0.5	-2.6**
	(0.5)	(1.5)	(1.2)	(1.2)
Observations	441,172	123,621	158,931	158,620
Adj. R-Square	0.076	0.078	0.071	0.071
Treated group mean in pre-period	71.6	76.1	70.9	67.6
Implied employment effect	-11,737	3,898	3,062	-14,074
Panel B: Employed full-time in the last 12-months				
ATE	-3.9***	-1.0	-1.2	-8.0***
	(0.6)	(1.6)	(1.3)	(1.1)
Observations	441,172	123,621	158,931	158,620
Adj. R-Square	0.068	0.072	0.061	0.049
Treated group mean in pre-period	51.0	55.0	51.0	46.7
Implied employment effect	-66,628	-5,014	-8,047	-43,819
Panel C: Employed part-time in the last 12-months				
ATE	3.2***	1.8*	1.7	5.4***
	(0.6)	(0.9)	(1.1)	(0.7)
Observations	441,172	123,621	158,931	158,620
Adj. R-Square	0.028	0.020	0.022	0.040
Treated group mean in pre-period	20.6	21.1	19.9	20.8
Implied employment effect	54,890	8,912	11,109	29,745
Basic Triple-Difference	Yes	Yes	Yes	Yes
Add Demographic Controls	Yes	Yes	Yes	Yes
Add Unemployment Controls	Yes	Yes	Yes	Yes
Add Minimum Wage Controls	Yes	Yes	Yes	Yes

Note: This table reports the results of Equation (2.1), estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, who do not have a college degree, and who are not potential qualifying children as define in Appendix A.2.2, via the American Community Survey (ACS) from 2012-2017. Panels A, B, and C vary only in the dependent variables, which are indicators (multiplied by 100 for scale) for any employment, full-time employment, and part-time employment at any point in the last 12-months, respectively. Column 1 reports the results from the full sample, while columns 2 through 4 restrict the sample by the count of adults (inclusive of the individual in my sample). All models include demographic, state unemployment, and minimum wage controls. Standard errors clustered at the state level. * 0.10 ** 0.05 *** 0.01

within these households.

I also examine the heterogeneity of the labor supply response across race and Hispanic origin, and disability status.[28] My results by race, reported in Table A.9, indicate that there are generally similar results for Hispanic mothers and non-Hispanic White mothers that broadly match that of the full sample. There is no statistically or economically significant effect of the CalEITC on the employment of Non-Hispanic Black women, and, as shown in Figure A.16, the pre-trends for women who identify as Non-Hispanic members of another racial group are fairly noisy, making it difficult to determine if there is any effect for this group. As I report in Table A.10 (and in Figure A.17), I find that single women who report one or more disabilities increase their part-time work by 5.0 percentage points, without any change in full-time work, while women without any reported disability appear to have a result more in keeping with the main sample.[29] This fits with prior work on the EITC and disability by Goodman et al. (2013), who found that while people with disabilities are almost as likely to claim the EITC, they are eligible for much smaller benefit amounts, due both to a lower probability of having children and lower earned income. These factors might lead the CalEITC to have a unique impact on some women with a self-reported disability.

Putting these results together, the results regarding the responsiveness across household characteristics highlight an important role for intra-household bargaining. Women have larger labor supply responses in larger households (by number of adults), and when they can claim fewer qualifying children. In a larger household, the benefit of more time on home production

[28] Prior work has considered the effect of the EITC across different racial groups. For work on labor supply and income inequality see Hardy et al. (2022), for work on the impact of IRS audit procedures, the EITC, and race see Elzayn et al. (2024), and for work on take-up and awareness see Iselin et al. (2023), Jones (2014), Thomson et al. (2022).

[29] Since 2008 the ACS asks a series of questions about whether a given individual has encountered difficulty across a range of areas (hearing, vision, cognitive, ambulatory, self-care, and independent living).

Table 2.6: Triple-difference estimates of the effect of the CalEITC on annual employment, by age of youngest qualifying child

	(1) Full sample	(2) 0-5	(3) 6-12	(4) 13+
Panel A: Employed in the last 12-months				
ATE	-0.7	0.9	-3.0***	-0.8
	(0.5)	(0.8)	(0.9)	(1.3)
Observations	441,172	340,019	312,789	296,620
Adj. R-Square	0.076	0.079	0.082	0.084
Treated group mean in pre-period	71.6	65.7	75.6	78.5
Implied employment effect	-11,737	6,856	-16,232	-3,002
Panel B: Employed full-time in the last 12-months				
ATE	-3.9***	-2.9***	-4.9***	-4.8***
	(0.6)	(1.0)	(1.1)	(1.4)
Observations	441,172	340,019	312,789	296,620
Adj. R-Square	0.068	0.066	0.068	0.073
Treated group mean in pre-period	51.0	42.8	55.1	62.8
Implied employment effect	-66,628	-22,760	-26,654	-17,687
Panel C: Employed part-time in the last 12-months				
ATE	3.2***	3.7***	1.9*	4.0***
	(0.6)	(1.1)	(1.0)	(1.1)
Observations	441,172	340,019	312,789	296,620
Adj. R-Square	0.028	0.032	0.033	0.037
Treated group mean in pre-period	20.6	22.9	20.6	15.7
Implied employment effect	54,890	29,616	10,422	14,685
Basic Triple-Difference	Yes	Yes	Yes	Yes
Add Demographic Controls	Yes	Yes	Yes	Yes
Add Unemployment Controls	Yes	Yes	Yes	Yes
Add Minimum Wage Controls	Yes	Yes	Yes	Yes

Note: This table reports the results of Equation (2.1), estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, who do not have a college degree, and who are not potential qualifying children as define in Appendix A.2.2, via the American Community Survey (ACS) from 2012-2017. Panels A, B, and C vary only in the dependent variables, which are indicators for any employment, full-time employment, and part-time employment at any point in the last 12-months, respectively. Column 1 reports the results from the full sample, while columns 2 through 4 restrict the sample to include those with no qualifying children and the number reported (1, 2, or 3+). All models include demographic, state unemployment, and minimum wage controls. Standard errors clustered at the state level. * 0.10 ** 0.05 *** 0.01

might be worthwhile given the presence of other sources of household earnings, or there might be more opportunities for utility-enhancing exchanges of household responsibilities. In a smaller household, without other sources of income or household labor, such a shift might be much more difficult.

2.6 Effect of the CalEITC on Earned Income

Given my empirical results, there is clear evidence of a shift from full-time to part-time work after the implementation of the CalEITC. I therefore would expect earned income to fall, with an increase in density of workers around the level of annual earnings that would maximize the CalEITC, and a decrease in density higher up the earnings distribution. The size and nature of this decline in earned income is a critical element of determining the welfare effects of the CalEITC, as I will discuss in depth in Section 2.7 and Section 2.8, and plays a role in assessing which model of labor supply to use to evaluate the effects of tax policies. In this section, I provide suggestive evidence of a large decline in earned income, consistent with a discrete labor supply transition. However, I also discuss how certain limitations in the ACS earnings data make a full accounting for the earnings effects of the CalEITC difficult.

First, I directly estimate the effect of the CalEITC on earned income, using the triple-difference strategy I used in Section 2.5.2. The results, presented in Table 2.7, are from three regressions: an OLS estimate on the full-sample, an OLS estimate restricted to only those with positive earnings, and a Poisson regression, designed to take into account non-earnings (i.e. those with $0 in earned income) (Chen and Roth, 2024). Across all specifications, I find a decrease in earnings of between 4 and 7 percent, or between $600 and $1,400 dollars annually, not condi-

tioning on those in the labor force. These represent intent-to-treat estimates, or estimates of the effect of the CalEITC on the earnings of single mothers in California without focusing on those who adjusted their labor supply. A simple back-of-the-envelope calculation, given my estimate of a three-percentage-point shift from full-time to part-time work, is that the earnings of impacted mothers could fall by between $20,000 and $47,000.[30] These estimates are large, though not inconsistent with full-time workers making over $30,000 annually shifting to part-time work with annual earning that would maximize the CalEITC. As I will discuss in depth later in the section, there are certain reasons to be skeptical about features of the earnings data in the ACS, but the size of the earnings results imply labor supply changes that are discrete, rather than continuous, given the framework laid out in Section 2.3.

Second, following Kosonen and Matikka (2023), I visualize the CalEITC-induced shift in the earnings distribution between 2012 and 2017. I divide my sample of single women into eight groups: 2012-2014, and 2015-2017, with and without qualifying children, and in California verses in the pool of control states. For each group (ex. single mothers in 2012-2014 in California), I calculate the share of the group in each $6,000 real (2017 USD) earned income bin. By comparing changes in the shares in each bin across the eight groups, I can visually re-create the logic of my triple-difference strategy, seeing how the difference in the earnings distribution between single women with and without children changed in California before and after the implementation of the CalEITC, compared to the same exercise in my pool of control states.

These plots are presented in Figure 2.10. In Figure 2.10a, I show the results for California. The solid blue line represents the difference in the distribution of income between single mothers

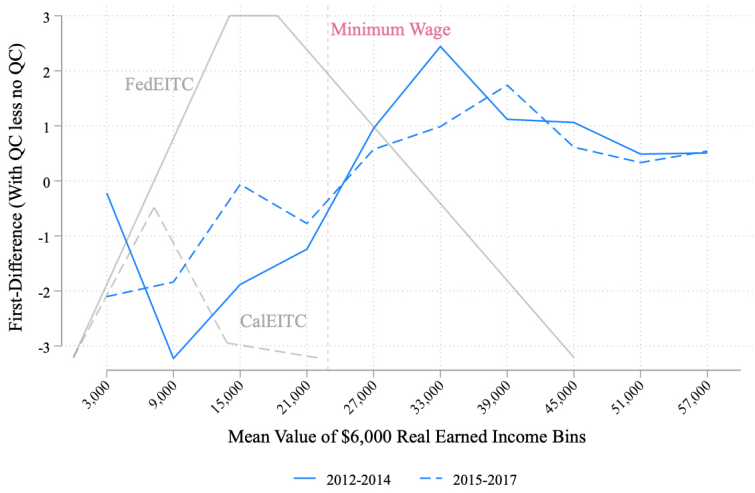
[30] Assuming a 3-percentage point shift from full-time to part-time work, and no change in earnings in the rest of the population of single mothers, I multiply the estimated earnings effect β by $\frac{100}{3}$.

Table 2.7: Triple-difference estimates, effect of the CalEITC on annual earnings

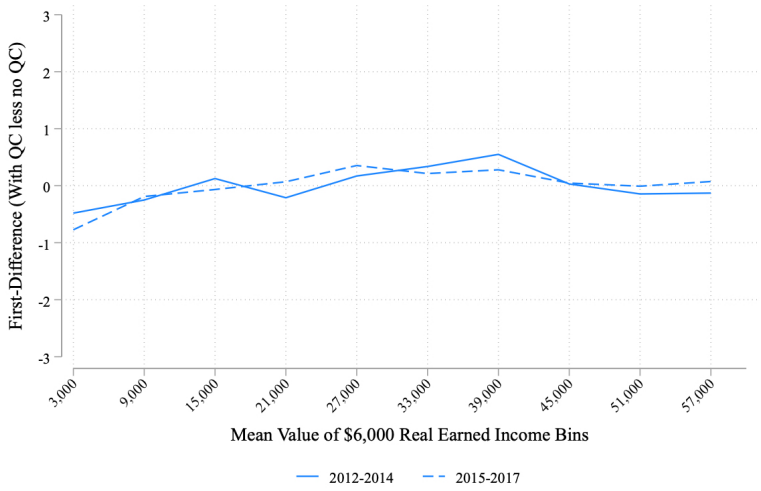
	(1) Basic Triple Difference	(2) Add Demographic Controls	(3) Add Unemployment Controls	(4) Add Minimum Wage Controls
Panel A: OLS - No restrictions				
ATE	-608.1***	-847.9***	-1368.3***	-1413.3***
	(178.4)	(204.9)	(377.9)	(497.3)
Observations	441,172	441,172	441,172	441,172
Adj. R-Square	0.011	0.102	0.102	0.102
Treated group mean in pre-period	21,238	21,238	21,238	21,238
Panel B: OLS - positive income only				
ATE	-1233.3***	-1282.0***	-1958.2***	-1758.2***
	(212.1)	(216.5)	(404.5)	(575.1)
Observations	326,229	326,229	326,229	326,229
Adj. R-Square	0.011	0.118	0.118	0.118
Treated group mean in pre-period	29,704	29,704	29,704	29,704
Panel C: Poisson				
ATE	-0.04***	-0.05***	-0.08***	-0.08***
	(0.01)	(0.01)	(0.02)	(0.03)
Observations	441,172	441,172	441,172	441,172
Pseudo-R-squared	0.015	0.136	0.136	0.136
Treated group mean in pre-period	21,238	21,238	21,238	21,238
Effect in USD	-695	-944	-1,452	-1,464
Triple-Difference	Yes	Yes	Yes	Yes
Add Demographic Controls	No	Yes	Yes	Yes
Add Unemployment Controls	No	No	Yes	Yes
Add Minimum Wage Controls	No	No	No	Yes

Note: This table reports the results of Equation (2.1), estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, not potential qualifying children as define in Appendix A.2.2, and who do not have a college degree via the American Community Survey (ACS) from 2010-2017. The dependent variable is annual earnings, defined as wages plus self-employment income, in real 2019 dollars. Panels A estimates an OLS model, and includes individuals with no earned income. Panel B estimates the same model excluding non-workers. Panel C estimates a Poisson model via the Stata command *ppmlhdfe*, including those with no earned income. In columns 2, 3, and 4, demographic controls include indicators for 5-year age bins, race and Hispanic origin groups, educational attainment, bins for the age of the individual's youngest qualifying child, and the number of adults in an individual's household. Standard errors clustered at the state level. * 0.10 ** 0.05 *** 0.01

Figure 2.10: Changes in the earnings distribution over time, California verses control states



(a) California



(b) Control states

Note: These figures show the difference in the distribution of earned income (in real 2017 dollars) between individuals with and without qualifying children over two periods: 2012-2014 (solid line) and 2015-2017 (dashed line). Panel (a) shows the results for California and panel (b) for the pool of control states limited to states without substantial changes in EITC policy from 2008-2019, The distribution is for the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, not potential qualifying children as define in Appendix A.2.2, and without a college degree. Panel (a) also plots the federal and California EITC schedule for a single parent with two qualifying children in 2017, as well as the level of earnings a full-time worker (40 hours a week, 52 weeks a year) would make at the California minimum wage in 2017.

and non-mothers in 2012 through 2014, while the dashed blue line shows the same difference for 2015 through 2017. There is an increase in the relative density between $6,000 and $18,000 dollars, just at and above the CalEITC-maximizing point, and a decrease in density higher up the earnings distribution, above the annual earnings level of a full-time worker, working 52 weeks per year, at the binding minimum wage in 2017 (represented by the dashed vertical line). This same pattern is not seen in Figure 2.10b where I plot the same difference for the set of control states. Here, for both periods, the difference between the share of the sample of single women with and without children is approximately 0 for all bins. These results imply that the CalEITC directly impacted the distribution of earnings, an effect that can be rationalized by a model of discrete labor supply, where workers choose between a set of employment options, maximizing their utility between the set of available choices.

Importantly, this shift does not occur in the regions that would be predicted by a model of continuous labor supply. In that setting, I would expect a decline in employment among workers at the federal EITC kink point, who shift their earnings marginally downward to maximize their CalEITC. In fact, I see the opposite; The earnings level where I observe a decline in density is above that of a full-time worker, earning minimum wage, who worked 52 weeks out of the year. Further, it is consistent with the approximate magnitude of the decline in earnings observed in Table 2.7. As shown in Figure A.18a, I see the same pattern when constructing the same exercise by wage and salary income, which implies that the effect is not driven by some sort of strategic behavior regarding self-employment income.[31] I see an increase in density among workers that are making between $6,000 and $18,000. This is somewhat higher up the income distribution

[31] I also see a similar (and marginally larger) result if I report the result using nominal earnings as shown in Figure A.18b.

than I would have anticipated solely based on parameters of the CalEITC, and matches the hours-worked per week results presented in Figure A.11. This response could be rationalized by a model where, due to adjustment or information frictions, individuals are unable to target their annual earnings appropriately to maximize their CalEITC receipt (Chetty et al., 2011). Alternatively, it could be explained by a salience effect: the CalEITC increases the salience of refundable tax credits in general, leading to some pool of tax units shifting to a level of earnings that maximizes the federal EITC and CTC. Finally, it could be a feature of noisy ACS earnings data: there could be some error given the nature of self-reported earnings data or the timing of reference period.

This third feature is worth considering briefly. Namely, that there are several potential concerns with the self-reported earnings data in the ACS. First, self-reported income in surveys can be biased, with work using merged survey and administrative data has shown that survey-based income measures often underestimate actual income (Jones and Ziliak, 2022, Meyer and Mittag, 2019, Meyer et al., 2021, 2022). Second, the ACS has a unique reference period; the 12 months prior to the survey. If an individual's earnings over a reference period are notably different from their annual earnings over the course of a tax year, then my results, particularly for 2015, will be biased. This is less of an issue for earlier or later years, as the entire 12-month reference period falls fully within the treated or control period, but it does raise additional concerns about accuracy and reliability. Third, missing observations for earned income are common in the ACS: a full quarter of the observations in my sample have an "allocated" earning values, representing cases where the Census Bureau imputed missing income data. My overall earnings results are robust to excluding these imputed earnings records (see Table A.11), but this is still a point of concern, especially when estimating the effect of the CalEITC on the distribution of earnings.

These issues all suggest that my earnings results likely contain a considerable amount of

noise, which makes more detailed estimates of the distributional effects of the CalEITC difficult. I present some results along these lines in the appendix. There, I follow Bastian (2020) and report the results of Equation (2.1), where the dependent variable has been modified to equal one if an individual is employed (however defined) within a given earnings range. The results, shown in Figure A.19, are nosier that those reported in Figure 2.10. I observe declines in full-time work across the earnings distribution. Somewhat comforting, however, is the fact that the increase in part-time work is concentrated at or right above the level of earnings that would maximize the receipt of the CalEITC.

Overall, these results show a decline in earnings consistent with a discrete labor force transition from full-time to part-time work. However, given issues with the earnings records available in the ACS, it is difficult to pin down the exact nature of the shift in the earnings distribution. This represents an important factor for further investigation: as I will discuss in the next several sections, analyzing the welfare implications of the CalEITC requires knowing the counter-factual income of a full-time worker.

2.7 Elasticities

To put my results in the context of the broader literature, I construct what the labor economics literature refers to as a "mobility" elasticity, which captures the degree to which a worker might re-locate from one point in the earnings distribution to another in response to a change in the relative average tax rates (Saez, 2002, Kosonen and Matikka, 2023, Kreiner et al., 2015). Importantly, unlike the literature on the elasticity of taxable income, this measure considers the effect of changes in the average tax rates faced by taxpayers at different levels of annual earnings,

reflecting that this response appears to be a discrete, rather than marginal, changes in labor supply (Saez, 2002, Kreiner et al., 2015).[32] I discuss the methodology for calculating these elasticities in Appendix A.4.

Constructing this elasticity requires making two assumptions. The first is minor; namely, I assume that, as implied by my main specification, the entire increase in part-time work is driven by a shift from full-time work, and I ignore any additional decrease in full-time work that is not offset by the part-time effect. Second, and more importantly, in order to estimate the mobility elasticity, I need identify the counter-factual earnings these single mothers would have earned but for the CalEITC. This is necessary to calculate the change in the after-tax rates that this sample would have faced. As I discussed in Section 2.6, it is not immediately clear what this counter-factual level of income might be: in fact, it is possible that women with a range of counter-factual levels of earnings responded to the incentives of the CalEITC. Given this uncertainty, I calculate the elasticity given three different assumptions regarding the counter-factual earnings of single mothers that respond to the CalEITC. The first is the annual earnings of a women working 40 hours per week, 52 weeks per year, at the binding 2017 minimum wage in California ($21,840). The second and third are the median and mean annual earnings of single mothers in California in 2014, conditional on working full-time (converted into real 2017 dollars) ($30,655 and $36,413, respectively). As I will show, these estimates can change the elasticity estimates dramatically.

I find mobility elasticities of 1.1, 1.4, and 1.7, for each of the counter-factual earnings levels described above (minimum wage, median, and mean, respectively). Larger counter-factual

[32] All the elasticities below take into account federal individual income taxes and payroll taxes, as well as California's individual income tax. However, I do not account for the marginal incentives of non-tax programs, like the Supplemental Nutrition Assistance Program (SNAP) and Temporary Assistance to Needy Families (TANF). This is because of issues with data availability in the ACS that complicate the calculation of benefit eligibility. I discuss the likely bias this introduces below.

income levels lead to larger estimated elasticities, reflecting the fact that, in percent terms, the after-tax income gains from the CalEITC are smaller if you assume a higher counter-factual income. In the context of the literature on extensive-margin effects of the federal EITC, there is an ongoing debate about what is considered a "reasonable" extensive margin elasticity (Schanzenbach and Strain, 2021, Kleven, 2024). Kleven (2024) argues that, outside of the 1993 reform, the distribution of elasticity estimates is centered on 0, and that the higher average elasticities from the 1993 reform (a mean of 0.63) are likely due to confounding effects from welfare reform and a booming economy in the 1990s. Schanzenbach and Strain (2021) responded (to Kleven (2019), an earlier working paper version of Kleven (2024)) by arguing that, when appropriately controlling for business cycles and looking at the population of women with less educational attainment, there are robust effects of the EITC on labor supply, with elasticities ranging from 0.78 to 1.66.[33] In this context, all my estimates are on the high-end of the range of elasticities, especially so when allowing larger drops in earnings when moving from full-time to part-time work. In fact, these estimates are larger than all but the estimates for the 1986 reform. Kleven argues that, given the small size of the change in average tax rates that women faced during that reform, it might be uniquely prone to larger elasticity estimates driven by small degrees of bias in the estimation of the effect of the reform. These mobility elasticities are higher than that found by Kosonen and Matikka (2023), looking at the response of students in Norway. However, this is not necessarily surprising, given that single mothers have generally been found to have higher labor supply elasticities than other populations.

[33]These elasticities were calculated by Kleven (2024), based on estimates from Schanzenbach and Strain (2021).[34] There is one estimated elasticity from the 2009 EITC reform that I omit, which was -0.03, though work by other researchers has found small effects for that period (Bastian and Lochner, 2022, Bastian and Jones, 2021).

2.8 Marginal Value of Public Funds

Finally, I estimate the MVPF of the CalEITC, following the methodology outlined in Hendren and Sprung-Keyser (2020). The MVPF is a form of cost-benefit analysis that measures how much welfare (in dollar terms) can be delivered for a given dollar of government spending. Importantly, this approach takes into account the fiscal externalities generated by the observed behavioral responses to a given policy: for example, how changing labor supply impacts government tax revenue, or how a health intervention on children might impact government budgets via long-term effects on future utilization of government services. The definition of the MVPF is expressed in Equation (2.4).

$$\text{MVPF} = \frac{\text{Willingness-to-Pay}}{\text{Program Cost}} \tag{2.4}$$

I estimate the numerator of Equation (2.4) by assuming that tax payers value one dollar in a refundable tax credit at \$1, and, via the envelope theorem, that individuals who change their labor supply in response to the CalEITC are – to a first-order approximation – indifferent between working part time and their counterfactual employment status, inclusive of the welfare effects of government benefits.[35] Therefore:

$$\text{Willingness-to-Pay} = \text{CalEITC}_{\text{D}} - \text{CalEITC}_{\text{B}} \tag{2.5}$$

Where $\text{CalEITC}_{\text{D}}$ is the direct cost of the CalEITC program between 2015 and 2017, and

[35]This second assumption is non-trivial: recent work has explored the potential for justifications for EITC-type programs based on correcting for non-optimal behavioral responses to labor markets. For example, Lockwood (2020) argues that EITC programs are optimal in settings when present-biased workers choose sub-optimal levels of labor supply that the EITC can correct for. In this context, a social planner might value the behavioral response of taxpayers at some non-zero value.

CalEITC_B is the cost of the CalEITC due to behavioral responses of workers, or the value of the CalEITC recived by workers who change their labor supply in order to receive the credit. The former comes from the California Franchise Tax Board (2021), and I estimate the latter using my estimates of the labor supply effects of CalEITC.

I estimate the denominator of Equation (2.4) by calculating the programmatic cost of the CalEITC, inclusive of fiscal spillovers from other programs. In other words, if the labor supply effects induced by the CalEITC induce an change in net revenue, those costs should be incorporated into the total cost estimate along with the direct cost of the CalEITC. To estimate these fiscal spillovers, based on my estimated coefficients, I draw a random sample of single mothers working part-time in the CalEITC earnings range to represent tax units who changed their labor supply in response to the CalEITC. For this sample, I simulate their federal and state taxes, including federal Individual Income Taxes (IITs), federal Payroll Taxes, and California IITs both with and without the assumed labor supply change, and take the difference between the two as my estimate of the fiscal spillover of the CalEITC.[36] Importantly, I do not attempt to estimate either the fiscal spillovers from other social programs or the longer-term effects of the policy due to data limitations.[37] Importantly, to avoid double-counting, when accounting for changes in California IIT revenue, I exclude the changes due directly to the CalEITC. Therefore, I estimate the total program cost as follows:

[36] In this estimation procedure, after selecting my random sample, I hold their demographics and unearned income constant, but assign each individual the level of earned income that would maximize their CalEITC receipt.

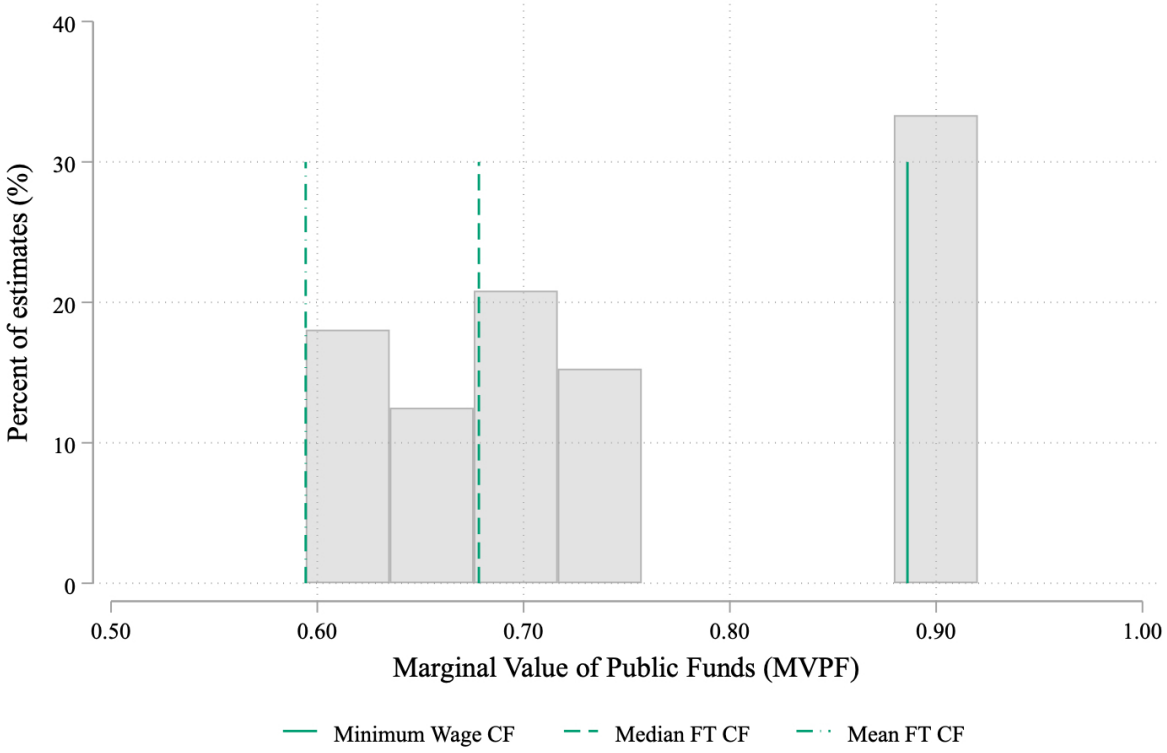
[37] The direction of the bias of these omissions is ambiguous. First, while prior research has found that the federal EITC can increase the long-term outcomes of children, much of that benefit came from the labor supply effects (Bastian and Michelmore, 2018). It is unclear what the effect of the long-term shift from full- to part-time work might be. Second, it is also unclear how the labor supply transition induced by the CalEITC might impact eligibility for other social safety net programs. Much of this would depend on whether the earnings of part-time workers is low (high) enough to gain (lose) eligibility for these programs. Further, the fact that the full-time to part-time response is concentrated in households with three or more adults makes it unclear if the women's eligibility would change for any program whose eligibility is determined at the household, rather than individual, level.

$$\begin{aligned}\text{Program Cost} = \text{CalEITC}_D - \text{Federal IIT Tax Changes} \\ - \text{Federal Payroll Tax Changes} - \text{California IIT Tax Changes}\end{aligned} \tag{2.6}$$

I produce a range of estimates, varying several aspects of the estimation process as well as the specifications of the estimated behavioral effect, as discussed in Section 2.5.3. Importantly, I consider the same three potential estimates of the counter-factual annual earnings as used in Section 2.7: the implied wage for a full-time worker earning the minimum wage, and the median and mean values based on 2014 data. Across all of these estimates, I estimate a MVPF of between 0.6 and 0.9 (see Figure 2.11). This implies it costs the government between \$1.10 and \$1.6 to provide \$1 in benefits to low-income workers via the CalEITC. I highlight three estimates in this range, denoted by the vertical lines on the figure. Each is based on the specification with all controls, using my preferred sample, but varies the counter-factual annual income level of a full-time worker. The MVPF is lower when I assume that the counter-factual full-time level of income is higher, driven by the fact that the federal government loses more revenue when a higher-earner reduces their labor supply.

Across all models, the MVPF is less than 1 because the behavioral effect of the CalEITC reduces federal revenues from payroll taxes and the IIT. I break down this effect in Figure 2.12, where I plot the net effect on the different sources of tax revenue under three different scenarios, based on different counter-factual levels of annual earnings. Note that across all cases, state individual income taxes play almost no role. In the first case, where the representative full-time worker who transitions to part-time work due to the CalEITC would have earned just above

Figure 2.11: Distribution of Marginal Value of Public Funds estimates



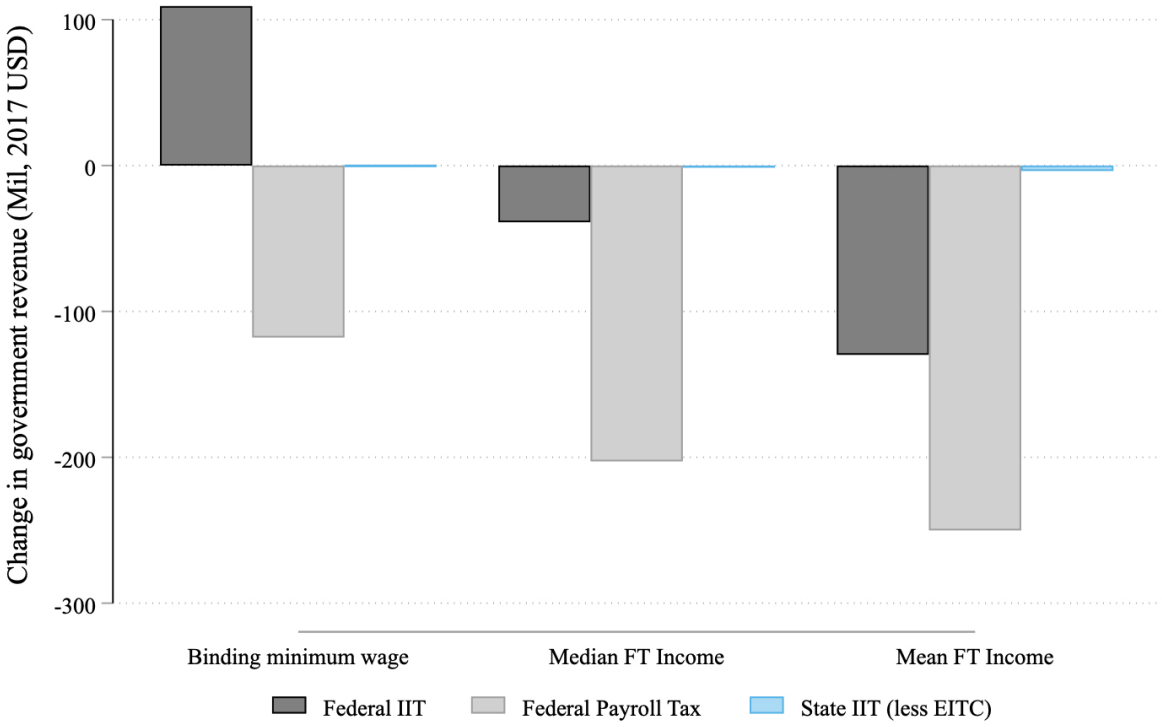
Note: This figure plots the distribution of Marginal Value of Public Funds (MVPF) estimates, calculated as described in Section 2.8. The different estimates vary in specifications of the estimates of the employment effects of the CalEITC, the types of heterogeneity of effects allowed (none, by year, and by count of qualifying children (QC)), the assumed counterfactual full-time level of earnings (the annual earnings of a full-time worker, working year-round at their binding local minimum wage, plus the median and mean annual earnings of a single mother in California in 2014 [in 2017 USD]). The numerator of the MVPF is always the full 2015-2017 cost of the program less the increase in CalEITC receipt driven by the observed behavioral change. The denominator is the full 2015-2017 cost of the program plus any fiscal spillovers due to behavioral changes effecting the federal or state tax system. The vertical lines represent three representative models, all estimated with heterogeneity across the number of QC, with demographic and economic controls, but iterating over the counter-factual full-time earning level.

$20,000, the decline in federal revenue from payroll taxes (the grey bar) is offset by increasing revenue from the federal IIT. This is because of a decline in their eligibility for the Federal EITC: if a worker had been working full-time, earnings approximately $22,000 annually, then they would receive close to their maximum allowed federal EITC credit (conditional on their number of qualifying children). Therefore, moving from full-time to part-time work would reduce their federal EITC amount, increasing federal IIT revenue. Alternatively, if a worker was earning over $30,000, then the taxpayer has already moved far enough along the phase-out of the federal EITC to mean that transitioning between full-time and part-time work actually increases their federal EITC receipt, resulting in a loss of federal IIT revenue on top of payroll tax revenue.

2.9 Conclusion

In this paper, I examine the labor supply response of single mothers to the implementation of the CalEITC. I find robust evidence that the CalEITC induced a shift from full-time to part-time work of approximately 45,000-60,000 single mothers between 2015 and 2017. This finding is robust to a range of specifications and methodologies, and is supported by a matching decline in annual earnings. This labor supply response is consistent with a discrete, rather than continuous, change in labor supply, where workers respond to changes in incentive by moving between a limited number of potential working arrangements. In this setting, this is best represented by workers facing a limited set of choices regarding how many hours to work per week. Using these results, I estimate that single women have a mobility elasticity of between -1.1 and -1.6, which are large elasticities relative to the existing EITC literature. Further, I find that the CalEITC has a MVPF of between 0.6 and 0.9, indicating it costs more than a dollar to the government (Federal

Figure 2.12: Implied fiscal spillovers of the CalEITC under different assumptions



Note: This figure plots the fiscal spillovers via the Federal Individual Income Taxes (IIT), federal payroll taxes, and state IIT (less the CalEITC), that are incorporated into the Marginal Value of Public Funds (MVPF) estimates, calculated as described in Section 2.8, across three different assumptions regarding the counter-factual level of earning of a full-time worker who responds to the CalEITC. In each case the estimated behavioral effect on full- and part-time work comes from a model including a full set of demographic and economic controls (minimum wages and unemployment interacted with the number of qualifying children), accounting for heterogeneity across the number of qualifying children. Each bar expresses the change in US government revenue in real 2017 USD. The three counter-factual levels of earnings are the annual earnings of a full-time worker, working year-round at their binding local minimum wage, plus the median and mean annual earnings of a single mother in California in 2014 (in 2017 USD).

and state) to provide $1 in benefits to CalEITC recipients.

I would highlight two important limitations of this paper, both stemming from the ACS data. First, the ACS data is not a longitudinal sample, meaning that I cannot track single mothers across years. This would be important to understanding the labor-force attachment and earnings of these women prior the CalEITC, as well as the duration of their part-time work spells. There may be different welfare implications of permanent verses temporary part-time work spells that would be worth exploring. Second, the earnings data in the ACS are likely a biased and noisy measure of the true taxable earnings of workers. This makes estimating any number of parameters difficult, but the most important of these is the counter-factual income of workers who would have been working full-time but for the CalEITC. This is a key parameter in determining the welfare impacts of the CalEITC. Both of these issues could be address by turning to higher-quality, longitudinal administrative earnings records, and I believe this is the most fruitful avenue for future work related to this credit. These data would both have the appropriate reference period (the calendar year) and match what appears on an individuals tax return. A further benefit of these data would be the ability to utilize the panel nature of most of these sources, allowing direct estimation of movement from one level of annual earnings to another. This would pin-down more precisely the dynamics of the transition induced by the CalEITC. Ideally these records would also have information on the utilization of other social safety net programs, which could be incorporated into the MVPF estimate.

I also find that the full-time to part-time transition is concentrated among large, multi-adult households, indicating an important role for intra-household bargaining in allowing women the flexibility to respond to the incentives of the CalEITC. Further work on the CalEITC inform our understanding of household labor supply decisions; namely, if characteristics outside the tax unit

but inside the household inform the elasticity of a given worker. For example, if the CalEITC induces a shift in parent time use, a possibility explored in Table A.12, this could impact the labor supply decisions of other household members or the longer-term outcomes of children, questions that have been explored by others in the context of the federal EITC (Bastian and Lochner, 2022, Bastian and Michelmore, 2018).

Chapter 3: Elderly State Income Tax Incentives and Migration: Evidence from Administrative Tax Data

Abstract[1]

Across both domestic and international contexts, national and sub-national governments have crafted tax policies to attract or retain residents, motivated at least in part by the belief that individuals and families move between jurisdictions to minimize their tax burden. One common example of such policies in the United States are state-level elderly income tax benefits. This research uses confidential administrative tax records from 1999 through 2019 and a range of empirical approaches to estimate the elasticity of interstate elderly migration in response to these tax preferences. We find that elderly filers are marginally more responsive to overall after tax rates than their younger counterparts, but find no evidence of migration responses to elderly-specific tax incentives. These results align with findings of the prior literature looking at the pre-2000s period. Existing elderly-specific tax incentives seem not to serve as a means to attract or retain elderly residents, and rather operate as a transfer to and within the elderly population.

[1]Co-authored with Karen Conway (UNH) and Jonathan Rork (Reed College). These results are preliminary. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors and do not necessarily reflect the views or the official positions of the U.S. Department of the Treasury or the Internal Revenue Service. All results have been reviewed to ensure that no confidential information is disclosed. We would like to thank Kevin Pierce and Amanda Eng at the Statistics of Income Division for their invaluable help on this project. We would also like to thank the seminar participants at the U.S. Treasury's Office of Tax Analysis and the 2024 National Tax Association Annual Meeting, especially Gopi Shah Goda, for their comments.

3.1 Introduction

Our research addresses the confluence of two growing phenomena; the aging of societies and the competition among governments to use tax policies to attract or retain residents. Recent work has established that certain potentially mobile groups–such as star scientists and inventors (Akcigit et al., 2016, Moretti and Wilson, 2017), athletes (Kleven et al., 2013) and veterans –are fairly responsive to income tax differentials, but the responsiveness of the rich more generally is less clear (Kleven et al., 2014, Agrawal and Foremny, 2019, Moretti and Wilson, 2023, Young et al., 2016, Rauh and Shyu, 2019). The elderly (ages 65+) and near-elderly are increasingly more affluent than the general population (Conway and Rork, 2016) and are potentially quite mobile, especially once they are retired and no longer have labor market ties. The elderly are also the fastest growing age group in many economies. The decade of 2010-20 marked the largest ever numeric or percentage gain in the elderly, and one in six persons in the US (16.8%) was over age 65 in 2020.[2]. Governments appear to have noticed, as tax breaks targeting the elderly and retired have expanded both within the US (e.g. Brewer et al., 2017) and across countries (e.g. Kalin et al., 2024). As the size of this group grows, so do the revenue and other consequences of these policies. Yet research on tax-induced US elderly migration mostly ends with the 2000 US decennial Census long form (e.g. Conway and Rork, 2006, 2012a), before the retirement of the Baby Boomers and the enactment of many generous income tax breaks. Those earlier studies also had at best limited information on higher-income groups, who have been the target of many of these policies.

We fill those gaps by using confidential administrative tax records from 1999 through 2022

[2]See this discussion of the issue by researchers at the US Census Bureau: https://www.census.gov/library/stories/2023/05/2020-census-united-states-older-population-grew.html.

and a range of empirical approaches to estimate the elasticity of interstate elderly and near-elderly migration in response to these tax preferences. Besides providing an update, these data offer several critical advantages over past work. First, the tax data provide a much larger sample (e.g., the full universe of taxpayers versus a 1 or 5% population sample), which is essential to accurately observe low-frequency events like elderly interstate migration (at approximately 1% a year). This is especially important when it comes to observing particular sub-groups, like the wealthy. Second, the detailed income data included in tax records allows us to observe both retirement income, which is under-reported in survey data (Bee and Mitchell, 2017, Brady and Bass, 2021) and the income characteristics of the very wealthy, which are obscured by top coding. Given these advances and how little is known about recent US elderly migration, one key contribution is documenting the patterns of and changes in interstate elderly migration, for different income groups, over the last two decades and contrasting them to their younger counterparts.

Our main empirical analyses combine this rich migration data with state income tax policies and take a multi-pronged approach to exploring the impact of these tax policies on location choices. Our main estimation uses the entire sample of states and years and Poisson Pseudo-Maximum Likelihood (PPML) to estimate how state-to-state flows change in response to state income tax changes. Consistent with past work, our results indicate that elderly tax breaks have a minimal effect on elderly migration, a result that holds up across a wide range of model specifications. Future work will supplement this approach with an investigation of several highly visible and potentially impactful specific policy changes occurring in different states at different times, using a difference-in-difference approach.

The rest of the paper is organized as follows. Section 3.2 discusses what is known about interstate migration, especially for the elderly and in response to taxes, and then provides a brief

primer on the different types of tax breaks enjoyed by elderly US taxpayers. Section 3.3 describes the confidential, administrative data we use and the nuances of creating migration measures, as well as the construction of state income tax measures and other characteristics that may influence that migration. Section 3.4 presents the results from the full panel, migration-flow PPML analyses. Section 3.5 describes the proposed analytical approach of the DiD investigations, Section 3.6 estimates the overall fiscal cost of elderly-specific tax policies polices, and Section 3.7 concludes and discusses future directions for this work.

3.2 Background

Migration research tends to focus on the working population; how it responds to wage and housing market differences and the consequences a mobile population has on economic efficiency and growth (e.g., see Kleven et al. (2020) and Jia et al. (2023) for recent surveys). Of particular interest has been exploring reasons for the observed decline in U.S. interstate worker migration (e.g. Molloy et al., 2017) and a greater emphasis on the geography of labor markets, such as commuting zones (CZs) or census-based statistic areas (CBSAs) (e.g. Bayoumi and Barkema, 2019, Olney and Thompson, 2024). Because most elderly have left or are soon leaving the labor market, wages are not a factor and may lead to geographic sorting by age Komissarova (2022).

The emergence of millionaire taxes in the U.S. and changes in wealth and/or estate taxes around the world has led to a growing focus on the migration behavior of the very rich, beginning with Bakija and Slemrod (2004) who found statistically significant but small impacts of income and estate taxes. Young and Varner (2011) find a minimal impact of the New Jersey millionaire tax overall but higher levels for those of retirement age or receiving primarily investment income.

Subsequent research of the millionaire tax in California finds more mixed results Varner et al. (2018), Rauh and Shyu (2024), while Young et al. (2016) in a nationwide investigation including a cross-border analysis finds small impacts in income tax policies among millionaires. In their study of billionaires, Moretti and Wilson (2023) find they are sensitive to estate tax rates, especially as they age.[3] Other rich and likely mobile individuals, such as star scientists and football stars, are also found to respond to tax differentials Akcigit et al. (2016), Moretti and Wilson (2017), Kleven et al. (2013).

The elderly are both disproportionately high income and have characteristics that might imply they are highly mobile; not only do they have weak ties to the labor market but in the US they are perhaps the only group whose primary health insurance (Medicare) is also not sensitive to job or location. Yet, research on their tax-induced migration mostly precedes the 2000s and the retirement of the Baby Boomers. That federal and state estate taxes have been dramatically reduced since 2001 suggests that income taxes may become increasingly important for all but the very rich. Several state tax policies exist that target the elderly specifically and reach much further down into the income distribution, to the much larger groups of moderate- and higher-income households. We therefore investigate the interstate migration behavior of the elderly and near-elderly over a range of income levels, not just the rich.

[3]Numerous studies examine tax-induced migration of the rich at the country-level, mostly within Europe, including Kleven et al. (2014) (Denmark), Jakobsen et al. (2024) (Denmark and Sweden), Agrawal and Foremny (2019)(Spain), Advani et al. (2023) (UK), and especially for Switzerland (e.g. Brülhart and Parchet, 2014, Brülhart et al., 2022, Köthenbürger et al., 2023). See Kleven et al. (2020) and Advani and Tarrant (2021) for summaries and critical reviews.

3.2.1 What do we know about inter-state (Near-) Elderly Migration?

Interstate migration tends to decline with age (Molloy et al., 2011), perhaps with a slight uptick among the near-elderly (e.g., ages 55-64) as they approach retirement and no longer have children at home. Within the elderly population, migration behavior further changes as one ages. Near-elderly and younger elderly tend to move for amenities (weather and environment, cost of living, tax policies) but as they age they often move to get assistance as their health and finances decline, sometimes "returning home" to where they once lived and their children still reside (Conway and Rork, 2011). This pattern can lead to a "double whammy", where amenity-rich states recruit and benefit from the younger, healthier and richer elderly with the presumption that many will return home before drawing on their health care and income support systems. Migration into inter-state elderly migration became more difficult after the elimination of the decennial census long form in 2000. This is because migration is a relatively rare event (1% a year) and the geographic patterns are quite stable over time. A large dataset is therefore necessary to accurately identify any policy-induced changes in elderly migration. The remaining publicly available datasets are inadequate due to either sample size or measurement issues (see Conway and Rork (2016) for further discussion). For example, the widely used American Community Survey (ACS) is a 1% sample and therefore observes a relatively small number of moves, especially if broken down further into age and income categories. Its definition of residence combined with the year-round interview process introduces substantial noise in identifying true moves; top-coding of income further limits identifying the very high income. Another potential data-source, the IRS-SOI migration flows are not broken down by age and income, and, as discussed below in Section 3.3, potentially face other data issues over our sample period.

As a result, what we know about tax-induced interstate elderly migration comes primarily from Conway and Rork (2006, 2012b) and Bakija and Slemrod (2004), all based on migration trends preceding the 2000s. That research finds limited, at most, responsiveness to income and estate taxes among the elderly. However, both the US elderly population and tax policies have changed since 2000, and a recent study by Kalin et al. (2024) calls these earlier findings into question. These authors investigates the effect of a Portugal policy that granted a full tax exemption of foreign pensions for retirees moving there and finds "substantial international mobility responses to the reform," (Kalin et al., 2024). The largest effects come from the most educated from the highest income countries. Evidence of strong agglomeration and spatial clustering may amplify the process and subsequent consequences. Further confirmation comes from the estimated reduction in out-migration from Finland once it regained the ability to tax pensions at the source. Kalin et al. (2024) go on to explore the implications for optimal tax policy and conclude that targeting certain groups may be justified and can lead to "permanent changes in location patterns."

Maroto et al. (2024) provide additional evidence of the substantial positive consequences of induced elderly migration using data from France, while Komissarova (2022) uses a spatial equilibrium model to estimate the welfare gains from geographic sorting into work and retirement locations. The consequences to elderly individuals, such as life expectancy, may also be substantial Finkelstein et al. (2021). However, state income tax breaks for the elderly, at least for the high-income elderly, are found to have little impact on state economic growth Brewer et al. (2022). This raises the question of whether elderly in the US respond to state income tax policy. The lack of evidence for the US alongside the potential consequences of such migration underscore the contributions of this research.

3.2.2 State Income Tax Policies Targeting the Elderly

Every state in the United States with a personal income tax has at least one provision that explicitly benefits elderly taxpayers. The federal personal income tax also benefits elderly taxpayers but generally to a lesser extent. The three main types of income tax breaks are: 1) additional exemptions or deductions given on the basis of age, 2) the reduced taxation of social security benefits, and 3) favorable tax treatment of "retirement income", which is often quite broadly defined. Conway and Rork (2016) describe these policies and their potential distributional and revenue consequences in detail, while Conway and Rork (2012a) provide the historical context behind them. This discussion draws heavily from that research.

In Table 3.1, we report the status of state-level elderly preferences in 1999, the first year of our dataset, and 2023. First, age exemptions and deductions tend to be the most modest and stable among the three policies; three-quarters of states with personal income taxes offer them and the number of states remains the same throughout our sample period.[4] However, even here, there is substantial variation across states, as well as complexity in how these policies are applied. For example, several states allow taxpayers to choose between taking an aged-based exemption or a pension-income exemption. Another nuance is that an extra standard deduction (e.g., the current federal policy) only benefits non-itemizes. All of these features, combined with their interaction with the existing state tax rates, mean that simply estimating the effect of the existence of an age-specific deduction, exemption, or credit does not capture the full effect of these policies. This is a major reason we, in Section 3.3.2, construct our summary measure of the elderly tax benefit in each state.

[4]Note that Tennessee eliminated its limited income tax in 2021. The numbers in Table 3.1 therefore report for 44 states in 1999 and only 43 states in 2023.

Second, in the years prior to the 1983 federal legislation that began taxing a portion of Social Security benefits (SSB) for high income households, SSB was tax-free at the federal and state level.[5] The majority of states did not follow these reforms so SSB continue to be tax-free for all households. Of those states that did begin taxing SSB, many offer more generous treatment and none tax more than the federal policy. Moreover, many states subsequently eliminated or greatly reduced those taxes. Table 3.1 reveals that 1 state eliminated taxation on SSB during our main sample period, and five adjusted their taxation of SSB to be more generous than that limit imposed by the federal individual income tax system.

The most active and potentially generous state policy is the treatment of "retirement income". These policies also tends to be more complex, as they can vary across multiple dimensions, including the size of the exemption, whether the benefit phases out with income, and even the definition of retirement income, including, importantly, whether or not SSB is included in the definition of pension income. Over our sample period, there has been a wide array of changes across these dimensions, including three states (Connecticut, Iowa, and New Hampshire) moving to fully exempting pension income, and another three states (Georgia, Michigan, and New Jersey) increased their exemption amounts to over $70,000. The definition of retirement income in these credits can greatly increase the generosity of these tax breaks, as they can include types of investment income that are not tied to age or retirement. Overall, Table 3.1 shows substantial movement in these policies–both expansions and contractions–since 1999.

[5]Specifically, the amount of SSB subject to tax are based on "combined income", which equals the sum of gross income, tax-exempt interest, and 50% of SSB. (Single/married) households above the $25,000/32,000 threshold must pay tax on $.50 of every dollar that their combined income exceeds the threshold up to a maximum of 50% of SSB. The second thresholds ($34,000/44,000) operate similarly, taxing SSB at a rate and maximum of 85%. These thresholds are not indexed for inflation and so reach a wider set of households each year.

Table 3.1: Elderly preferences in state individual income tax systems

Policy Type			1999 # of States	2023 # of States
Extra Exemption, Deduction or Credit Based on Age				
	Offered		33	32
	Not Offered		11	11
Social Security Benefits				
	Fully Exempt		31	32
	Partially Exempt		13	11
		More favorable than federal limit	3 (CT, IA, WI)	9 (CT, KS, MI, MO, NE, NM, RI, UT, VT)
		Same as federal limit	10 (CO, KS, MN, MO, MT, ND, NE, NM, UT, WV)	2 (MN, MT)
Pension & Other Retirement Income				
	Fully Exempt		5 (AL***, HI, IL, MS, PA)	8 (AL***, CT*, HI, IA, IL, MS, NH, PA)
	Not Exempt		13	14
	Partially Exempt		26	21
		Exempt up to $10,000	8 (IA, MO*, MT*, NC, NJ, NM*, OK*, WV)	2 (MT*, WI*)
		Exempt: $10,001 - $20,000	5 (AR, LA, MN*, OR* UT*)	6 (AR, LA, MO*, OK, OR*, SC)
		Exempt: $20,001 - $40,000	8 (CO, DE, GA, MD**, NY, SC, TN*, VA)	4 (DE, NY, RI*, VA)
		Exempt: $40,001 - $60,000	0	1 (CO)
		Exempt: $60,001 - $70,000	1 (MI)	3 (KY, ME**, MD**)
		Exempt: $70,001 +	1 (KY)	3 (GA, MI, NJ*)
	Credit for pensions			
		Exempt up to $10,000	3 (IN*, OH*, WI*)	2 (OH*, UT*)

Note: This table is derived from the Bakija (2019) (for 1999) and the Institute for Taxation and Economic Policy (for 2023). The table includes all 50 states plus D.C. In 1999, 41 states plus D.C. levied an income tax whereas NH and TN had more limited income tax systems, for a total of 44. TN eliminated its limited income tax in 2021. Amounts given are the maximum amount possible for a married household; singles are generally half of the amount. * indicates an AGI cutoff, after which the benefit no longer applies. ** indicates that SSB count toward the pension exemption. *** indicates that pensions are defined benefit plans only.

3.3 Data

We utilize confidential data from the IRS Statistics of Income (SOI) Division, drawn from both Federal tax returns (Form 1040s) and information returns. The primary data are drawn from the SOI Databank (Chetty et al., 2018), which is a de-identified panel of all U.S. individuals, linked to their F1040s and information returns, as well as information from the Social Security Administration regarding their date of birth, date of death (if applicable), and gender. Recently extended through 2022, this dataset allows us to observe, for each individual, demographic and income characteristics, as well as their geographic location, defined as zip codes. We combine data from the SOI Databank with information from other IRS sources when required; for example, with information from retirement (F1099-R, F1099-SSA) or investment income (1099-INT, Schedule D) tax forms. The details of the construction of our main datasets are contained in Appendix B.2.

The SOI Databank identifies, for each tax year and individual, two potential locations: one based on the address on their tax return and the second based on the most common (modal) location across the universe of information returns associated with the individual. Prior research, as well as the the official IRS migration figures, have most often used address information from tax returns (Young et al., 2016, Pierce, 2015), and many other researchers using IRS data for non-migration purposes have done likewise (for example, Chetty et al. (2013) use F1040 zip codes to examine how the labor supply effects of the Earned Income Tax Credit (EITC) vary by neighborhood-level knowledge of the EITC). However, other researchers have relied on the information returns (Yagan, 2019). We choose to use the most common information return zip code associated with a given tax return due to a potential data issue with the tax return addresses

during the 2014-2017 period, which will be discussed in depth in future research.[6]

We construct two different datasets to track migration in this project, each based on the SOI Databank. The first is a set of state-by-state flows, which record the count of individuals who moved between state o (origin) and d (destination) for a given pair of years. For each year t, we take the sample of individuals who filed taxes in year t and $t-1$, who were alive in both years, who were not filing as a married filing separately return, and for whom we observe the requisite demographic, geographic, and income characteristics (namely, adjusted gross income (AGI)). We also drop all individuals who, at any point in time, were flagged as associated with a fraudulent return, as the act of committing fraud can resemble a move in our data.[7] After assigning each individual to cells based on their age, marital status, and year $t-1$ AGI brackets, we collapse the dataset to construct a count of individuals moving between each pair of states within a given demographic cell.[8] Importantly, this means for each pair of states, we have two flows (from New York to Florida, and from Florida to New York). The basic distribution of this sample by age group and AGI bracket is reported in Table B.1. Second, we construct an individual-level sample, composed of a stratified random sample: 100 percent of all individuals ever on a tax return with real AGI (2019 USD) over $1 Million, 50 percent of all remaining individuals ever on a tax return with real AGI (2019 USD) over $500,000, and 20 percent of the remaining population.[9] We restrict this sample in the same way as with the flows, dropping individuals associated with

[6] If an individual has no information returns, we do use the F1040 information, but this is fairly rare. In future work, we intend to consider the same impacts using solely information return zip-codes, and solely tax return zip codes.

[7] Fraudulent returns are potentially problematic because the act of stealing and using an individual's social security number to claim their tax refund often appears as a move. If an individual living in New York has their identify stolen, and an individual in Florida files a fraudulent return in their name, it will appear as a New York to Florida move. This was a particular problem in the early 2010s.

[8] We use fixed, inflation-adjusted bins instead of percentiles because the affluence of the elderly has grown during this time period relative to other age groups. This approach makes the groups more comparable over time.

[9] Randomization is based on the universe of social security numbers in the SOI Databank, meaning that the randomization is not conditional on filing taxes, or any other characteristics.

fraudulent returns, those with missing data, and those who are not observed in both year t and $t-1$. From this, we construct a state-level panel we use for our individual-level DiD design.

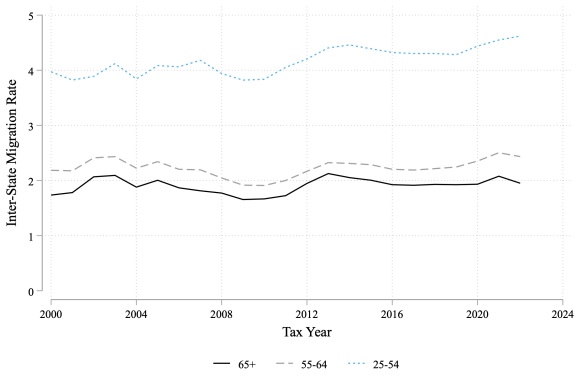
3.3.1 Interstate Elderly Migration in the 2000s

Aggregate measures of these state-to-state migration flows are more useful for investigating the salient patterns and meaningful changes in elderly migration since 2000. The first measure is the overall migration rate; i.e., the number of people moving across state lines in a given year divided by the total number of people. Figure 3.1a plots this rate for the three main age groups, aged 25-54, 55-64 and 65+. Immediately apparent is the well-known tendency for migration to decline with age and the stability of these rates over time. As documented elsewhere (Molloy et al., 2011), migration declined for all age groups in the years around the Great Recession.

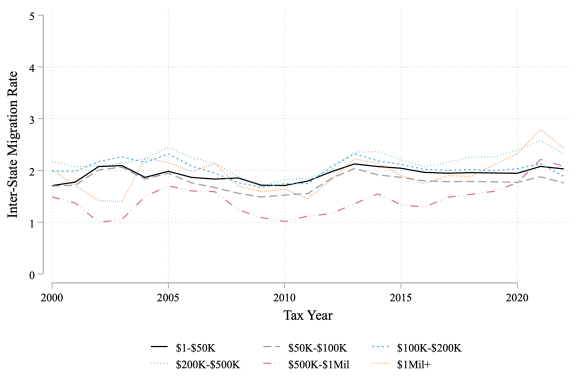
A notable difference, however, is that the migration rates for all three groups are considerably higher than what appears from other sources (e.g. Conway and Rork, 2016, Molloy et al., 2011). The administrative tax data yields an elderly migration rate that is approximately double that of the 1% annual rate (or less) often reported. While our sample is limited to the universe of tax filers, it represents 79 percent of the US population and thus that limitation seems unlikely to explain this discrepancy.[10] This is likely due, at least in part, to the use of information return addresses: as observed by the creators of the SOI Databank, the average level of migration across all age groups is higher when using information return addresses rather than the addresses associated with the tax return. However, it is also true that levels of migration are higher using any measure of location derived from tax records than when using survey-derived instruments Foster

[10] In 2010, the US adult population (over 25) was 203.9 million. In that same year, our sample included 151.2 million observations.

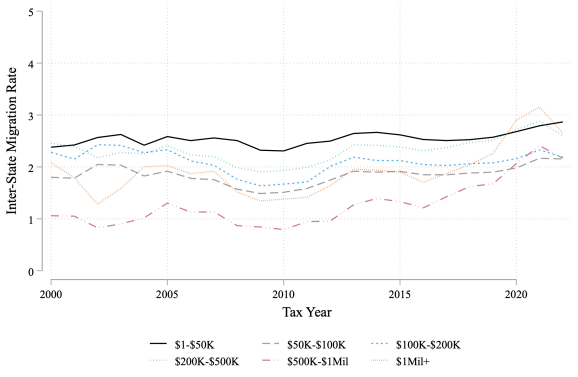
Figure 3.1: Inter-state migration by age and and AGI: 2000-2022



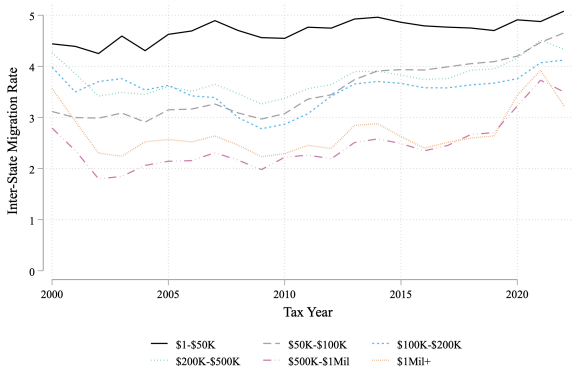
(a) All ages (25+)



(b) Elderly (65+)



(c) Near-Elderly (55-64)



(d) Young (25-54)

Note: Data for these figures are derived from the SOI Databank, from 1999-2022. The sample is limited to individuals who, for a give year t filed taxes in year t and $t-1$, for whom we observe their corrected adjusted gross income (AGI) and who report positive AGI in year t, who did not file a Married Filing Separately return, and who is not matched to an individual associated with a fraud claim. AGI brackets are in 2019 USD.

(2023).

The migration patterns by AGI differ between younger and older (See Figure 3.1). The most mobile groups for the young (Figure 3.1d) are those with the lowest incomes, potentially capturing the youngest (post-college age) who are the most mobile and have the lowest incomes. For the near-elderly groups the difference is less stark (Figure 3.1c) and changes completely for the elderly (Figure 3.1b). Here, the lowest rates of migration are for the highest AGI groups ($500,000 to $1,000,000 and $1,000,000+). Notably, however, they have been increasing since 2016 with sharp increases during 2019-21, perhaps due to the Covid-19 pandemic. The pandemic disrupted many facets of location choice, which is why our main analysis ends in 2019.

Both the higher rates of migration and strong uptick observed in recent years imply that the existing wisdom – that these tax breaks have little impact – deserves revisiting. We begin by exploring where these individuals are moving using the net in-migration rates for each state (in-migrants minus out-migrants, divided by the total number in the state). Table 3.2 lists the net in-migration rates for every state at the start, middle and end of our main analysis period (2000-2019). It shows stable patterns; the top 'importing' states remain largely the same; the big players are Florida, Arizona, Nevada and increasingly South Carolina,–e.g.., those with an annual rate of 1 percent or more–all amenity-rich states. Other more surprising big importers are Delaware, Maine, South Dakota and Idaho; the rest are mostly central mountain (MT, OR) or southeastern (NC, TN, GA) states. The top exporters are also fairly stable, with Alaska, D.C., New York, New Jersey, Illinois and California consistently in the group. Put simply, few big changes emerge from this table.

To explore this stability more formally, we turn to the state-to-state migration flows and a basic regression:

Table 3.2: Net in-migration rates by state, year, and age-groups.

	Near-Elderly (55-64)			Elderly (65+)		
	2000	2010	2019	2000	2010	2019
Alabama	0.2	0.3	0.6	(0.0)	0.3	0.4
Alaska	(0.1)	(1.1)	(1.1)	0.2	(1.0)	(0.9)
Arizona	2.6	1.0	2.2	1.0	1.0	2.1
Arkansas	0.4	0.7	0.8	(0.2)	0.1	0.3
California	(0.8)	(0.4)	(0.8)	(0.3)	(0.2)	(0.8)
Colorado	0.0	(0.0)	(0.2)	0.1	0.2	(0.2)
Connecticut	(0.9)	(0.6)	(1.0)	(0.2)	(0.3)	(0.7)
Delaware	1.1	1.1	2.0	0.5	0.4	1.4
District of Columbia	(1.4)	(0.7)	(1.4)	(1.2)	(1.2)	(1.0)
Florida	4.3	1.2	1.8	0.9	0.7	1.4
Georgia	0.1	0.1	0.0	0.4	0.4	0.3
Hawaii	0.2	0.1	(0.2)	(0.2)	(0.1)	(0.3)
Idaho	0.9	0.6	2.0	0.6	0.6	1.8
Illinois	(0.7)	(0.4)	(1.0)	(0.3)	(0.4)	(0.9)
Indiana	(0.4)	(0.2)	(0.1)	(0.1)	(0.1)	(0.1)
Iowa	0.0	(0.1)	(0.4)	0.0	(0.3)	(0.3)
Kansas	(0.2)	(0.1)	(0.1)	0.0	(0.0)	(0.1)
Kentucky	0.3	0.1	0.1	0.2	(0.1)	0.1
Louisiana	(3.5)	0.3	(0.3)	(1.2)	(0.1)	(0.2)
Maine	1.1	1.1	0.4	0.6	0.7	0.2
Maryland	(0.5)	(0.2)	(0.9)	0.1	(0.2)	(0.7)
Massachusetts	(0.5)	(0.3)	(0.7)	(0.1)	(0.2)	(0.6)
Michigan	(0.5)	(0.4)	(0.2)	(0.3)	(0.2)	(0.3)
Minnesota	0.1	(0.9)	(0.3)	0.1	(0.4)	(0.2)
Mississippi	0.2	0.3	0.1	0.3	0.2	0.1
Missouri	0.5	(0.0)	0.0	0.2	(0.1)	(0.1)
Montana	0.6	0.5	0.8	0.1	(0.0)	0.2
Nebraska	(0.3)	(0.1)	(0.3)	0.0	(0.2)	(0.2)
Nevada	3.1	0.2	1.8	2.0	0.4	1.3
New Hampshire	0.8	(0.3)	0.4	0.7	0.1	0.2
New Jersey	(0.6)	(0.6)	(0.8)	(0.4)	(0.4)	(0.7)
New Mexico	0.3	0.5	0.6	(0.1)	(0.1)	0.2
New York	(1.0)	(0.5)	(0.9)	(0.7)	(0.6)	(0.9)
North Carolina	1.1	0.6	0.7	0.5	0.4	0.6
North Dakota	(0.5)	0.1	(0.5)	0.2	(0.2)	(0.2)
Ohio	(0.3)	(1.2)	(0.2)	(0.1)	(0.7)	(0.2)
Oklahoma	0.2	0.2	0.1	0.0	(0.0)	0.1
Oregon	0.1	0.3	0.3	(0.1)	0.3	0.2
Pennsylvania	0.0	(0.1)	(0.3)	(0.1)	(0.2)	(0.3)
Rhode Island	0.2	(0.5)	(0.3)	0.3	(0.4)	(0.4)
South Carolina	1.2	3.1	2.1	0.7	2.1	1.6
South Dakota	0.8	0.7	0.9	0.5	(0.1)	(0.0)
Tennessee	0.4	0.7	0.8	0.2	0.3	0.6
Texas	0.0	0.2	0.6	0.1	0.2	0.5
Utah	(0.0)	0.4	0.3	0.3	0.3	0.4
Vermont	0.8	0.1	(0.2)	0.6	0.2	(0.5)
Virginia	0.0	(0.2)	(0.2)	0.3	(0.1)	(0.1)
Washington	(0.0)	0.0	(0.0)	0.1	0.1	(0.1)
West Virginia	0.4	0.1	(0.0)	(0.1)	(0.3)	(0.4)
Wisconsin	0.1	(0.1)	(0.1)	(0.0)	(0.1)	(0.1)
Wyoming	(0.2)	(0.5)	0.5	(0.2)	(0.3)	(0.0)

Note: Data for this table are derived from the SOI Databank, from 1999-2022. The sample is limited to individuals who, for a give year t filed taxes in year t and $t-1$, for whom we observe their corrected adjusted gross income (AGI) and who report positive AGI in year t, who did not file a Married Filing Separately return, and who is not matched to an individual associated with a fraud claim. AGI brackets are in 2019 USD. Net-in-migration is defined for state d and year t as the number of individuals moving into d in year t less the number of individuals who left d between $t-1$ and t, divided by the population of d in $t-1$.

$$\log(1 + Y_{odt}) = \delta_t + \gamma_{od} + \log(p_{ot}) + \log(p_d t) + \log(1 + Y^y_{odt}) + \epsilon_{odt} \tag{3.1}$$

Where δ_t and γ_{od} are year and flow fixed effects, and p_{ot} and p_{dt} represent the origin and destination state populations, making this the well-known gravity model of migration which captures broader geographic trends in population growth. From this, 1 – R-squared shows the amount of variation not explained by persistent patterns (flow fixed effects) or overall time trends (year fixed effects). We can also add the migration flow of the young (y^y_{odt}) to see how much of the remaining variation tracks other groups. We also compute a similar model using the Poisson Pseudo-Maximum Likelihood approach (Correia et al., 2020), to account for potential issues with the log specification (Chen and Roth, 2024).

Table 3.3: R2 values from gravity model.

	Full Sample		$1-$50K		$50-$100K		$100K-$500K		$500K+	
	(1) Basic	(2) Young Flow	(3) Basic	(4) Young Flow	(5) Basic	(6) Young Flow	(7) Basic	(8) Young Flow	(9) Basic	(10) Young Flow
Elderly: Log(1+count)										
N	51,000	51,000	51,000	51,000	51,000	51,000	51,000	51,000	51,000	51,000
R2	0.924	0.925	0.919	0.92	0.871	0.874	0.832	0.836	0.791	0.819
Near-Elderly: Log(1+count)										
N	51,000	51,000	51,000	51,000	51,000	51,000	51,000	51,000	51,000	51,000
R2	0.926	0.927	0.918	0.919	0.892	0.895	0.878	0.879	0.749	0.778
Elderly: PPML										
N	45,310	45,310	44,482	44,482	39,307	39,307	39,859	39,859	7,291	7,291
Pseudo-R2	0.949	0.953	0.93	0.938	0.918	0.921	0.936	0.936	0.875	0.875
Near-Elderly: PPML										
N	48,852	48,852	47,656	47,656	44,229	44,229	44,482	44,482	10,212	10,212
Pseudo-R2	0.952	0.959	0.937	0.947	0.912	0.919	0.935	0.935	0.856	0.857

Note: This table reports the R2 (or pseudo-R2 for the PPML model) estimated via Equation (3.1). Data for this table are derived from the SOI Databank, from 2000-2019. The sample is limited to individuals who, for a give year t filed taxes in year t and $t-1$, for whom we observe their corrected adjusted gross income (AGI) and who report positive AGI in year t, who did not file a Married Filing Separately return, and who is not matched to an individual associated with a fraud claim. AGI brackets are in 2019 USD. Net-in-migration is defined for state d and year t as the number of individuals moving into d in year t less the number of individuals who left d between $t-1$ and t, divided by the population of d in $t-1$.

The results are shown in Table 3.3, and we can see that the regressions yield R-squared (or pseudo-R2) values that are over 0.9 for the full sample and decreasing with the level of income.

For instance, the R-squared value for the full elderly sample is 0.924, 0.919 for the bottom of the income distribution (AGI between $1 and $50,000), and 0.791 for the top of the income distribution (AGI = $500,000+). Adding the migration flow for the young (25-54) does not substantially impact the results. One important consideration is that the number of empty flows increases with income, as the size of the overall sample declines: this is explicitly addressed by the PPML approach, and importantly we see pseudo-R2 values using this approach that are between 0.05 and 0.1 units higher than when estimating the standard fixed effects regression with log(1-y) as the dependent variable. Overall, there is little variation in migration flows that is not explained by persistence and overall time trends, but whatever variation there is appears concentrated among higher-income taxpayers.

This approach, however, considers all of the flows, many of which are quite small. A substantial proportion of migration is accounted for by the top 30 flows and so it's informative to see what they are and how they may have changed over time. Past research reports that the concentration of elderly flows, measured as the percentage of all migration accounted for by the top 30 flows, has steadily declined from 35% in 1980 to 23% during 2006-10 (Conway and Rork, 2016, (Table 3)). Our data reveals a continuation of that decline, dropping from 27% in 2000 (matching Conway and Rork (2016) using decennial census data) to 22% in 2022.[11] While the migration flows of the young are less concentrated overall at 15% in 2000, their migration has steadily become slightly more concentrated, at 18% in 2022.

As expected, the largest flows are associated with the largest population states. We therefore reconstruct these rankings standardized by the average of the origin and destination state

[11] We find a substantially higher concentration rate of 26-28% during the 2006-10 period than Conway and Rork (2016) report using the ACS. This discrepancy suggests that the decline has continued in recent years and shows once again the shortcomings of using the ACS to measure elderly migration.

populations and the patterns in concentration are similar. Table 3.4 shows the top 30 migration flows, ranked by the share of migrants per 1,000 individuals, using the average of the population across the two states. For each flow, we also report the share of total inter-state moves captured by a given flow. This exercise once again reveals the stability of elderly migration. The top elderly migration flow is the same as it has been for 50+ years – New York to Florida (count or per capita). Florida is prominent – as both a destination and an origin (reflecting possible return migration) throughout the top 30 in every year. New York, New Jersey and the upper Midwest are prominent as well (more often as senders) and cross-border moves (e.g., NY-NJ) are common. The most obvious change is California's growing role as a top sender.

Taken together, these descriptive analyses suggest that elderly migration in the 2000s is fairly stable since 2000 and similar to previous decades. We do see more changes in the last few years of our data, 2020-22, which coincides with the Covid-19 pandemic and its many ramifications (i.e., remote work). Even there, however, the main geographic patterns persist.

3.3.2 State Income Tax Benefits for the Elderly

To study the degree of state tax benefits for the elderly, we construct a series of state individual income tax profiles. For each year, marital status, age (less than 62 or over 62) and AGI group, we pull a random sample of 1,000 tax units (after applying the restrictions regarding fraud, missing data, etc… applied earlier), without limiting the sample based on geography. We then pull from the SOI Databank and other databases information from the individuals tax return and information returns to construct the elements required by the NBER TAXSIM model (V35). To abstract from noise in the annual measures of income in each sample, we take the average income

Table 3.4: Top 30 per-capita elderly migration flows

Rank	2000		2010		2019		2022	
	Flow	Percent	Flow	Percent	Flow	Percent	Flow	Percent
1	NY to FL	2.38	NY to FL	2.53	NY to FL	2.13	NY to FL	2.14
2	FL to NY	1.73	NY to NJ	1.73	WA to AZ	0.49	NJ to FL	1.07
3	NJ to FL	1.31	MD to DC	1.31	CA to AZ	1.31	MA to NH	0.27
4	MI to FL	1.36	MO to KS	1.13	NJ to FL	1.01	WA to AZ	0.43
5	FL to MI	1.25	OR to WA	1.20	NC to SC	0.45	NC to SC	0.42
6	OH to FL	1.21	KS to MO	1.11	MI to FL	1.00	CA to AZ	1.09
7	MA to FL	0.93	VA to MD	0.43	MA to NH	0.26	CA to TX	1.41
8	FL to MA	0.88	NJ to NY	0.96	CO to AZ	0.35	NY to NJ	0.72
9	FL to OH	0.98	MD to VA	1.08	OR to WA	0.34	MA to FL	0.75
10	FL to NJ	0.85	WA to OR	0.23	CA to NV	0.93	IL to FL	0.85
11	WA to AZ	0.45	NC to SC	0.91	IL to FL	0.92	MD to DE	0.17
12	PA to FL	1.03	FL to GA	0.76	MA to FL	0.79	OR to WA	0.30
13	MA to NH	0.32	MA to NH	0.28	FL to NY	1.05	MI to FL	0.83
14	CT to FL	0.57	GA to FL	0.16	MD to DE	0.18	CA to NV	0.80
15	OR to WA	0.36	SC to NC	0.86	OH to FL	0.86	FL to GA	0.69
16	IN to FL	0.65	NH to MA	0.32	NY to NJ	0.67	FL to NY	0.92
17	DC to MD	0.21	NJ to PA	0.77	PA to FL	0.91	PA to FL	0.81
18	IL to FL	0.88	NH to ME	0.29	WA to OR	0.27	GA to FL	0.65
19	NY to NJ	0.82	KY to IN	0.26	FL to GA	0.66	OH to FL	0.73
20	CO to AZ	0.31	NY to FL	0.70	CA to TX	1.10	FL to NC	0.67
21	FL to IN	0.58	VA to NC	0.30	MN to AZ	0.27	CO to AZ	0.26
22	MN to AZ	0.34	VA to DC	0.66	GA to FL	0.64	WA to OR	0.25
23	CA to NV	1.01	DC to VA	0.25	OR to AZ	0.24	SC to NC	0.29
24	FL to CT	0.49	NC to VA	0.66	CT to FL	0.55	VA to NC	0.34
25	AZ to MN	0.32	PA to NJ	0.80	SC to NC	0.29	WA to ID	0.18
26	MI to IL	0.71	FL to NY	0.54	FL to MI	0.68	DC to MD	0.12
27	FL to PA	0.79	GA to AL	0.25	VA to NC	0.35	FL to MI	0.63
28	AZ to WA	0.33	GA to SC	0.60	WA to ID	0.18	OR to AZ	0.20
29	FL to IL	0.71	CT to MA	0.24	FL to NC	0.61	MD to VA	0.24
30	KS to MO	0.27	AL to GA	0.75	DC to MD	0.12	NC to FL	0.55
Top 30 share		24.04		22.07		19.60		18.77

Note: Data for this table are derived from the SOI Databank, from 1999-2022. The sample is limited to individuals who, for a give year t filed taxes in year t and $t-1$, for whom we observe their corrected adjusted gross income (AGI) and who report positive AGI in year t, who did not file a Married Filing Separately return, and who is not matched to an individual associated with a fraud claim. The flows are ranked by the number of movers per 1,000 residents, with residents defined as the average of the origin and destination state populations in the year prior to the move.

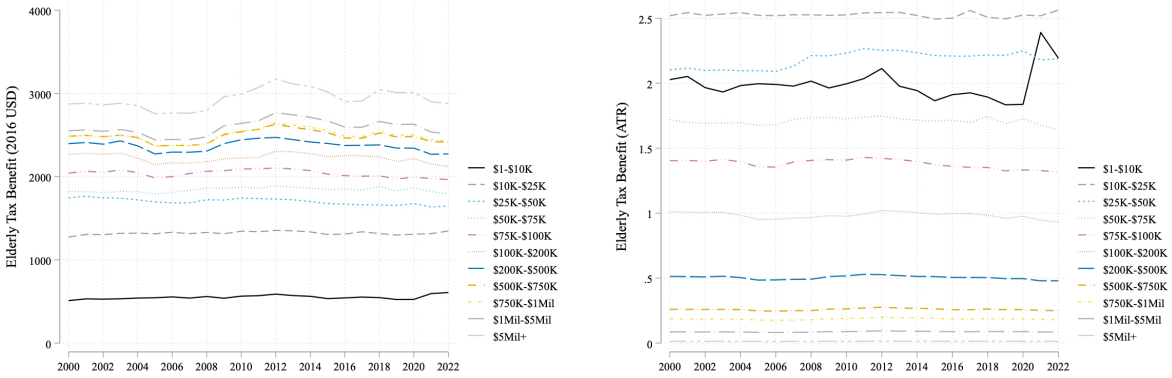
values (after adjusted for inflation) for each marital status, age, and AGI group across the entire period (2000-2022). This creates a representative taxpayer for each group. In order to provide a sense of the relative benefit the tax code of a given state might provide to the elderly, we create an alternative version of each of the elderly taxpayers that is non-elderly (de-aged), by assigning each elderly individual (and their spouse, in the case of the married filing jointly taxpayers) to be aged 45, and by reallocating any source of retirement income (pensions, social security, etc…) to be "other non-property income".[12] In essence, this turns off the set of elderly-specific tax preferences that are a function either of the age of a taxpayer (or their spouse) and/or the type of income being reported. After this step, we have, for each AGI and marital status group, three representative tax filers; under 62, over 62 and unchanged, and over 62 and de-aged. We then create an inflation adjusted version of these representative taxpayers for each tax year, and and run all of these through TAXSIM, once for each state. By comparing the tax liability of the elderly and de-aged representative tax filers, we can determine the scale of the elderly tax preference across each state over time, abstracting from changes in the composition and level of income.

Figure 3.2 reports the average state tax benefit to being elderly (ETB) for different AGI groups in terms of either the difference in overall tax liability (in 2019 USD) (reported in Figure 3.2a) and that difference divided by AGI, yielding the difference in the average tax rate (ATR) (reported in Figure 3.2b). As expected, the differential liability is larger the higher one's AGI. In real terms, it stays constant relatively constant across all groups, though there is an uptick for higher-income groups starting in 2008. The magnitudes are substantial: even for moderate

[12]There are a few isolated cases where NBER TAXSIM has coded a state-specific policy in some way that interacts with other non-property income in an unusual way. The two most common variants are states with individual income taxes that only apply to interest or dividends income (New Hampshire and Tennessee) and states with unique treatment of alimony (included in other non-property income). In these cases, we reassign retirement income to be interest income.

income households (50-75K) the average benefit is almost $2,000. Focusing on the differential ATRs, the pattern is reversed; the differential is inversely related to income, though with the same basic time-trend. These patterns fit with what we know about the type of elderly tax breaks that states offer: the dollar value of most tax breaks increase with income up to a point where they reach a maximum and, in some cases, begin to be phased out.

Figure 3.2: Average elderly tax benefit across AGI brackets for MFJ returns.



(a) Tax Liability (2019 USD)

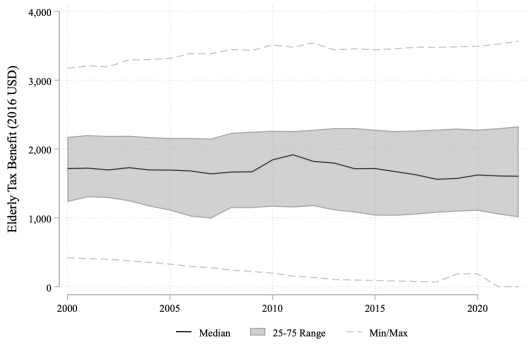
(b) Average Tax Rate

Note: Data for these figures are derived from the SOI Databank, from 2000-2022. Each line is the average across all states with individual income taxes, calculated using the income profiles described in Section 3.3.2. The elderly benefit is the difference between the taxes owned (in terms of total liability or average tax rates) between our elderly and de-aged elderly representative individuals.

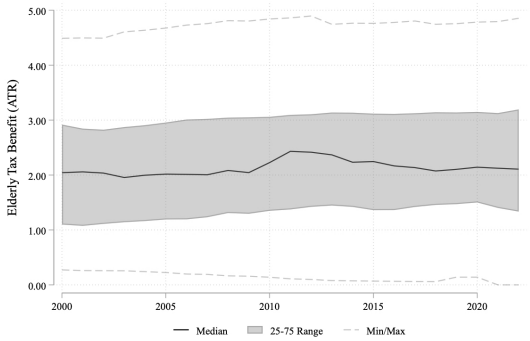
These average benefit measures obscure the wide variation across the states and how that changes during our sample. Figure 3.3 explores these aspects by reporting the median, the 25-75% range, the minimum and the maximum of these measures over time, each for selected AGI group. Figure 3.3e and Figure 3.3f, for very-high-income individuals reporting between $1 million and $5 million in AGI (real 2019 USD), confirm that the difference in rates is in fact tiny across all states, despite the sizable dollar amount existing in some states/years. Higher income households (those making between $100,000 and $200,000 in AGI) see only slightly lower dollar

amounts and therefore much bigger differences in their ATRs, with most states providing around 1% but some exceeding 4%, as shown in Figure 3.3c and Figure 3.3d. Moving further down the income distribution this pattern continues, with $25,000-$50,000 AGI households enjoying nearly a 3 percentage point differential in ATR (Figure 3.3a and Figure 3.3b). These figures also reveal that the variation in these policies differs by AGI. While the 25-75 percentile bands are fairly constant, the maximum value increases for the higher income groups while staying largely flat for the lower income group. These figures also mask that a few states decreased their tax breaks substantially. We investigate these cases further in Section 3.5. Further, for those at the bottom of the income distribution, there is both a shift in overall generosity, as shown by the median, in the middle of the sample, and a general widening of range of tax benefits offered, suggesting while large changes are rarer among this income group, states are adjusting their tax policy in smaller but important ways. This reflects the movement by a few states (like Georgia) to greatly expand their tax breaks to upper income retired households.

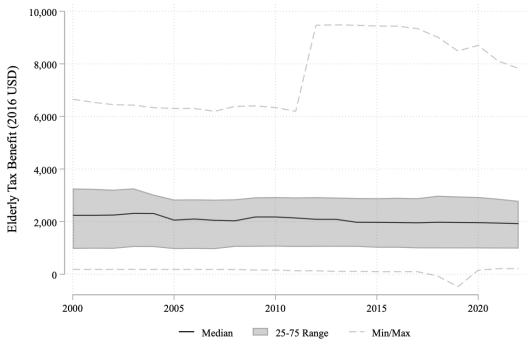
Figure 3.3: Distribution of elderly tax benefit across AGI brackets for MFJ returns.



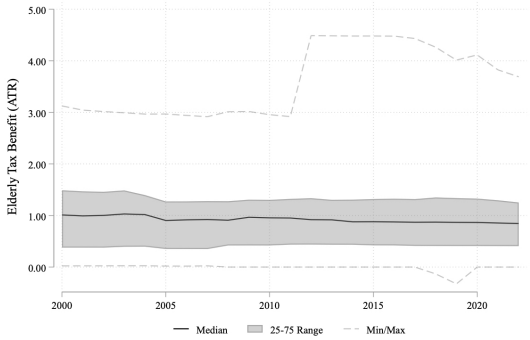
(a) Tax Liability (2019 USD), $25K-$50K in AGI



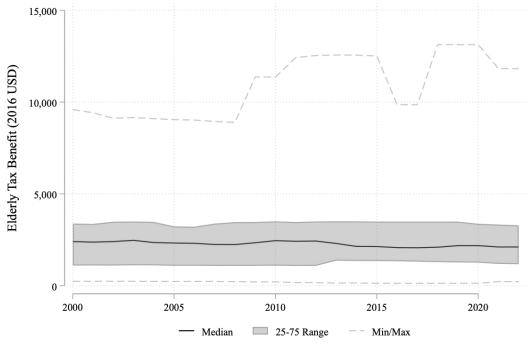
(b) Average Tax Rate, $25K-$50K in AGI



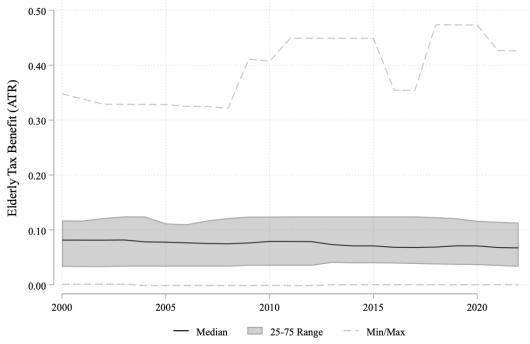
(c) Tax Liability (2019 USD), $100-$200K in AGI



(d) Average Tax Rate, $100-$200K in AGI



(e) Tax Liability (2019 USD), $1 Million-$5 Million in AGI



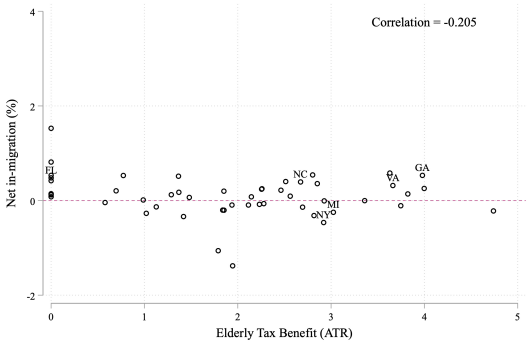
(f) Average Tax Rate, $1 Million-$5 Million in AGI

Note: Data for these figures are derived from the SOI Databank, from 2000-2022. Each figure reports the median, 25-75 percentile range, and maximum and minimum value for a given AGI bracket (in real terms) across states. The values are calculated using the income profiles described in Section 3.3.2. The elderly benefit is the difference between the taxes owned (in terms of total liability or average tax rates) between our simulated elderly and non-elderly individual.

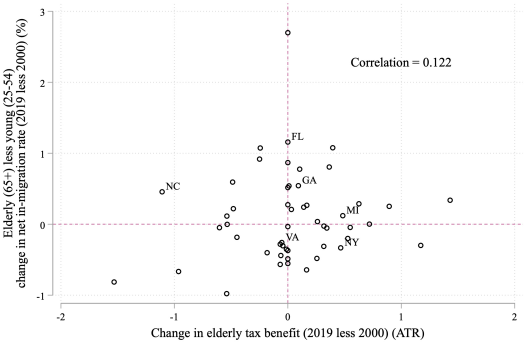
3.3.3 Are changes in Tax Policies and Migration Patterns Correlated?

Before turning to our main econometric approaches, we first examine whether there is an obvious pattern between changes in tax policy and elderly migration in the raw data. The above analyses demonstrate that while elderly migration patterns and state tax breaks overall are fairly stable, there remains nontrivial variation in both, and they tend to be present in the same groups, middle to higher income households. In Figure 3.4, we plot the net in-migration rate of each state against its ETB in terms of the ATR. In the first column (Figure 3.4a, Figure 3.4c, and Figure 3.4e) we plot the overall elderly benefit against net in-migration in 2000. Here, we can see there is no clear relationship between the generosity of elderly-specific tax programs and net in-migration in 2000. The second column (Figure 3.4b, Figure 3.4d, and Figure 3.4f) plot a similar figure that shows the correlation between the change in elderly benefits between 2019 and 2000 (on the y-axis) and the change in elderly in-migration over the same period (on the x-axis), after subtracting the same difference among the young. This is a rough difference-in-difference approach, which uses the change in net in-migration among the young to control for general trends in migration over the 2000s and 2010s. Again, we observe no clear relationship between changes elderly benefits and migration, and and all associated correlations are quite small and statistically insignificant.

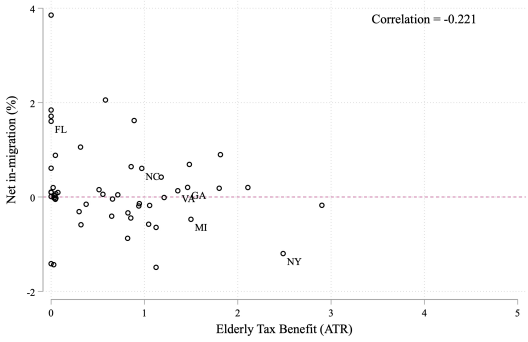
Figure 3.4: Relationship between net in-migration and elderly tax benefits



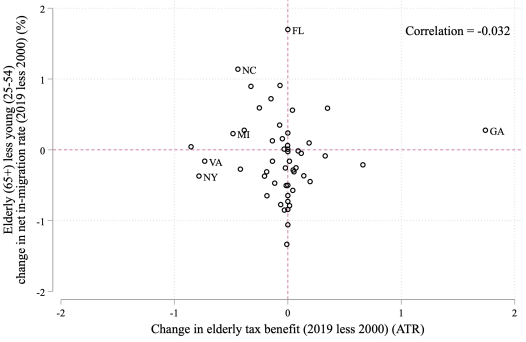
(a) TY 2000: $25,000-$50,000



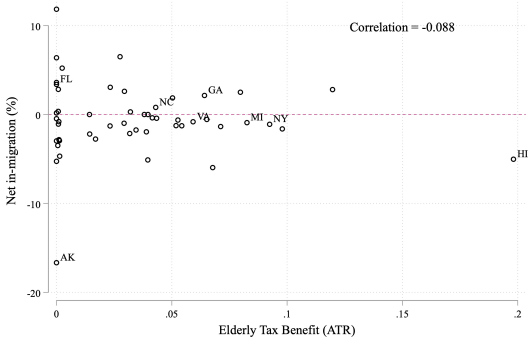
(b) TY 2019 less 2000: $25,000-$50,000



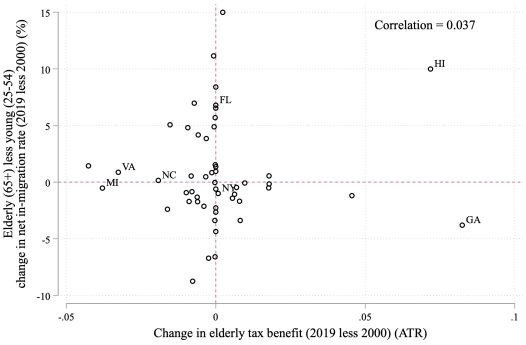
(c) TY 2000: $100,000-$200,000



(d) TY 2019 less 2000: $100,000-$200,000



(e) TY 2000: $1,000,000+



(f) TY 2019 less 2000: $1,000,000+

Note: Data for these figures are derived from the SOI Databank, from 2000-2019. Each plots net in-migration (y-axis) against the elderly state benefit in terms of average tax rates (x-axis). Figure 3.4a and Figure 3.4b are limited ot those taxpayers making between $25,000 and $50,000, with Figure 3.4a plotting the characteristics from 2000, while Figure 3.4b plots the change in the elderly benefit on the x-axis (2019 less 2000) and the difference in the change in the net in-migration rate (2019 less 2000) between the old verses the young) on the y-axis. The first is designed to show the relationship between elderly tax benefits and migration at the beginning of the sample, and the latter is design to observe the correlation between the change in taxes and the change in elderly in-migration (after controlling for changes in overall migration via subtracting the difference among the young. Figure 3.4c, Figure 3.4d, Figure 3.4e, and Figure 3.4f show the same statistics for two other AGI groups: $100,000 to $200,000 and $1,000,000+.

3.4 PPML Estimates

We estimate a comprehensive estimate of the effect of elderly tax benefits on elderly migration by applying a Poisson Pseudo-Maximum Likelihood (PPML) regression approach to our full sample containing the migration flows for all states (and the District of Columbia) from 2000-2019. This most closely tracks the estimation strategy employed by Moretti and Wilson (2017), who looked at the effect of state taxes on the migration behavior of scientists. Similar to that work, the main concern with identification in this structure is that there may be some factor that impacts elderly migration (say, economic conditions or other policy changes) that is correlated with state-level individual income tax changes. The specification below controls in several key ways for likely sources of ommitted variable bias.

3.4.1 Empirical Strategy

To account for the large number of migration flows that have a count of zero, we estimated the following model via Poisson Pseudo-Maximum Likelihood approach (Correia et al., 2020):

$$M_{odaft} = \exp\left(\beta \text{Tax}_{odaft} + \gamma X_{odt} + \delta_t + \phi_{odf} + \rho_{or,dr,t} + \epsilon_o daft\right) \tag{3.2}$$

Where M_{odaft} is the flow from origin state o to destination state d between period t and $t-1$ for AGI group a and filing status f. Tax_{odaft} captures the difference in tax policy between the destination and origin state, discussed in depth below. δ_t and ϕ_{odf} are time and flow-by-filing status fixed effects that are included in every model and set up the basic difference-in-difference structure: comparing the effect of changes in elderly benefits within a given flow and filing status

across time. X_{odt} is a set of origin-destination flow time-variant controls. These, expressed in terms of the difference between the lagged values between the origin and destination states, include controls for logged state population, per-capita new mortgages, the unemployment rate, per capita state and local revenue from sales, property, corporate, and estate taxes, overall and health and hospital per capita expenditures, and controls for elderly-specific property tax breaks derived from American Community Survey data (Ruggles et al., 2023). These controls account for the fact that state policymaking with regards to tax policy might be correlated with economic or policy factors that could also impact the migration decisions of elderly individuals. $\rho_{or,dr,t}$ are origin-region by destination-region by year fixed effects, included in some models. These are designed to account for time-varying regional factors (say, the shift in relative environmental conditions between the northeast and the pacific northwest) that might be correlated with both elderly migration and state policies.

We estimate this model separately for each AGI group, allowing the effect of tax rates (as well as any control variables) to vary with income. This is important both to determine if taxes influence migration differentially across the income distribution, but also because certain controls might have very different effects on lower or higher-income taxpayers. For example, Moretti and Wilson (2023) showed that for the very wealthy, estate taxes can influence migration decisions, but this is not likely to be the case for anyone outside of the top of the income distribution. We exclude the "stayers", i.e. the flows that represent the pool of individuals who did not move between states (e.g., New York to New York flow), as well as those with negative or $0 in AGI, who are a small share of total returns and for whom our measure of elderly preferences might not be appropriate. Standard errors are clustered at the origin-year, destination-year, and pair level, following Moretti and Wilson (2017), to account for the joint concerns about correlated errors

within year across origin or destination states and across time within a given flow.

We follow past work (e.g. Moretti and Wilson, 2017, Kalin et al., 2024) in expressing our tax measures in beneficial terms; e.g., we include the proportion of each dollar of income not taxed, or 1 – ATR. In our first specification, we estimate the effect of this tax rate on migration by defining Tax_{odaft} to be:

$$\text{Tax}_{odaft} = \overbrace{\left[(1-\tau^{e}_{daft-1})\right]}^{\text{Elderly after-tax rate in the destination state}} - \overbrace{\left[(1-\tau^{e}_{oaft-1})\right]}^{\text{Elderly after-tax rate in the origin state}} \tag{3.3}$$

Where the superscript e indicates the average tax rate calculated using our income profiles for our hypothetical elderly individual. The beta on this term will show the degree to which a one percentage point change in the difference between the after-tax rates increases or decreases migration. We calculate this value separately for elderly (65+) and non-elderly (25-54) individuals, using two separate sets of profiles. By comparing across age groups, we can compare the responsiveness of filers by age: the elderly (65+), the near-elderly (55-64), and the young (25-54).

We also want to isolate the effects of the differential treatment on the basis of age. Beyond allowing us to dis-aggregate the effect of eldelry-specific reforms, this approach also likely better abstracts from other features of the public sector. For example, an increase in the overall ATR is likely also associated with increases in spending or decreases in other taxes, owing to the government's budget constraint. In contrast, a decrease in the tax benefit of being elderly has a much more targeted effect on the decision-maker and a much smaller balanced-budget effect.

For this reason, our next specification includes two tax variables, ETB_{odaft} and NET_{odaft}. This approach separates the beneficial tax treatment of being elderly from the baseline tax burden.

Both are expressed as the difference between the destination and origin state in the year prior to the move. ETB_{odaft} captures the difference in the Elderly Tax Benefit (in terms of ATR – e.g., shown in Figure 3.2b to range from zero to 2.5), and NET_{odaft} captures the difference in overall taxes (expressed as 1-ATR) facing both the elderly and non-elderly, which we refer to as the non-elderly tax (NET). This specification allows us to see if the elderly respond differently to tax breaks targeting them rather than the overall tax structure.

Specifically, these two measures are calculated as:

$$\begin{aligned}\text{ETB}_{odaft} &= \overbrace{\left[(1-\tau^{e}_{daft-1})-(1-\tau^{ne}_{daft-1})\right]}^{\text{Elderly benefit in the destination state}} - \overbrace{\left[(1-\tau^{e}_{oaft-1})-(1-\tau^{ne}_{oaft-1})\right]}^{\text{Elderly benefit in the origin state}} \\ &= \left[\tau^{ne}_{daft-1}-\tau^{e}_{daft-1})\right] - \left[\tau^{ne}_{oaft-1}-\tau^{e}_{0aft-1})\right]\end{aligned} \tag{3.4}$$

Where the superscript ne indicates the average tax rate calculated using our income profiles for our hypothetical de-aged elderly individual. This variable captures the degree to which the destination state offers a targeted tax break (in terms of a lower ATR) to its elderly citizens as compared to the origin state. Again, these targeted tax breaks likely have far smaller balanced-budget effects and, like in the case of Georgia and Iowa, may be quite visible.

We define NET as the difference in the after-tax rate between destination and origin state for our de-aged elderly representative tax-filer. :

$$\text{NET}^{eb}_{odaft} = \left[(1-\tau^{ne}_{daft-1})-(1-\tau^{ne}_{oaft-1})\right] \tag{3.5}$$

Substituting these into equation (3.2) for TAX yields our estimating equation:

$$M_{odaft} = \exp\left(\beta_1 \text{ETB}_{odaft} + \beta_2 \text{NET}_{odaft} + \gamma X_{odt} + \delta_t + \phi_{odf} + \rho_{or,dr,t} + \epsilon_o daft\right) \tag{3.6}$$

Intuitively, this specification breaks the elderly taxpayer's burden into 2 pieces: the proportion of income that everyone gets to keep (NET_{odaft}) and the targeted, extra proportion they receive because they are elderly (ETB_{odaft}). ETB + NET equals the total 1-ATR for the elderly. We might expect the two betas to differ either because individuals have a different awareness of targeted policies or because the unobserved balanced budget effects differ. If neither complication exists, then we would expect them to be equal. The bottom line is that location choice theory suggests that both betas should be positive because an increase in either variable suggests the individual keeps a greater proportion of their income in the destination state relative to the origin state and thus is incentives to move.

When estimating Equation (3.6), we also present results for our three age groups, though the interpretation changes slightly. Here, we view the estimates for the young as a placebo check, and the estimates for the near-elderly capture the degree to which those approaching retirement proactively consider migration.

3.4.2 Results

Table 3.5 reports the PPML results from estimating Equation (3.2), with the tax variable defined as in Equation (3.3), for three different age groups (elderly, near elderly and the young) and eleven different AGI categories. This regression includes region-year fixed effects, as well as time-varying controls. For the elderly (Panel A), our estimates imply that a one-percentage-point

increase in the after-tax rate differential between destination and origin state increases migration by between 3 and 6 percent for those elderly filers who, in the prior-year, reported AGI between $200,000 and $1,000,000. To put these figures into context, the average elderly migration rate for this income group, as plotted in Figure 3.1b, is 1.4 percentage points, meaning that migration increased less than 0.09 percentage points as a result of these tax changes. For the near-elderly, we see similarly sized effects over a slightly wider income range, with the largest effects observed for the highest-income tax filers (those reporting over $5 million in AGI); a one percentage point after-tax differential increases migration by 6.4 percent. For the young, the response is more constant: an increase in migration of between 2 and 3 percentage points for tax filers making between $100,000 and $5,000,000 dollars in real AGI during the prior year. Overall, these results imply that higher-earning taxpayers do respond to changes in tax policy, though not by huge amounts, and that the elderly are marginally more responsive to overall changes in tax rates than their younger counterparts. The implied elasticities for the elderly range from 0.12 (for those with AGI between $200,000 to $500,000) to 0.23 (for those with AGI between $750,000 to $1,000,000), much lower than those estimated when looking at star scientists, athletes, those with very high incomes, and foreigners in other contexts (Moretti and Wilson, 2017, Kleven et al., 2013, 2014, Rauh and Shyu, 2024).[13]

The results are broadly consistent when excluding controls and region-year fixed effects, as shown in Table B.2, especially for the elderly and near-elderly. However, in this specification, the young appear less responsive than in the model with the full set of controls, which corroborates the notion that younger taxpayers are less mobile than the elderly. We also conduct a placebo

[13]To calculate this rough elasticity, we divide the percent estimate in Table 3.5 by $\frac{1}{\text{E}(1-\text{ATR}_{sat})}$, or the implied percent increase in the after-tax rate of a 1 percentage point increase in the after-tax rate in a given state for a given income group relative to the average after-tax rate in this sample (between 3.6 and 4.2 percent for these income groups.

test in Table B.3, checking to see whether or not tax filers respond to the 2-year lead of the tax variable. We find that, for the elderly, the very-wealthy seem to be marginally responsive to the 2-year lead, implying that the true effect as estimated in Table 3.5 may be biased upwards, but generally the results seem robust to this test.

Table 3.5: Effect of after-tax rates on inter-state migration by age, via Poisson Pseudo Maximum Likelihood

	(1) $1-$10K	(2) $10K-$25K	(3) $25K-$50K	(4) $50K-$75K	(5) $75K-$100K	(6) $100K-$200K	(7) $200K-$500K	(8) $500K-$750K	(9) $750K-$1M	(10) $1M-$5M	(11) $5M+
Panel A: Elderly (65+)											
Elderly 1-ATR Differential	0.0160	0.0145	0.0255	0.0128	0.0083	0.0133	0.0338*	0.0487***	0.0533***	0.0153	-0.0087
	(0.0121)	(0.0226)	(0.0227)	(0.0213)	(0.0160)	(0.0147)	(0.0139)	(0.0143)	(0.0153)	(0.0109)	(0.0229)
Observations	96311	96805	96729	96330	95380	95076	90003	61211	45203	56474	23462
Population in Sample	1494962	2528798	2268435	1516117	964631	1284561	442304	62306	26920	53375	8812
Percent effect	1.62	1.46	2.58	1.29	.836	1.34	3.43	4.99	5.48	1.54	-.863
Panel B: Near-Elderly (55-64)											
Non-Elderly 1-ATR Differential	0.0216	0.0341	0.0338	0.0187	0.0330	0.0363**	0.0323**	0.0195	0.0352**	0.0293**	0.0622**
	(0.0135)	(0.0227)	(0.0193)	(0.0219)	(0.0177)	(0.0129)	(0.0102)	(0.0100)	(0.0119)	(0.0107)	(0.0201)
Observations	96425	96786	96824	96767	96140	96121	91371	69505	54673	63909	28391
Population in Sample	1287717	2454713	3024718	2048480	1306447	1888677	684425	93635	39966	78095	11780
Percent effect	2.18	3.47	3.43	1.88	3.35	3.69	3.28	1.97	3.58	2.97	6.42
Panel C: Young (25-54)											
Non-Elderly 1-ATR Differential	0.0128	0.0259	0.0226	0.0162	0.0237	0.0298**	0.0205*	0.0200*	0.0226**	0.0222**	0.0241
	(0.0075)	(0.0137)	(0.0131)	(0.0180)	(0.0134)	(0.0107)	(0.0080)	(0.0082)	(0.0080)	(0.0078)	(0.0139)
Observations	96900	96900	96900	96900	96881	96824	95133	78698	65578	72271	39394
Population in Sample	1.06e+07	2.10e+07	2.25e+07	1.16e+07	6225112	7306045	2075389	229200	88296	149150	19885
Percent effect	1.28	2.63	2.29	1.63	2.39	3.02	2.07	2.02	2.29	2.24	2.44

Note: This table reports the results of the Poisson Pseudo Maximum Likelihood estimation of the effect of changes in after-tax rates on inter-state migration, estimated via Equation (3.2) using the tax measure defined in Equation (3.3). Data for this table are derived from the SOI Databank, from 2000 through 2019. We start with a 100 percent sample of individuals 25 years or older who, for a give year t, filed taxes in year t and $t-1$, for whom we observe their corrected adjusted gross income (AGI) and who report positive AGI in year t, who did not file a married filing separately return, and who is not matched to an individual associated with a fraud claim. These data are collapsed to create a count of individuals in a given flow, defined by year, origin and destination state, AGI bracket, age group, and filing status (married or single, with other statuses included in single). The outcome variable is the count of individuals in a given if a given individual flow. Each panel presents the results for a different age group (65+, 55-64, and 25-54), and each row for different AGI groups (assigned via year $t-1$ AGI, with the AGI brackets in 2019 real USD). The conversion of the coefficients into percent terms ($\exp(\beta)-1$) are reported at the bottom of each panel. All models include flow by filing status fixed1 effects, origin-region by destination-regions by year fixed effects, year fixed effects, and controls for state-level population, policy, and economic characteristics, expressed as the difference between the destination and origin states. Standard errors are clustered at the origin-year, destination-year, and flow level * p<0.05 ** p<0.01 *** p<0.001.

Next, we report the results of estimating Equation (3.6) in Table 3.6, breaking the effect of the change in after-tax rates into the two components: ETB and NET. When looking at the ETB, we find no effect of of the after-tax rate differential on the migration response of the elderly, implying that elderly-specific tax breaks are not driving migration decisions. Rather, we see that the positive effect seen in Table 3.5 load on the NET variable. The coefficients for the ETB variable statistically insignificant across all income groups, while the sign and magnitude of the NET variable lines up with the results from Table 3.5. We see similar results for the near-elderly, except for the very-wealthy, were we see a substantial decline in elderly migration corresponding to elderly-specific tax breaks. Recall, however, that for this income group, these tax breaks are incredibly small, so this result is likely spurious. The young are similarly not responsive to the ETB. The ETB estimates for the young is akin to a placebo test and can be subtracted from the elderly parameter estimates to obtain a rough DDD estimate. It is therefore reassuring to see that the ETB estimates are all close to zero for the young. These results strongly suggest that targeted tax breaks – in isolation – have no measurable impact on location decisions. This result is robust to excluding region-year fixed effects and our controls (see Table B.4) and we similarly see no effect of the 2-year lead of the ETB on the elderly population (see Table B.5).

The main take-away from this analysis is that while the high income elderly are responsive to income tax burdens, perhaps even more so than their younger counterparts, the predicted effects on migration are modest. Moreover, tax cuts targeting the elderly, as measured by ETB, have no discernible effect. We conduct a range of additional robustness checks on these results, including limiting the sample of states to exclude those without individual income taxes (Table B.7) and using F1040 addresses rather than our measure of migration (Table B.8: these results do not change.

Table 3.6: Effect of elderly-specific after-tax rates on inter-state migration by age, via Poisson Pseudo Maximum Likelihood

	(1) $1-$10K	(2) $10K-$25K	(3) $25K-$50K	(4) $50K-$75K	(5) $75K-$100K	(6) $100K-$200K	(7) $200K-$500K	(8) $500K-$750K	(9) $750K-$1M	(10) $1M-$5M	(11) $5M+
Panel A: Elderly (65+)											
ETB	0.0077	0.0014	0.0136	-0.0014	-0.0192	-0.0264	-0.0349	-0.0477	0.0738	-0.0300	0.5742
	(0.0129)	(0.0263)	(0.0280)	(0.0287)	(0.0213)	(0.0215)	(0.0351)	(0.0635)	(0.1113)	(0.0210)	(0.3971)
NET	0.0488	0.0310	0.0394	0.0269	0.0304	0.0338*	0.0367**	0.0455**	0.0539***	0.0222	-0.0087
	(0.0305)	(0.0231)	(0.0214)	(0.0199)	(0.0170)	(0.0154)	(0.0137)	(0.0141)	(0.0152)	(0.0119)	(0.0231)
Observations	96311	96805	96729	96330	95380	95076	90003	61211	45203	56474	23462
Population in Sample	1494962	2528798	2268435	1516117	964631	1284561	442304	62306	26920	53375	8812
Percent effect - ETB	.77	.138	1.37	-.141	-1.91	-2.6	-3.43	-4.66	7.66	-2.95	77.6
Percent effect - NET	5.01	3.15	4.02	2.72	3.08	3.43	3.74	4.65	5.54	2.25	-.87
Panel B: Near-Elderly (55-64)											
ETB	0.0146	0.0055	0.0166	0.0020	-0.0077	-0.0135	-0.0260	-0.0018	-0.0802	0.0328	-0.5038*
	(0.0138)	(0.0296)	(0.0326)	(0.0270)	(0.0219)	(0.0129)	(0.0200)	(0.0458)	(0.0810)	(0.0185)	(0.2441)
NET	0.0397	0.0353	0.0411	0.0263	0.0274	0.0258*	0.0277**	0.0192	0.0340**	0.0305**	0.0615**
	(0.0223)	(0.0246)	(0.0248)	(0.0209)	(0.0182)	(0.0132)	(0.0103)	(0.0103)	(0.0127)	(0.0108)	(0.0205)
Observations	96425	96786	96824	96767	96140	96121	91371	69505	54673	63909	28391
Population in Sample	1287717	2454713	3024718	2048480	1306447	1888677	684425	93635	39966	78095	11780
Percent effect - ETB	1.47	.556	1.68	.203	-.769	-1.34	-2.56	-.176	-7.71	3.33	-39.6
Percent effect - NET	4.05	3.59	4.2	2.67	2.77	2.62	2.81	1.94	3.46	3.09	6.34
Panel C: Young (25-54)											
ETB	0.0008	0.0019	0.0029	0.0019	-0.0067	-0.0005	0.0075	0.0203	0.0889	0.0086	-0.2830
	(0.0094)	(0.0221)	(0.0221)	(0.0217)	(0.0146)	(0.0153)	(0.0195)	(0.0321)	(0.0598)	(0.0125)	(0.3852)
NET	0.0187	0.0265	0.0229	0.0179	0.0176	0.0251*	0.0219*	0.0216*	0.0270**	0.0233**	0.0260
	(0.0124)	(0.0164)	(0.0155)	(0.0151)	(0.0137)	(0.0112)	(0.0086)	(0.0088)	(0.0089)	(0.0080)	(0.0143)
Observations	96900	96900	96900	96900	96881	96824	95133	78698	65578	72271	39394
Population in Sample	1.06e+07	2.10e+07	2.25e+07	1.16e+07	6225112	7306045	2075389	229200	88296	149150	19885
Percent effect - ETB	.079	.186	.289	.19	-.664	-.0497	.755	2.05	9.3	.866	-24.7
Percent effect - NET	1.89	2.68	2.32	1.81	1.77	2.54	2.21	2.19	2.74	2.35	2.64

Note: This table reports the results of the Poisson Pseudo Maximum Likelihood estimation of the effect of changes in after-tax rates on inter-state migration, estimated via Equation (3.6) using the tax measure defined in Equations (3.4) and (3.5). Data for this table are derived from the SOI Databank, from 2000 through 2019. We start with a 100 percent sample of individuals 25 years or older who, for a give year t, filed taxes in year t and $t-1$, for whom we observe their corrected adjusted gross income (AGI) and who report positive AGI in year t, who did not file a married filing separately return, and who is not matched to an individual associated with a fraud claim. These data are collapsed to create a count of individuals in a given flow, defined by year, origin and destination state, AGI bracket, age group, and filing status (married or single, with other statuses included in single). The outcome variable is the count of individuals in a given if a given individual flow. Each panel presents the results for a different age group (65+, 55-64, and 25-54), and each row for different AGI groups (assigned via year $t-1$ AGI, with the AGI brackets in 2019 real USD). The conversion of the coefficients into percent terms ($\exp(\beta)-1$) are reported at the bottom of each panel. All models include flow by filing status fixed effects, origin-region by destination-regions by year fixed effects, year fixed effects, and controls for state-level population, policy, and economic characteristics, expressed as the difference between the destination and origin states. Standard errors are clustered at the origin-year, destination-year, and flow level * p<0.05 ** p<0.01 *** p<0.001.

3.4.3 Heterogeneity By Age

One question is whether or not the migration response of elderly individuals is heterogeneous: there are multiple stages of life covered by the elderly population, and difference forces drive migration during these stages. For example, those just turning 65 and retiring might move to a new home once they disconnect from the labor force, while older Americans might return to live closer to family (the pattern of return migration discussed above). To explore this, we break the elderly population into those aged 65-74 and those older than 75, and re-estimate the effect of elderly-specific tax breaks. Importantly, these policies do not distinguish on the basis of age within the elderly population, so they face the same after-tax rates, conditional on similar income mixes. We report the results of these models in Table 3.7. The results for those aged 65-74 are broadly consistent with the results in Table 3.6: no statistically significant differential effect of elderly-specific tax policies on migration. The results for those aged 75 or older are generally thee same, with the exception of those reporting AGI between $750,000 and $1,00,000. Here, an increase in the elderly-specific after-tax rate differential of one percentage point increases migration by 60 percent, corresponding to an elasticity of 2.5. This is large, through not unreasonably so relative to some findings (See the elasticity estimate of 3.0 found by Rauh and Shyu (2024) in the context of California). However, the fact that no other similar income groups saw such a large change, and given that the same magnitude of a response is found when using a 2-year lead of the tax variable in Table B.6, this result is likely spurious.

Table 3.7: Effect of elderly-specific after-tax rates on inter-state migration across the elderly, via Poisson Pseudo Maximum Likelihood

	(1) $1-$10K	(2) $10K-$25K	(3) $25K-$50K	(4) $50K-$75K	(5) $75K-$100K	(6) $100K-$200K	(7) $200K-$500K	(8) $500K-$750K	(9) $750K-$1M	(10) $1M-$5M	(11) $5M+
Panel A: Elderly, 65-74											
ETB	0.0137	0.0032	0.0172	0.0026	-0.0218	-0.0284	-0.0447	-0.0820	-0.0581	-0.0149	0.6763
	(0.0152)	(0.0301)	(0.0327)	(0.0304)	(0.0223)	(0.0196)	(0.0339)	(0.0774)	(0.1286)	(0.0239)	(0.5964)
NET	0.0605	0.0385	0.0402	0.0280	0.0325	0.0337*	0.0456**	0.0461**	0.0526**	0.0252	-0.0072
	(0.0328)	(0.0265)	(0.0252)	(0.0215)	(0.0180)	(0.0158)	(0.0144)	(0.0162)	(0.0163)	(0.0131)	(0.0254)
Observations	95494	96387	96368	95646	93670	93214	85438	54036	37940	48326	18866
Population in Sample	846956	1428287	1370294	972557	652671	902489	307440	42523	18169	36030	6094
Percent effect - ETB	1.37	.319	1.73	.264	-2.15	-2.8	-4.37	-7.88	-5.65	-1.47	96.7
Percent effect - NET	6.23	3.93	4.1	2.84	3.3	3.43	4.66	4.71	5.4	2.55	-.718
Panel B: Elderly, 75+											
ETB	0.0004	-0.0009	0.0090	-0.0089	-0.0148	-0.0147	-0.0005	0.0361	0.4663*	-0.0565	0.4799
	(0.0118)	(0.0234)	(0.0226)	(0.0273)	(0.0227)	(0.0270)	(0.0388)	(0.0735)	(0.2135)	(0.0330)	(0.7055)
NET	0.0330	0.0203	0.0374*	0.0235	0.0247	0.0333*	0.0186	0.0414*	0.0606*	0.0199	-0.0185
	(0.0285)	(0.0212)	(0.0180)	(0.0187)	(0.0175)	(0.0159)	(0.0150)	(0.0181)	(0.0246)	(0.0186)	(0.0349)
Observations	94145	96083	95589	94164	90364	90478	77787	40359	25690	36255	9672
Population in Sample	648006	1100511	898141	543560	311960	382072	134864	19783	8750	17345	2708
Percent effect - ETB	.0422	-.0851	.9	-.882	-1.46	-1.46	-.047	3.68	59.4	-5.49	61.6
Percent effect - NET	3.35	2.05	3.81	2.38	2.5	3.39	1.87	4.23	6.25	2.01	-1.83

Note: This table reports the results of the Poisson Pseudo Maximum Likelihood estimation of the effect of changes in after-tax rates on inter-state migration, estimated via Equation (3.6) using the tax measure defined in Equations (3.4) and (3.5). Data for this table are derived from the SOI Databank, from 2000 through 2019. We start with a 100 percent sample of individuals 25 years or older who, for a give year t, filed taxes in year t and $t-1$, for whom we observe their corrected adjusted gross income (AGI) and who report positive AGI in year t, who did not file a married filing separately return, and who is not matched to an individual associated with a fraud claim. These data are collapsed to create a count of individuals in a given flow, defined by year, origin and destination state, AGI bracket, age group, and filing status (married or single, with other statuses included in single). The outcome variable is the count of individuals in a given if a given individual flow. Each panel presents the results for a different age group (65-74, and 75+), and each row for different AGI groups (assigned via year $t-1$ AGI, with the AGI brackets in 2019 real USD). The conversion of the coefficients into percent terms ($\exp(\beta)-1$) are reported at the bottom of each panel. All models include flow by filing status fixed effects, origin-region by destination-regions by year fixed effects, year fixed effects, and controls for state-level population, policy, and economic characteristics, expressed as the difference between the destination and origin states. Standard errors are clustered at the origin-year, destination-year, and flow level * p<0.05 ** p<0.01 *** p<0.001.

3.5 Case Studies of Isolated Policy Changes

Our tax measures capture the many, often subtle ways that the tax benefits of being elderly change over time, which lead to policy variables that are almost constantly changing in the states with personal income taxes. However, it seems likely that taxpayers may be less aware of these smaller, indirect changes and instead are primarily impacted by large, highly visible ones. We therefore intend to investigate the effects of a few highly visible changes to policy, before turning to our long panel, migration flow analysis that includes all of the states and years of our data (2000-2019) in Section 3.4.

This approach has the advantage that we do not have to explicitly model what policy measure (e.g., the difference in the ATR versus the difference in tax liability) individuals are responding to. It also allows us to conduct the event history analyses common to difference-in-difference analyses. The relative overall stability of the tax measures, as shown in Figure 3.2 and Figure 3.3, lends support to this approach as well. Nonetheless, as described below, to abstract from policy changes in other states, we limit our analyses to individuals living in states without an income tax to study in-migration patterns and to individuals currently living in the affected state when studying out-migration. We can therefore study asymmetries between moving in or out of state and, because we use individual-level data, can control for characteristics such as age.

We intend to choose policy changes that 1) reflect both increases and decreases in elderly tax preferences, 2) affect different parts of the income distribution, 3) operate through different policies, and 4) occur in different time periods while also having sufficient data on years before and after the policy change to identify an effect. Below, we describe two examples: a dramatic increase in elderly tax benefits in Georgia (in 2006 and again in 2012), and the phasing out of the

taxation of social security benefits in Iowa starting in 2007.

3.5.1 Empirical Strategy

For each reform, we estimate the effect of the tax change on the in- and out-migration of the elderly (or near-elderly). We employ a DID approach, comparing the migration patterns of the impacted age group to that of the young, before and after the reform. This approach identifies the causal effect of each elderly-specific tax reform under the assumption that, but for the given tax change, the trend in the migration of the young would be an appropriate counter-factual (up to a level shift) for the elderly or near-elderly. Using the individual-level dataset, we estimate the following difference-in-difference specification:

$$y_{it} = \alpha + \beta \text{Treated}_{it} \times \text{Post}_t + X_{it} + \delta_t + \zeta_a + \gamma_s + \epsilon_{it} \tag{3.7}$$

Where y_{it} is a binary variable equal to one if individual entered or exited the given state. To avoid reporting very-small estimated coefficients, we scale y_{it} by 100 such that the estimated effect are *percentage* changes rather than proportionate ones. Treated_{it} is an indicator equal to one if the individual was over the age of 65 (or between 55 64 in the near-elderly sample), and Post_t is an indicator equal to one if the observation is in the post-treatment period. δ_t and ζ_a are fixed effects for year and age groups, respectively. X_{it} is a vector of controls, including measures of marital status, age (and age squared), count of dependents, the individual's AGI group based on their prior-year income ($t-1$), and (for in-migration) controls for origin-state elderly-specific property taxes, and γ_s is an origin-state FE, only included when estimating the effect of a reform on in-migration. Standard errors are clustered at the origin-state and individual level for in-

migration, and at the individual-level for out-migration. Importantly, because the comparison is now between elderly and young taxpayers, we need to adjust our definition of AGI, which currently excludes certain types of elderly-specific income (social security, pensions, etc...). To make sure the comparison is between individuals with similar levels of resources, we add back into AGI these sources of income, creating our definition of modified AGI.[14]

For each model we estimate both a traditional DiD, as well as an event-study model, which allows for a consideration of pre-trends and any potential dynamic effects. We estimated the event-study specification using equation (3.8):

$$y_{it} = \alpha + \sum_{j=T-5}^{T+4} \text{Treated}_{it} \times \mathbb{1}(t = j) + X_{it} + \delta_t + \zeta_a + \gamma_s + \epsilon_{it} \tag{3.8}$$

Where T is the period when the policy changed occurred. We limit the sample to five years prior to and following a given reform, though in certain cases we extend the sample to the entire 2001 through 2019 period.

To study the decision to move into the affected state we limit the sample to individuals who, in period $t-1$, were living (according to their information return zip codes) in a state without an individual income tax (Florida, Nevada, New Hampshire, South Dakota, Tennessee, Texas, Washington, and Wyoming)[15] When estimating the effect of a given reform on out-migration, we instead restrict the sample to individuals who were living in the state in question in year $t-1$.

[14] Specifically, $\text{AGI}^M_{it} = \text{AGI}_{it} + \text{Soc. Sec}^{\text{non-taxed}}_{it} + \text{IRA Distr.}^{\text{non-taxed}}_{it} + \text{Pensions}^{\text{non-taxed}}_{it}$

[15] We exclude one no-income tax state, Alaska, given its small population and unusual geography and policy landscape.

3.5.2 Georgia

Our first analysis will look at the reforms to the taxation of individual income in Georgia that took place in 2006 and 2012. Georgia has always had an exemption for retirement income for anyone over the age of 62, which includes $4,000 of earned income in addition to income from pensions, dividends, capital gains, interest, royalties, rent, partnerships and S-corporations. In 2006 Georgia increased the exemption for a married couple from $30,000 to $50,000, an amount that continued to grow to $70,000 in 2008. In 2012, the exclusion increased by another $30,000 per person people aged 65 and over (instead of 62). Figure 3.5a and Figure 3.5b show clearly the effects of these policies that, as expected, have the strongest effects on higher income households. In terms of ATR, the effects are strongest for those in the $75,000 to $500,000 range, which also makes sense as the proportionate effect of the tax break diminishes once income gets very large. This again raises the question of whether individuals respond to ATR or liability differences. For tax units with lower AGI, there is either no effect or, in the case of the very bottom of the AGI distribution, a reduction in the elderly tax benefit.[16] For tax units with much higher earnings, the effect is still present, but is smaller as a share of their AGI. Figure 3.5c and Figure 3.5d show the net-in migration trends for Georgia for the elderly and young by AGI groups. Overall, the trends are consistently close to 0, with the notable exception of the spike in 2008 for the lowest AGI bin. This spike is present for both the elderly and the young, and is occurring among the group without any exposure to the elderly tax reforms. The aggregate data therefore shows no obvious evidence of policy-induced migration.

Because Georgia had two-rounds of elderly tax reforms, we estimate both our difference-in-

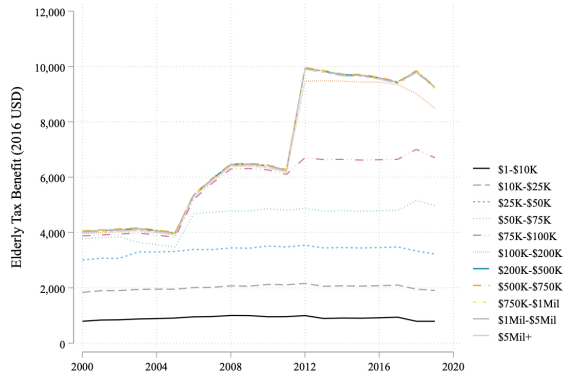
[16]This resulted from a separate change in 2010, when Georgia's small low-income tax credit was made non-refundable. That credit provided an additional $26 dollars per elderly resident for the very-lowest-income tax filers.

difference and event-study specifications over two time periods: 2001-2010 and 2001-2019. The former captures only the effect of the first round of reforms in 2006, while the second captures both. We present results on in- and-out migration, for both elderly and near-elderly, for the entire sample and separately by four AGI groups ($1-$50,000, $50,000-$100,000, $100,000-$500,000, and $500,000+). We would expect the largest effects to occur in the middle-two AGI groups, where the effect of the reform on the ATR is largest. Importantly, there is some concern in the case of Georgia with a confounding policy shock: the implementation of a 100 percent exemption for state property taxes for elderly residents; however, state property taxes made up only a small fraction of the property tax burden in Georgia, with local taxes representing the larger share.

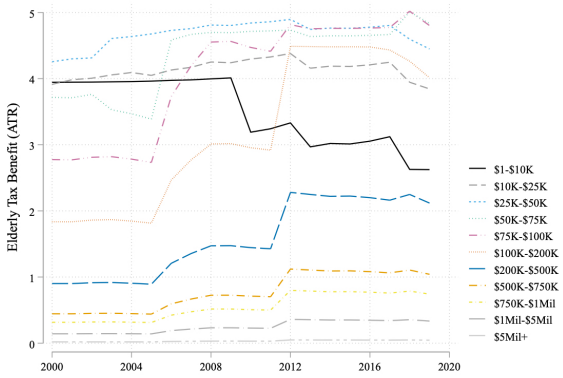
3.5.3 Iowa

In 2007, Iowa enacted a series of reforms to begin phasing out the taxation of social security benefits, designed to result in the elimination on the taxation of these benefits by 2014. A large portion (approximately one-third) of the elimination occurred immediately, resulting in a sharp increase in the elderly-specific benefit, as seen in Figure 3.6a and Figure 3.6b Like in the prior section, these figures report the impact on elderly tax benefits and Figure 3.6c and Figure 3.6d report the net-migration of the elderly and, for comparison, non-elderly (25-54) during the time period. The policy resulting in an increase in the elderly benefit of almost a $1,000 among the higher-income AGI groups that grew slightly with AGI, though the benefits as a percent of AGI are concentrated among tax units reporting between $25,000 and $200,000 in AGI: Lower than the impacted AGI groups in Georgia. Among these groups, there is no evidence of an adjustment of migration (in or out) after the implementation of the policy in 2007.

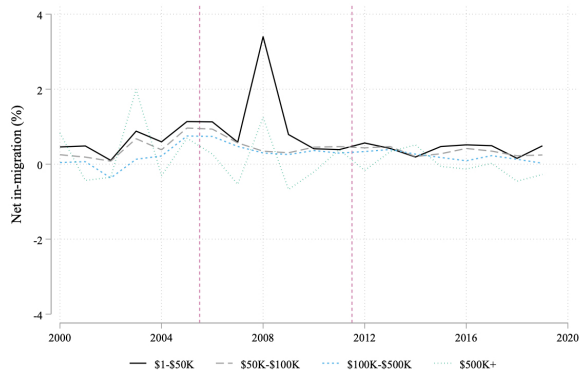
Figure 3.5: Georgia elderly tax benefits and net in-migration



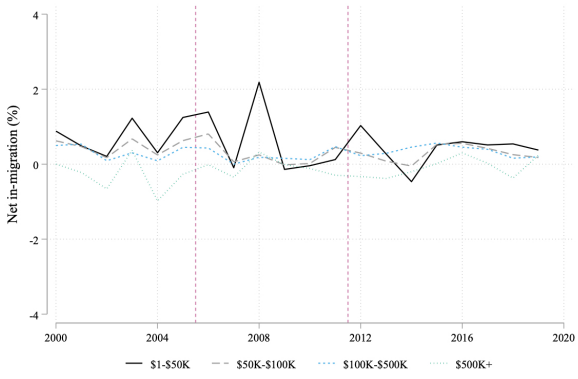
(a) Elderly Tax Benefit, Tax Liability (2019 USD)



(b) Elderly Tax Benefit, Average Tax Rate



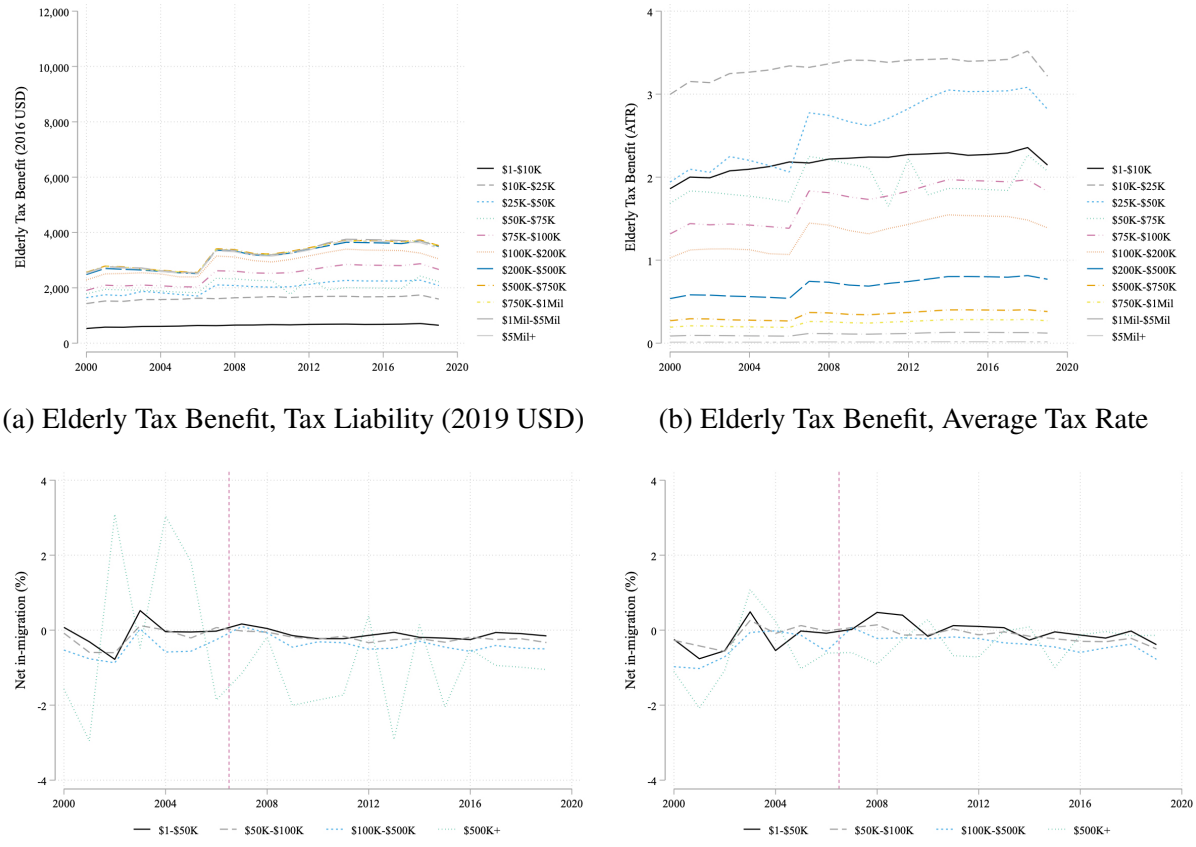
(c) Net in-migration, Elderly (65+)



(d) Net in-migration, Young (25-54)

Note: Data for these figures are derived from the SOI Databank, from 2000-2019. Figure 3.5a and Figure 3.5b plot the elderly tax benefit (in 2016 dollars or average tax rates), as calculated via our income profiles described in Section 3.3.2, for different AGI groups in Georgia, defined as the difference between the taxes owned (in terms of total liability or average tax rates) between our simulated elderly and young individuals. Figure 3.5c and Figure 3.5d plot the net-in migration among the elderly and near-elderly for larger AGI groups.

Figure 3.6: Iowa elderly tax benefits and net in-migration



(a) Elderly Tax Benefit, Tax Liability (2019 USD)

(b) Elderly Tax Benefit, Average Tax Rate

(c) Net in-migration, Elderly (65+)

(d) Net in-migration, Young (25-54)

Note: Data for these figures are derived from the SOI Databank, from 2000-2019. Figure 3.6a and Figure 3.6b plot the elderly tax benefit (in 2016 dollars or average tax rates), as calculated via our income profiles described in Section 3.3.2, for different AGI groups in Michigan, defined as the difference between the taxes owned (in terms of total liability or average tax rates) between our simulated elderly and young individuals. Figure 3.6c and Figure 3.6d plot the net-in migration among the elderly and near-elderly for larger AGI groups.

3.6 Cost of Elderly-Specific Policies

So far, our results imply that elderly (or near-elderly) individuals do not alter their migration decisions in response to individual income tax policies that target tax units based either on their age or source of income. This implies that these policies represent a transfer from the state (i.e. all taxpayers in the state, elderly or not) to a specific group of elderly residents whose income is taxed a lower rate due to these policies. In this section, we attempt to estimate the cost of these policies. Importantly, we capture the cost of all such programs that preference the the elderly based on either their age or the type of income, not just those enacted between 2000 and 2019 (the source of variation in the estimates above).

The full details of the costing analysis are described in Appendix B.2 and Appendix B.1, but in essence, we construct a version of the individual-level stratified random sample that is not limited to individuals observed in both year t and $t-1$ for 2019, the last year of our sample. After constructing all the required variables to run the tax units these individuals belong to through NBER's TAXSIM, we estimate the state tax liability for each individual twice: once with and once without the elderly preferences. Importantly, this approach assumes no behavioral change (migration, income shifting, etc...).

Our results, shown in Table 3.8, show that on-average, elderly benefits reduce individual income tax revenue by 5 percent, or around $20 billion dollars, in 2019. In two states, New Hampshire and Tennessee, the impact in percent terms is far larger (between 63 and 66 percent), due to the unique structure of individual income taxes in both states. New Hampshire only taxes interest income, and Tennessee only interest and dividend income, so our application here (treating elderly income as interest income for these two states) dramatically expands the tax base.

However, these two states are small enough in terms of IIT so as not to change the overall results significantly. South Carolina and Georgia have the largest per-tax-unit cost: around $775 USD for both states, and both are in the set of 7 states that see their individual income tax revenue drop by over 10 percent (New Hampshire and Tennessee, South Carolina and Georgia, plus Delaware, Michigan, and Mississippi).

3.7 Conclusion

This research provides a crucial update on the trends, patterns and policy determinants of elderly interstate migration in the United States. The elimination of the Census long form after 2000 greatly hampered the ability of researchers to study the migration of certain age and income groups with publicly available data. The elderly are especially challenging to study because their interstate migration rates have historically been substantially lower than those of other age groups, requiring a larger sample to accurately gauge meaningful changes over time. Since 2000, however, both the size of the elderly population and prevalence of targeted income tax breaks for them have increased, suggesting that the possible consequences of these policy actions have grown as well. Are the aging boomers different from earlier generations in their location decisions? Does cutting taxes for them and the older elderly recruit and retain residents? Or does it simply result in an ever-growing reduction in income tax revenues?

We seek to answer these questions using restricted, administrative data on the universe of tax filers since 1999 and a multi-pronged empirical approach that combines that data with state income tax policies. The answer to the first question is that elderly migration patterns are strikingly similar to that suggested by older, pre-2000 research. The administrative tax data we

Table 3.8: Elderly Benefit Cost Estimate for 2019

State	State IIT Revenue (Millions 2019 USD)	Percent Effect of Elderly Tax Policies	Implied Foregone Revenue (Millions 2019 USD)	Implied Average Benefit (2019 USD)
Alaska	-	0.0%	-	-
Alabama	4,195	8.5%	327	465
Arkansas	3,012	8.2%	229	541
Arizona	5,357	6.9%	345	313
California	100,080	2.9%	2,800	416
Colorado	8,172	4.8%	371	375
Connecticut	8,446	3.1%	257	401
Delaware	1,741	11.1%	174	1,100
District of Columbia	2,299	1.8%	41	341
Florida	-	0.0%	-	-
Georgia	12,177	12.4%	1,345	845
Hawaii	2,568	6.5%	158	656
Iowa	4,098	9.0%	337	669
Idaho	1,669	7.7%	119	425
Illinois	18,762	7.1%	1,238	589
Indiana	8,853	7.0%	577	548
Kansas	3,778	3.3%	122	261
Kentucky	4,629	9.0%	381	578
Louisiana	3,833	6.5%	235	353
Massachusetts	17,151	3.5%	586	467
Maryland	10,067	5.1%	484	471
Maine	1,709	8.2%	129	549
Michigan	10,150	12.4%	1,122	705
Minnesota	12,400	2.2%	272	286
Missouri	6,588	5.0%	313	327
Mississippi	1,968	12.0%	211	523
Montana	1,413	3.9%	53	290
North Carolina	13,259	6.1%	766	474
North Dakota	415	2.4%	10	76
Nebraska	2,546	2.5%	63	204
New Hampshire	123	63.0%	47	190
New Jersey	15,903	3.7%	570	349
New Mexico	1,590	3.0%	46	147
Nevada	-	0.0%	-	-
New York	48,087	4.2%	1,955	580
Ohio	9,313	5.9%	519	284
Oklahoma	3,565	8.0%	264	459
Oregon	9,847	6.5%	598	851
Pennsylvania	13,517	8.1%	1,013	476
Rhode Island	1,376	2.8%	38	211
South Carolina	4,759	13.4%	561	718
South Dakota	-	0.0%	-	-
Tennessee	202	66.1%	80	77
Texas	-	0.0%	-	-
Utah	4,979	2.1%	104	218
Virginia	14,918	6.0%	844	592
Vermont	861	2.8%	23	206
Washington	-	0.0%	-	-
Wisconsin	8,760	7.0%	570	586
West Virginia	2,097	7.7%	149	569
Wyoming	-	0.0%	-	-
Total	411,235	5.0%	19,583	369

Note: This table reports the estimates of the cost of elderly state individual income tax benefits for 2019, using data from the SOI Databank, other IRS sources, and the State and Local Finance (SLFI) Data Query System (DQS) The Urban Institute and The Urban-Brookings Tax Policy Center (2024). The first column reports the State (and local, in the case of DC) individual income tax (IIT) revenue via the US Census of Governments, collected and organized by SLFI DQS, in millions of real 2019 USD. The percent effect columns shows the estimated percent decrease in IIT revenue due to the elderly benefit, calculated using a stratified random sample of tax returns and the NBER TAXSIM model. The implied forgone revenue shows the cost in millions of real 2019 USD of these elderly policies, and the final columns shows the average elderly benefit per return (elderly and non-elderly).

use consistently shows rates of interstate elderly migration that are roughly twice as large (2 percent per year versus 1 percent) as those found in publicly available data. This is true even for research on the period since 2000 using the American Community Survey and other sources (Conway and Rork, 2016). This finding implies that the elderly are more mobile than previously thought and so the potential for tax-induced migration is also even stronger. However, the overall level of migration among the elderly is still lower than those of younger tax-payers, substantially so relative to individuals aged 25 to 55.

To answer the second question, we estimate an aggregate migration flow model for different age and AGI groups across all states and years. This approach has the advantage of capturing the effect of all tax changes, both large and small, direct and indirect, targeted and across-the-board. We find that the elderly are responsive to overall changes in tax rates, and that this response is larger than that of younger taxpayers. However, we find no measurable effect of targeted elderly tax breaks, implying that elderly-specific tax policies are not an optimal way to attract or retain elderly residents. We are planning on extending this analysis to focus on a series of case studies of large, easily visible policy changes targeting the elderly using a difference-in-difference approach applied to individual level data.

Finally, we estimate the cost of the set of individual income tax policies focused on elderly individuals, finding that across all states in 2019, elderly-specific policies reduced potential individual income tax revenue by five percent, or approximately $20 billion dollars in 2019. While this estimate is larger than the total additional lost revenue due to the policy changes we study during the 2000-2019 period, it provides a sense of the magnitude of the lost revenue due to policies that, fundamentally, do not appear to be responsible for shifting migration decisions. These results therefore suggest that the answer to the third question is likely yes – with the growing size

and affluence of the elderly, the fiscal consequences of these policies are likely growing as well. Our results thus far suggest that those migration effects will be small and that the fiscal cost of recruiting residents is quite large.

Chapter 4: Comment on "Income Inequality in the United States: Using Tax Data to Measure Long-Term Trends" by Auten and Splinter

Abstract[1]

We develop a framework for accounting for tax noncompliance when estimating top 1% income shares and apply it to parse disagreements between Auten and Splinter (AS) and Piketty, Saez, and Zucman (PSZ). AS' methods miss the changes in misreporting that likely accompanied the recent increase in the importance and concentration of pass-through business income. Correcting this via a distributional neutrality assumption, we find that AS' estimate of the change in the top 1% income share from 1962 to 2019 is underestimated by about 1.1 p.p, while PSZ slightly overestimate this change, by 0.1 p.p., due to an excess re-ranking effect.

[1]Co-Authored with Daniel Reck (University of Maryland). We thank Gerald Auten, Judy Hellerstein, Ethan Kaplan, Melissa Kearney, David Splinter, Gabriel Zucman, *JPE* Editor Andrew Atkeson and an anonymous referee for their helpful feedback on this comment. This work was begun and all primary analysis completed prior to John Iselin beginning his employment at the Congressional Budget Office. The views expressed in this paper are the authors' and should not be interpreted as those of the CBO.

4.1 Introduction

Has the top 1% share of income in the United States risen sharply since the 1980s? Two recent, high-profile attempts to answer this question – Piketty et al. (2018) [henceforth PSZ] and Auten and Splinter (2024) [AS] – reach diverging answers. Both studies begin with income reported on tax returns, like Piketty and Saez (2003) [PS]. They then introduce additional assumptions and data to transition from the distribution of *fiscal income* – Adjusted Gross Income minus capital gains and government benefits – to the distribution of *National Income.* Similarly to PS, PSZ estimate that the top one percent share of Pre-Tax National (PTN) Income has risen by 6.2 percentage points between 1962 and 2019, while AS estimate a much smaller 2.6 percentage point increase. In this comment, we study a major source of divergence between AS and PSZ: income that is unobserved on tax returns due to tax non-compliance. Misreported income accounts for 3-7% of National Income between 1960 and 2020 and its distribution generates one-third of the total divergence between PSZ and AS, more than any other factor (see AS Table 4). We do not engage with methodological differences besides misreporting and the distribution of pre-tax income.

We propose a framework for understanding and evaluating the assumptions made by both sets of authors; we use this to parse the points of contention that have arisen in prior debates.[2] Our approach allows us to parse the drivers of the divergence between PSZ and AS, which we review below, including how these methods account for re-ranking, complexities arising from negative income, and – most importantly – the effect of the underlying economic assumptions these authors are making regarding how much to utilize audit data verses the distribution of

[2]See Auten and Splinter (2021), Reck et al. (2021), Gale et al. (2023), Piketty et al. (2024), Auten and Splinter (2025a,b).

reported income.

Our main argument is that AS' methods do not account for dramatic changes in the importance and concentration of pass-through business income since the 1980s and their likely effect on the distribution of misreporting. We find that their estimates impose unrealistic structure on the dynamic evolution of pass-through misreporting during this period, and we attribute this to specific limitations of their microsimulation method for allocating misreporting. PSZ's estimates are closer to our preferred estimates because they are based on reported incomes. However, we argue that re-ranking effects in PSZ's approach may lead them to slightly over-estimate the increase in the top 1% income share. Our preferred estimates suggest that relative to the top 1% share of fiscal (pre-tax National) income excluding misreporting, adding misreported income increases the change in the top 1% share between 1962 and 2019 by 0.8 pp (0.6-0.7 pp), which is 1.5 pp (1.0) higher than AS' estimate and 0.2 pp (0.1 pp) lower than PSZ's.

4.2 Framework

We focus on fiscal income first and turn to PTN Income in Section 4. For individual i in year t, $y_{it} = r_{it} + m_{it}$, where y_{it} is fiscal income inclusive of misreporting (*corrected fiscal income*), r_{it} is reported fiscal income, and m_{it} is misreported income. We define the top 1% share of reported and misreported income (ranked by corrected income, denoted by superscript c) as:

$$R^c_{1t} \equiv \frac{\sum_{i \in \text{[Top 1\% by y]}} r_{it}}{\sum_i r_{it}}, \qquad \text{and} \qquad E^c_{1t} \equiv \frac{\sum_{i \in \text{[Top 1\% by y]}} m_{it}}{\sum_i m_{it}}.$$

The top 1% share of corrected income, Y^c_{it}, is the average of the top 1% share of reported (R^c_{1t}) and misreported (E^c_{1t}) income, weighted by the share of total income in each category

$(1 - M_t, M_t)$.

$$Y_{1t}^c = (1 - M_t)R_{1t}^c + M_t E_{1t}^c. \tag{4.1}$$

We use the estimates of total annual misreporting from the National Income and Product Accounts (NIPA) to quantify M_t. These estimates are mainly drawn from IRS Tax Gap estimates, which include misreporting *detected* in audits of randomly selected tax returns together with IRS estimates of *undetected* misreporting.[3] The detected component comprises about one third of total misreporting; the remainder is undetected.

We simplify the R_{1t}^c term by assuming that re-ranking from reported income r_{it} (observed on tax returns) to corrected income y_{it} (unobserved) income has a negligible effect on the concentration of *reported* income. Figure C.1 validates Assumption 1 empirically.[4] [5]

Assumption 1: Negligible Reranking for Reported Income.

$$R_{1t}^c = R_{1t}^r \equiv \frac{\sum_{i \in [\text{Top 1\% by } r]} r_{it}}{\sum_i r_{it}}.$$

[3]The estimate of total undetected misreporting comes from a Detection Controlled Estimation (DCE) method, which utilizes estimates of each auditor's effectiveness to determine how much misreporting would have been detected if all auditors have the detection propensities of the auditors with the highest propensities. DCE methods are reviewed in more detail in Guyton et al. (2023a). See also Feinstein (1990, 1991), Erard and Feinstein (2010, 2011).

[4]It cannot be satisfied perfectly, but the bias from its failure is about 0.2 percentage points for the level of the top 1% share and perhaps 0.1 pp for its change over time. Most misreporting is under-reporting, so re-ranking from reported to corrected income will mechanically make reported income move downwards in the distribution.

[5]NIPA draws totals from Tax Gap studies to account for misreporting of wages and salary income and non-farm proprietor income (sole proprietorship and partnership business income). NIPA does not rely on tax data to estimate total income from rents, S corporations, or farms. For these components, the difference between NIPA totals and fiscal income totals could be due to misreporting or other accounting/statistical discrepancies. AS argue that for these categories, the discrepancy is similar to misreporting estimated in Tax Gap studies, and we follow them in including these in total misreporting. We note that even if we knew that the differences between NIPA and tax totals for these components were due to statistical noise rather than misreporting, our preferred approach would be to employ a distributional neutrality assumption to transition from fiscal to National Income, like our preferred approach for misreporting.

We can now focus our attention on the remaining unknown in Equation (4.1): E^c_{1t}. AS estimate E^c_{1t} in two steps. They first simulate audit-detected under-reporting building on Auten and Langetieg (2023); then they use a "multiplier" method proposed in Auten and Splinter (2021) to distribute undetected under-reporting. The first step is empirically validated by audit data; the second step is ad-hoc. Their estimate for E^c_{1t} is stable over time at about 18%.

Assumption 2.1 (AS 2024) In each period t, $E^c_{1t} \approx 18\%$.

PSZ allocate misreporting in proportion to non-negative reported income by type of income j, $\max[r_{ijt}, 0]$, and then re-rank individuals after this allocation. We label their estimate E^{PSZ}_{1t}. Using their replication file, we find that E^{PSZ}_{1t} rises from about 20% to 40% between 1962 and 2019.

Assumption 2.2 (PSZ 2018) $E^c_{1t} = E^{PSZ}_{1t}$.

The gap between these two estimates, plotted in Panel 4.1a, is driven by two factors. The first is that reported incomes increase sharply in concentration in the top 1% over time; PSZ assume that this increase is accompanied by an increase in the concentration of misreported income, while AS suppose that it is not. The second, the focus of AS's critiques of PSZ, is a re-ranking effect whereby individuals already reporting high, positive income are allocated the largest amounts of misreporting, so that E^{PSZ}_{1t} increases by more than one would expect based on R^r_{1t}. To isolate these effects, we propose an estimate of E^c_{1t} based on *distributional neutrality by type of income*, i.e. that the concentration of income of each type is not affected by adding misreporting.

Assumption 2.3 (Our Preferred Approach) For each t, j, $E^c_{1jt} = \frac{\sum_{i \in [\text{Top 1 percent by r}]} \max[r_{ijt},0]}{\sum_i \max[r_{ijt},0]}$.

Assumption 2.3 does not specify a micro-level allocation of under-reporting, but in the aggregate it is identical to PSZ's method without re-ranking after allocating misreporting. Com-

paring Assumption 2.2 to 2.3 thus captures the excess re-ranking effect, while the difference between Assumptions 2.1 and 2.3 reflects the conceptual disagreement about whether misreported income increases in concentration when reported income does. In Appendix C.3 we discuss the use of positive income in Assumption 2.3 and present sensitivity analysis. Assumption 2.3 does *not* impose that those with negative reported incomes are perfectly tax compliant.[6] Compared to perfect distributional neutrality by type of income – $E^c_{1jt} = R^c_{1jt}$ – Assumption 2.3 implies a smaller increase in the top 1% income share over time.

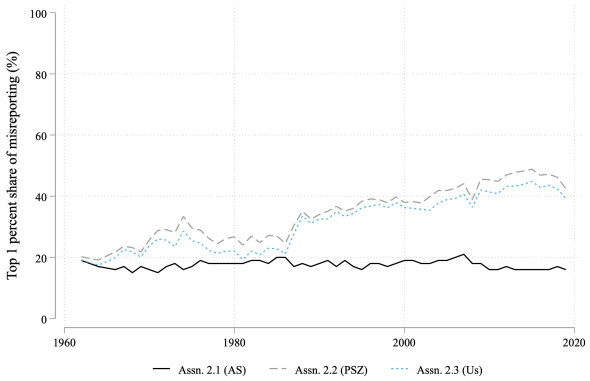
In Panel 4.1b, we plot the difference between the top one percent share of reported fiscal income (R^r_{1t}) and the estimate of Y^c_{it} under each of these assumptions about E^c_{1t}. In PSZ's specification and our preferred one, the top 1 percent share of fiscal income increases after correcting for to misreporting, and this effect grows by about 1 p.p. from 1962 to 2019. With AS's specification: the top one percent share instead falls by 0.7 p.p. over the same period. Intuitively, reported incomes are increasing in concentration and AS assume that misreported income is not, so mixing these two distributions flattens out the increase in the top 1% share over time. Panel 4.1c illustrates the implications of these assumptions for misreporting as a share of corrected income in the bottom 99% and top 1%. AS' estimates imply that this share is decreasing over time in the top 1% and increasing in the bottom 99%, while PSZ's and our preferred approach suggest that the misreporting share is roughly constant in the bottom 99% and increasing over time in the top 1%.

[6]Formally, it is straightforward to verify that

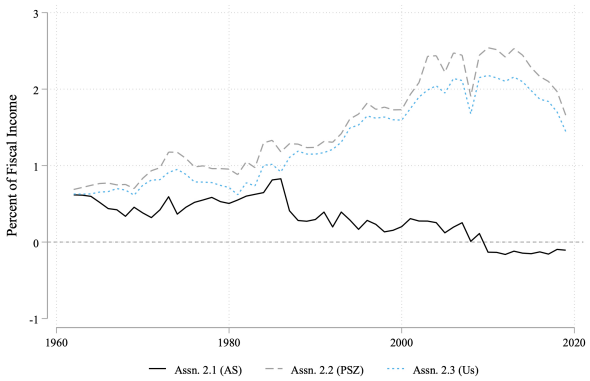
$$E^c_{1jt} = R^c_{1jt} \text{ for each } j \iff \frac{\sum_{i \in [\text{top 1 \% by y}]} m_{ijt}}{\sum_{i \in [\text{top 1 \% by y}]} y_{ijt}} = \frac{\sum_{i \in [\text{bottom 99\% by y}]} m_{ijt}}{\sum_{i \in [\text{bottom 99\% by y}]} y_{ijt}} \text{ for each } j.$$

Figure 4.1: Assumptions about Misreporting Allocated to the Top 1 Percent and Their Effects

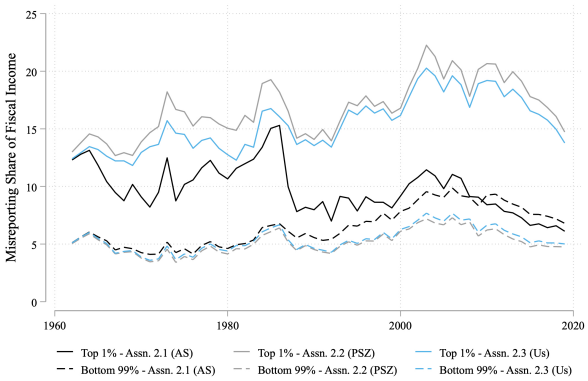
(a) Share of misreporting in the top 1% (E^c_{1t})



(b) Effect of misreporting on the top 1% share of fiscal income ($R^c_{1t} - Y^c_{1t}$)



(c) Misreporting as a share of corrected fiscal income, top 1% vs bottom 99%



Notes: For the first panel, the solid black line plots the share of total (detected and un-detected) misreporting allocated to the top 1 percent of the true income distribution in Auten and Splinter (2025a) (Assumption 2.1 (AS)). The dashed gray line represents the share of total misreporting allocated to the top 1 percent based on reported income following Piketty et al. (2018); namely, allocating misreporting proportional to reported positive income then re-ranking by corrected fiscal income (Assumption 2.2 (AS)). The dotted blue line shows the effect of eliminating the re-ranking effect, allocating misreporting proportional to reported positive income without re-ranking (Assumption 2.3 (Us)). For the second panel, we use the estimates of E^c_{1t} and plot the difference between the top 1 percent share of fiscal income (R^r_{1t}) excluding capital gains and the estimates of Y^c_{1t} derived from Assumptions 2.1, 2.2, and 2.3 (Our approach). The third panel plots the misreporting share of corrected income (fiscal income plus misreporting) for the top one percent and bottom 99, under the different assumptions.

4.3 Conceptual Argument

Fundamentally, then it seems that the main difference between the estimates of E^c_{1t} for AS and PSZ, as shown in Figure 4.1, stems not from issues of re-ranking or modest changes in the treatment of negative incomes, but a the disagreement regarding the degree to which reported income or audit data can provide more information misreporting at the top of the income distribution. We believe that, given the relationship between under-reporting and the rise in the concentration of pass-through business incomes at the top of the income distribution, random audit data likely does not capture key features of the evolution of high-income tax evasion. In our view, both approaches require making some fundamentally untestable (at least with existing random audit data) assumptions. In this section, we briefly review this issue and connect it do the debate between these authors.

We first observe that one economic factor explains virtually all of the divergence between these estimates: the rise in the importance of pass-through business forms (Smith et al., 2019, Kopczuk and Zwick, 2020).[7] Figure 4.2 establishes that the dynamics we observe in Figure 4.1 are entirely driven by the "rise of pass-throughs." Panel 4.2a shows that the top 1% share of reported non-wage business income (R^r_{1jt} for $j =$ non-wage business) grows from 25% to almost 60% over time; the concentration of reported wage income has changed relatively little. Panel 4.2b shows that pass-through income also became more important as a share of overall fiscal income during this period. The increase in top 1% misreporting rates for our preferred estimates in Panel 4.1c is driven by the increase in the importance and concentration of pass-through income

[7] Appendix C provides some further details involving changes to the tax system due to the Tax Reform Act of 1986, including substitution between C corporation and pass-through forms and partnership tax shelters limited by TRA86.

at the top, while the share of pass-through income that is misreported remains constant (see Appendix Figure C.4).

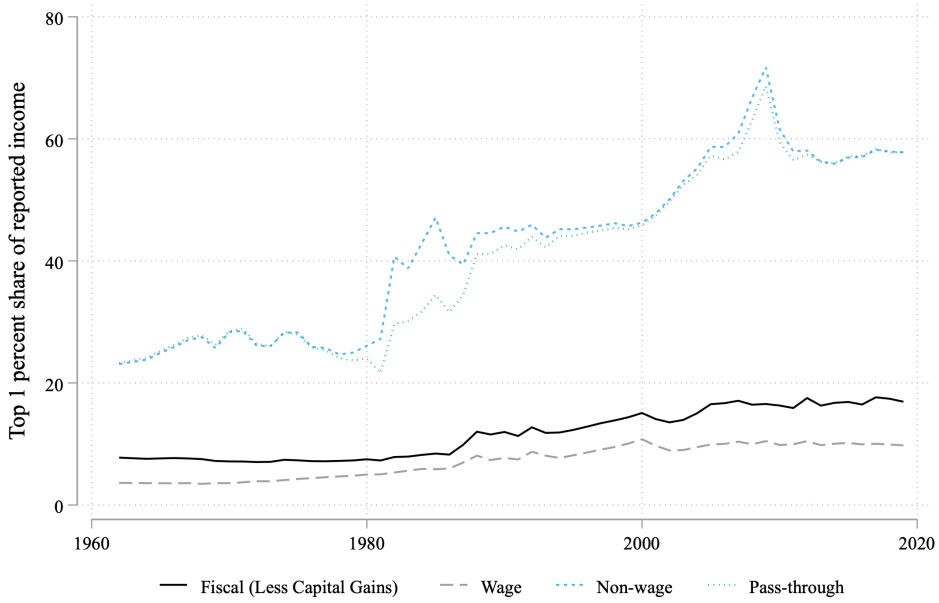
This change in the importance and concentration of business income is the main reason that, given (approximate) distributional neutrality by type of income, we find an increase in the top 1% share over time due to misreporting, and an increase in the share of income misreported in the top 1%. To verify this, we plot the misreporting shares from Panel 4.1c separately for wage and non-wage business income under Assumptions 2.2 and 2.3 in Figure C.4. This highlights that the increasing misreporting share under these Assumptions in Panel 4.1c is not driven by changes in the share of non-wage business income that is assumed to be misreported by the top 1%. About 45% of non-wage business income is estimated to be misreported in every year (per NIPA) and this figure is very similar in the top 1% and bottom 99% under Assumption 2.3. Rather, the overall increase in misreporting rates at the top in Panel 4.1c is driven by the increase in the importance of pass-through income at the top from Figure 4.2.

We observe that Assumption 2.1 (AS' approach) requires that the rise of pass-throughs observed in reported incomes was not associated with an increase in the share of unreported pass-through income belonging to the top of the distribution. This in turn implies that the rise of pass-throughs was associated with decreasing misreporting rates for pass-through income in the top 1%, and increasing misreporting rates in the bottom 99%. We do not have the data to break down AS' misreporting rates by type of income, but this must be the sole driver of the changes over time we see in Figure 4.1c.

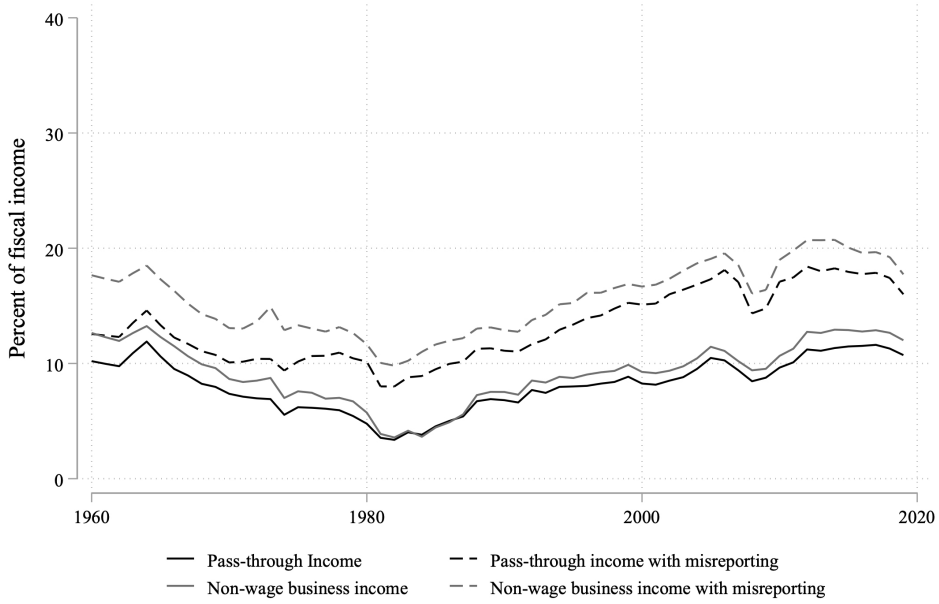
In contrast, AS' approach imposes that the misreporting rate for pass-through income is decreasing over time in the top 1% and increasing in the bottom 99%. We attribute this to limitations of their method: most pass-through income in the top 1% comes from partnerships and

Figure 4.2: Dynamics of the Composition of Fiscal Income

(a) Top 1 percent shares of reported income



(b) Non-wage business income as a share of total fiscal income



Notes: Data for the first panel comes from the publicly available DINA files, and data for the second panel are derived from the supplementary data associated with Auten and Splinter (2024). The first panel plots the rise in the top one percent share of business income (defined by reported sole proprietors', partnership, farm, rental, and s-corporation income) and pass-through income (defined by reported sole proprietors', partnership, and s-corporation income). Wage inequality rose slightly, but not as dramatically as non-wage business income. The second panel tracks the importance of non-wage and partnership income (defined as above) with and without misreporting.

S corporations, while the bottom 99% receives more sole proprietorship income. Sole proprietor income was comprehensively examined in random audits, but partnership and S corporation income was not, due to audit procedures (Guyton et al., 2023b).[8] Consequently, the first step of AS' microsimulations based on random audit data will miss pass-through misreporting especially in the top 1%. If we take total misreporting per NIPA at face value, then most entity-level misreporting in the top 1% must load onto undetected misreporting. In the second step, AS use the same mulitpliers for undetected misreporting in every year. Thus, neither step plausibly accounts for changes in misreporting due to the rise of pass-throughs.

The available data suggests that the rise of pass-throughs was associated with increases in misreporting in the top 1%. Indeed, taking NIPA totals as given and assuming distributional neutrality may be too conservative, because sophisticated types of pass-through misreporting are especially common in the top 1%. Guyton et al. (2023b) analyze data from a small-scale random audit study of S corporations in 2003-4 and estimate that about 19% of corrected S corporation business income is under-reported, without adjusting for undetected misreporting, which is higher than individual Tax Gap studies would suggest. A recent study of Portuguese expenditure data finds that high-income owners of closely held businesses routinely re-classify personal consumption expenditures as business expenses, to an extent that matters quantitatively for the magnitude of the top 1% income share in Portugal (Leite, 2024). [9] Guyton et al. (2023b)

[8]More specifically, the random audit data are missing *entity-level* misreporting of income because auditors in the individual NRP virtually never examined pass-through business returns of the business's owners; refer to Guyton et al. (2023b) and our reply to AS' response to the first draft of this comment for more detail. Slemrod (2007) suggests that the same limitation applies to TCMP data from earlier years and cautions against using random audit data to argue that compliance is relatively high in the high-income population.

[9]Without undetected misreporting, entity-level pass-through under-reporting of S corporation and Partnership business income is about 3% of corrected income (before DCE); the figure is about 9% if we undetected under-reporting identified by DCE. These figures are drawn from the rate of under-reporting for S corporation and Partnership income in Guyton et al. (2023b) Tables A1 (excluding DCE) and A6 (including DCE). These authors report that about 56% of detected under-reporting in these data is entity-level misreporting, so here we are scaling the rates of misreporting we find in the Appendix Tables by 60% to account for the share that is misreported at the entity level.

also find that sophisticated sheltering schemes that leverage pass-through ownership structures are far more prevalent at the top of the US income distribution than elsewhere. Multi-tiered partnership ownership structures are concentrated among top 1% owners. Recent work on audit data by Black et al. (2023) shows that auditing complex partnerships generates $20 of tax revenue for each $1 of administrative costs, twice that of auditing simple partnerships.

The exact quantitative implications of all this are uncertain. We lack representative, comprehensive data on misreported income, but the evidence we have suggests that it is reasonably conservative to assume a similar the rate of under-reporting of pass-through income between income groups. There is also uncertainty about several issues that are mostly unrelated to the rise of pass-throughs. This includes biases in estimated undetected evasion (Hemel et al., 2021), changes over time to estimation methods and audit procedures (Feinstein, 1990, Erard and Feinstein, 2010, Black et al., 2012, Johnson and Rose, 2019), and changes to offshore tax evasion (Guyton et al., 2023b, Laffitte, 2024). On balance, our review of the extant literature on these other issues suggests that they may matter significantly for the level of the top 1% income share, but they are unlikely to matter nearly as much as the rise of pass-throughs for its dynamic evolution.

4.4 From Fiscal to National Income

Finally, we return to the bigger picture: the distribution of National Income. In Table 4.1 and Figure C.3 we illustrate how allocating misreporting per Assumption 2.1 or 2.2 interacts with the other assumptions AS and PSZ make in transitioning from reported fiscal to PTN income.[10]

[10]Table 4 in Auten and Splinter (2024) breaks down the impact of different assumptions similarly, for the period 1962 and 2014. In Table C.1, we reproduce Table 4.1 for these years. There, using AS' PTN assumptions, we find that the gap between Assumption 2.1 and Assumption 2.2 is 1.5 percentage points, very close to the 1.6 shown in

Table 4.1: Estimates of the top 1 percent share, varying R^c_{1t} and E^c_{1t}

	Income Excl Mis-Reporting	Assn. 2.1 (AS)	Assn. 2.2 (PSZ)	Assn. 2.3 (Us)
		Reported Fiscal Income		
1962	8.3	8.9	9.0	8.9
2019	17.6	17.5	19.2	19.0
Change	9.3	8.6	10.3	10.1
		Pre-Tax National Income (AS)		
1962	10.8	11.1	11.2	11.1
2019	13.7	13.8	14.9	14.8
Change	2.9	2.6	3.7	3.6
		Pre-Tax National Income (PSZ)		
1962	12.5	12.8	12.9	12.8
2019	18.0	17.9	19.1	18.9
Change	5.5	5.1	6.2	6.1

Notes: This table reports the top 1 percent shares of income (Y^c_{1t}) for 1962 and 2019, as well as the increase in the top 1 percent share (in percentage points) over this period. Each panel corresponds with a different estimate of R^c_{1t}, and each column with a different estimate of E^c_{1t}, or the pre-misreporting top 1 percent share of reported fiscal income or pre-tax National Income under the assumptions of AS and PSZ.

We do this by re-defining income as PTN income within our framework: M_t is now total misreporting as a share of total PTN income, and R^r_{1t} (assumed equal to R^c_{1t}) is the top 1% share of PTN income excluding misreporting, whose distribution we obtain using the authors' data files. [11]

The assumptions AS make to transition from fiscal to PTN income flatten out the change in the top 1% income share over time. This brings their allocation of misreporting much closer to distributional neutrality for overall PTN income (but not neutrality by type of income). For fiscal income, AS's treatment of misreporting decreased the top 1% share by 0.76 percentage points, but

AS' Table 4. This comparison provides us some additional assurance that Assumption 1 is introducing negligible bias for our purposes.

[11] We calculate this figure separately for AS and PSZ using the following equation: $R^{\text{PTN, No M}}_{1t} = \frac{Y^{PTN}_{1t} - M_t * E_{1t}}{1 - M_t}$. This is based on re-arranging (4.1).

for PTN income, this drops to 0.24 percentage points.[12] With PSZ's method, switching from fiscal to PTN income results in a level shift downwards in the top 1% share, and adding misreporting increases the top one percent share less with a PTN baseline than with a fiscal income baseline. The dynamics change slightly, with misreporting increasing the top one percent share between 1962 and 2019 by 1.0 p.p. for fiscal income and by 0.8 p.p. with PTN income.

What one makes of these interactions obviously depends on what one believes about the other sources of divergence in these authors' estimates. We observe that the rise of pass-through forms and the expansion of excess depreciation allowances ("capital consumption" adjustments in NIPA) generates most of the remaining divergence and debate between these authors (see PSZ's recent comment and AS' reply). The other disagreements, around such issues as fringe benefits and retained earnings in C corporations, appear to be of minor quantitative importance.

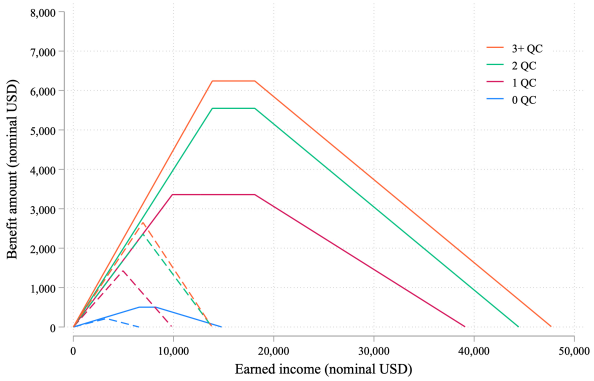
In conclusion, our preferred estimates are closer to those of PSZ for misreporting specifically. The available evidence suggests that the rise of pass-throughs should matter similarly for the reported and misreported business income of the top 1%. We propose a sounder implementation of PSZ's stated rationale for their method and find that this slightly dampens the increase in top-end inequality over time. A common feature of both studies is that due to re-ranking effects, using microsimulations to distribute misreporting makes it difficult to understand the macro-level, conceptual rationale behind each method. Here, it would be very useful to improve our theoretical understanding of distributional estimation with measurement error. We hope future research can shed some light on the formal relationship between structure imposed via microsimulations with re-ranking and macro-level structure like the assumptions of our framework.

[12] This finding also reconciles our illustrations with AS Figure B6. We observe some minor quantitative differences between AS Figure B6 and the analogous series in Figure C.3. This must be driven by Assumption 1 and/or the bundling in AS Figure B6 of capital consumption adjustments (e.g. excess depreciation allowances) together with misreporting. We note the divergence after 2008 coincides with the expansion of excess depreciation allowances.

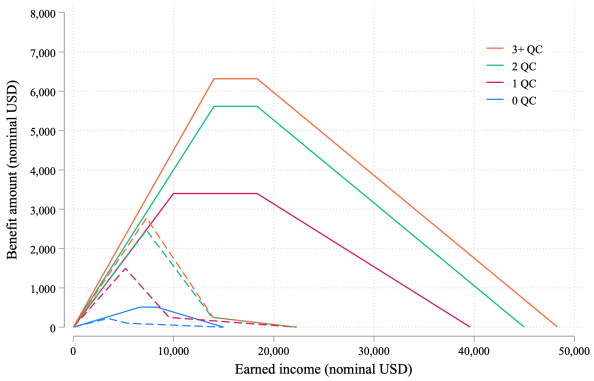
Appendix A: Chapter 2 Appendices

A.1 Additional Tables and Figures

Figure A.1: Federal and California EITC benefits schedule for 2015 and 2017.



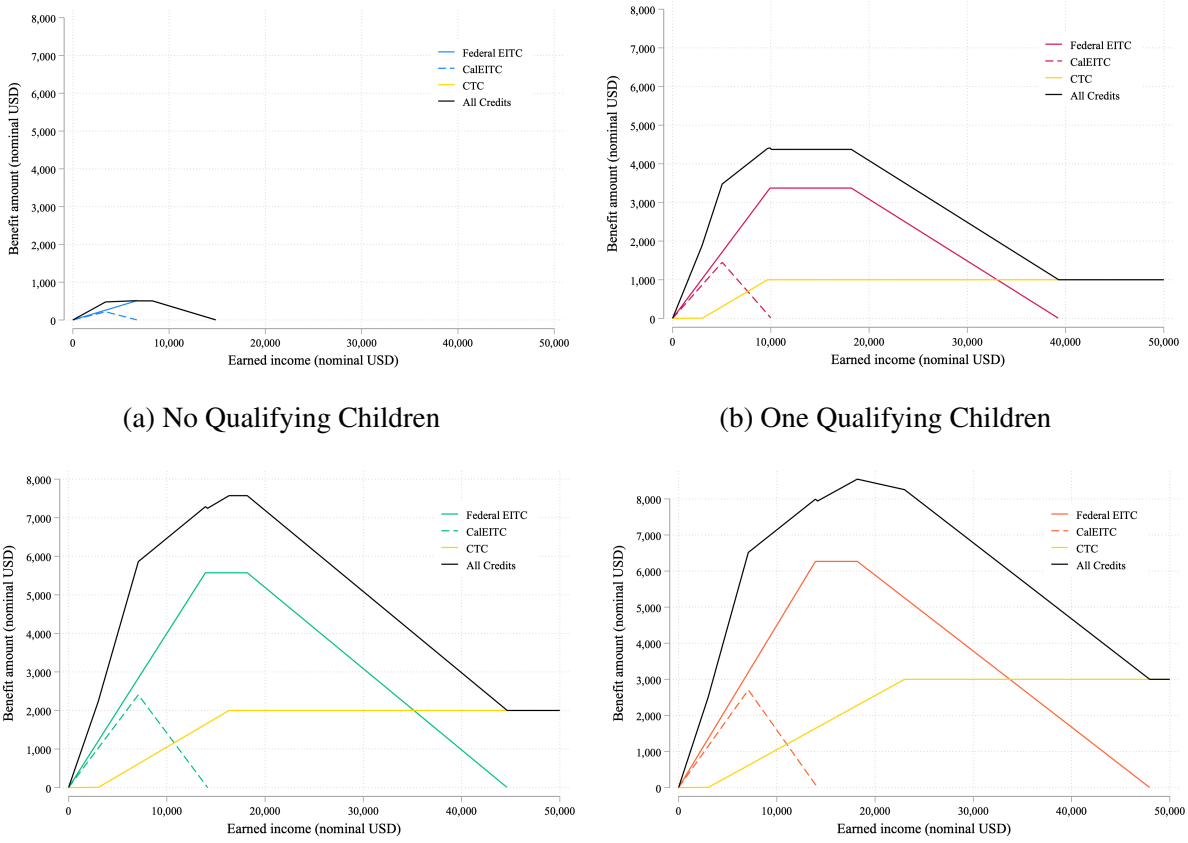
(a) Tax Year 2015, by count of qualifying children



(b) Tax Year 2017, by count of qualifying children

Note: The parameters of the federal EITC used to create this figure were obtained via the Tax Policy Center, while the parameters of the California EITC schedule were backed-out of the Earned Income Tax Credit tables within the California Form FTB 2514. The benefit schedules shown assume a single filer whose adjusted gross income is solely composed of wage income.

Figure A.2: Benefits schedules for the federal EITC, CTC, and CalEITC for TY 2016



(a) No Qualifying Children

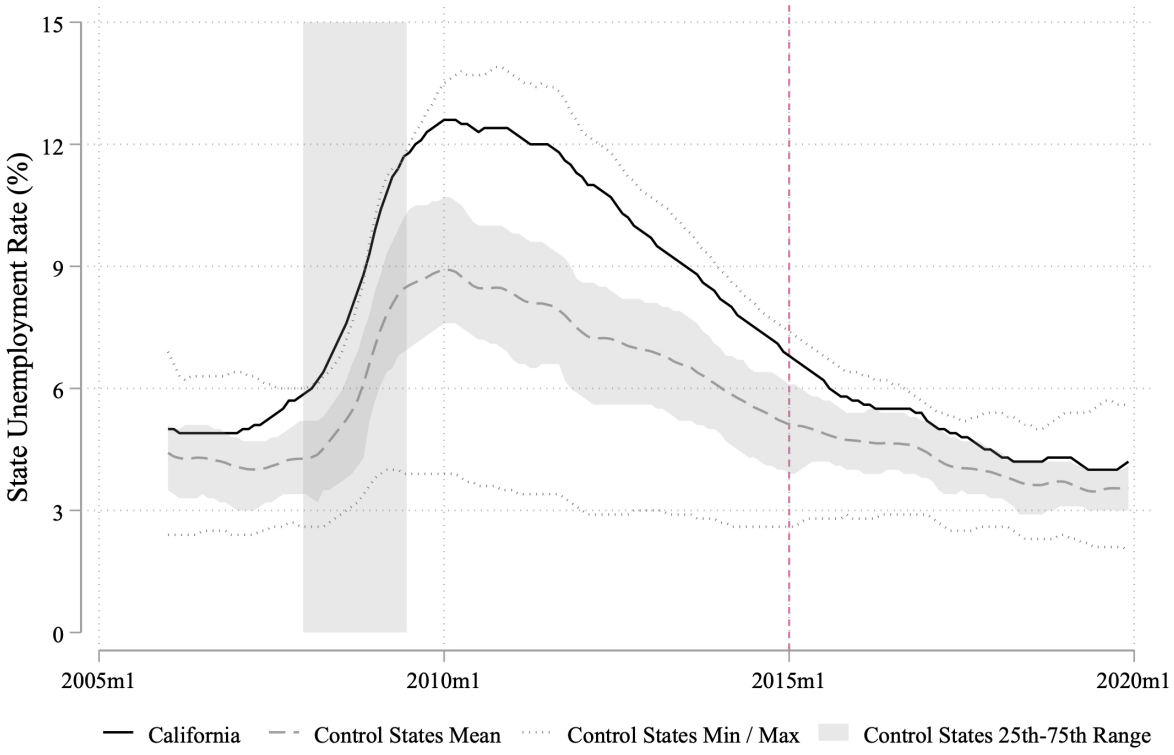
(b) One Qualifying Children

(c) Two Qualifying Children

(d) Three or More Qualifying Children

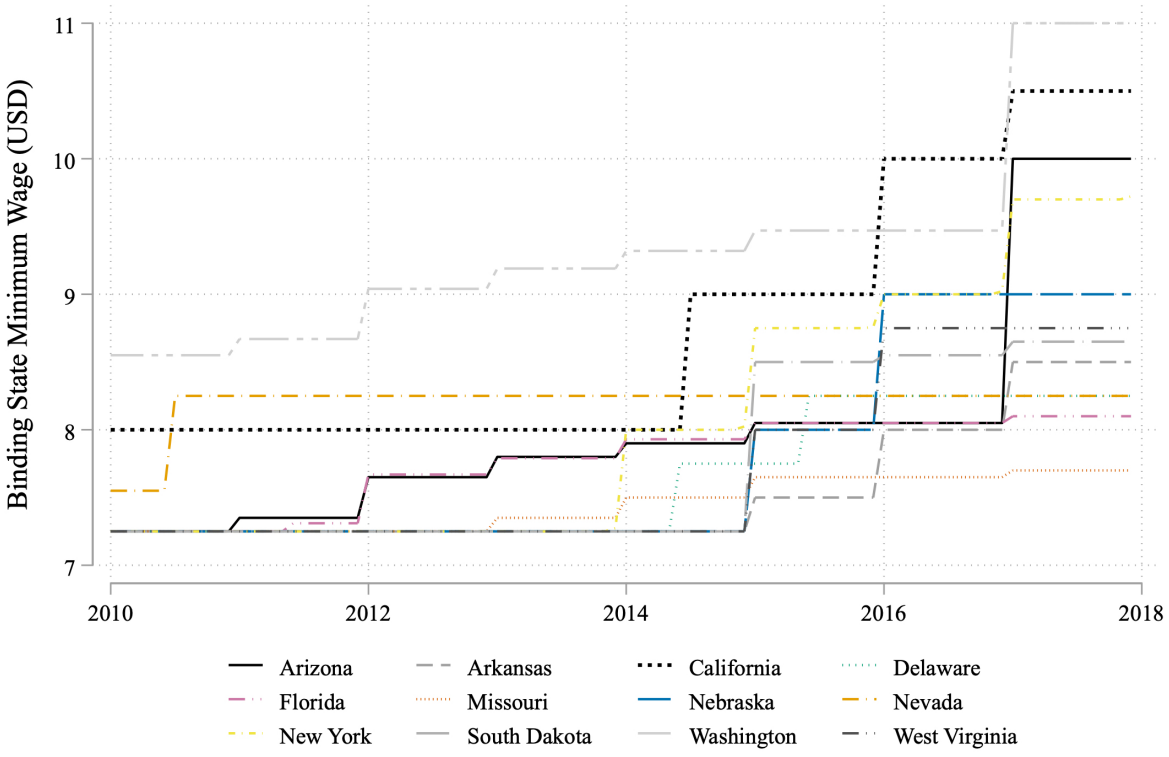
Note: The parameters of the federal EITC and CTC used to create this figure were obtained via the Tax Policy Center, while the parameters of the California EITC schedule were backed-out of the Earned Income Tax Credit tables within the California Form FTB 2514. The benefit schedules shown assume a single filer whose adjusted gross income is solely composed of wage income, as self-employment income was not made a part of the definition of earned income for the purpose of the CalEITC until 2017. The changes in color across the sub-plots for the federal EITC and CalEITC schedules aligns with the standard coloring across the count of qualifying children throughout this paper (blue for zero, red for one, green for two, and orange for 3 or more).

Figure A.3: State-level trends in unemployment (2006-2019)



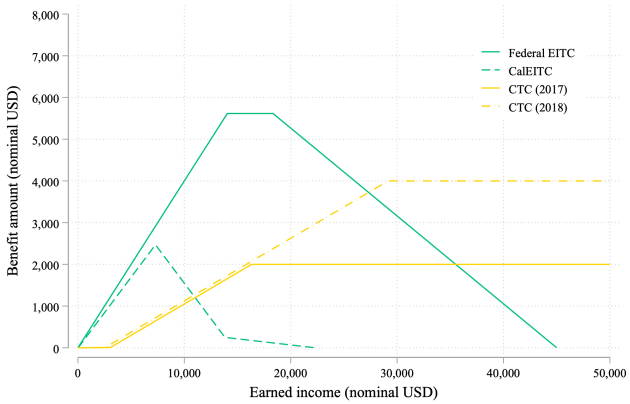
Note: Data from this figure are from the Bureau of Labor Statistics (BLS) Local Area Unemployment Statistics (LAUS). The shaded horizontal bar denotes the 2008-2009 recession. The solid black line is the monthly unemployment rates for California. The dashed gray line is the mean unemployment rate in the pool of control states. The gray area plot shows covers the 25-75 percentiles of the unemployment rate in control pool in any given month, and the dotted lines represent the maximum and minimum state unemployment rate among those states. The vertical purple dashed line shows the beginning of 2015. California, during the recovery from the Great Recession, has one of the highest levels of unemployment.

Figure A.4: Binding state minimum wages in control pool, 2010 through 2017

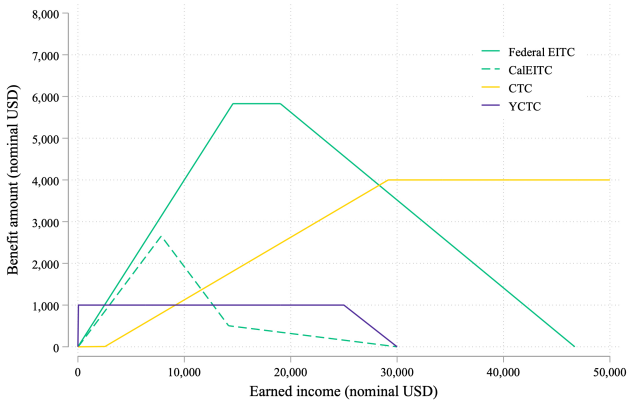


Note: This figure plots the monthly state-level binding minimum wages among the pool of states used in my main analysis sample who had a state minimum wage in excess of the federal minimum wage at any point between 2010 and 2017. Data on minimum wages are from Vaghul and Zipperer (2022).

Figure A.5: Benefits schedules for a single filer with 2 QC: 2018 and 2019 with TCJA CTC and Young Child Tax Credit



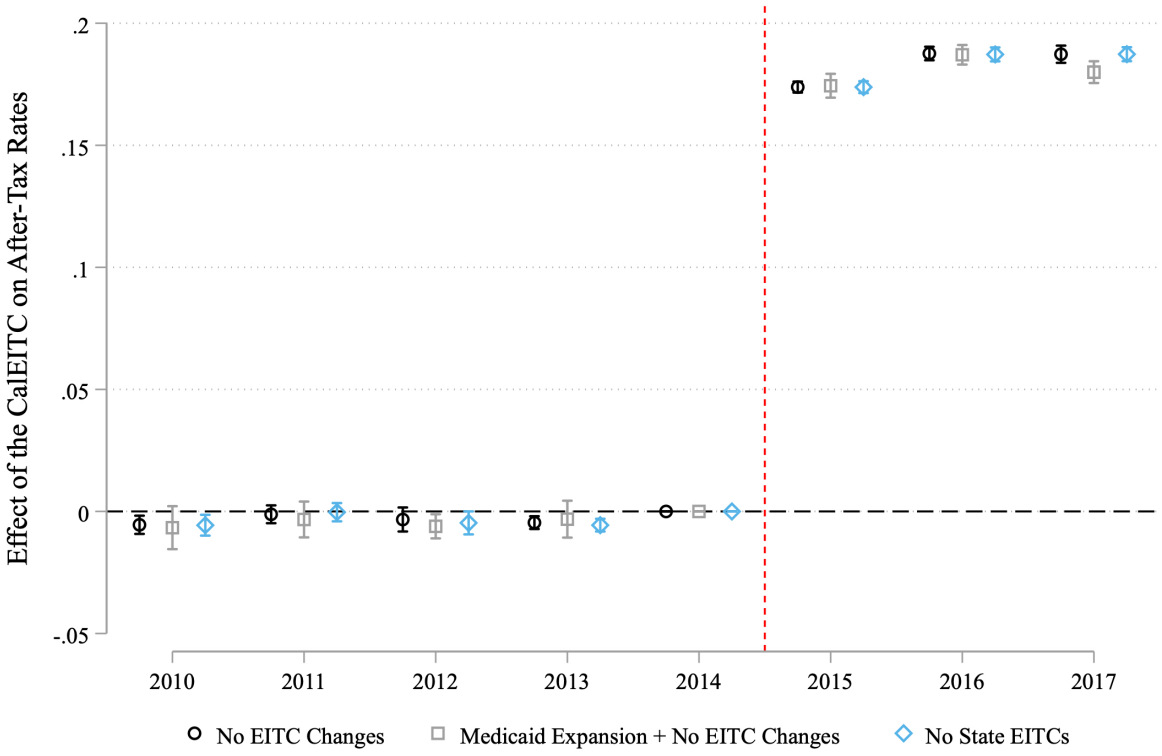
(a) TY 2018 Schedules with Pre- and Post-TCJA CTC



(b) TY 2019 Schedules with Young Child Tax Credit

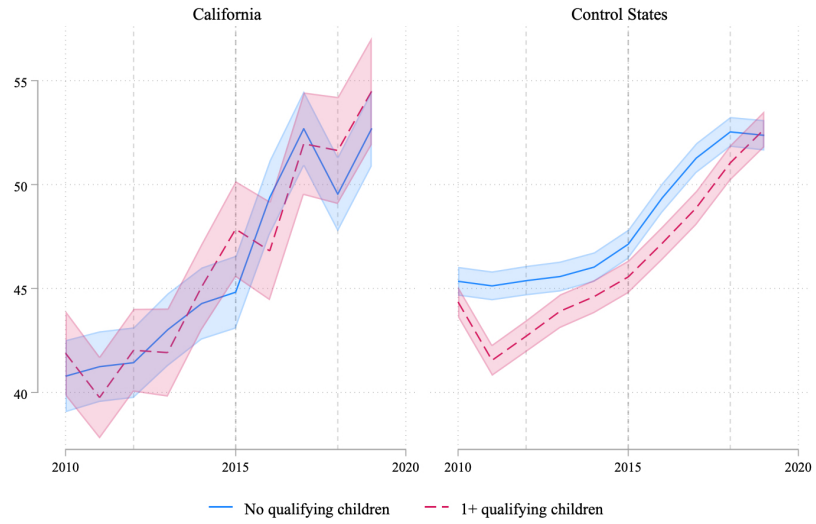
Note: The parameters of the federal EITC and CTC used to create this figure were obtained via the Tax Policy Center, the parameters of the California EITC schedule were backed-out of the Earned Income Tax Credit tables within the California Form FTB 2514, and the Young Child tax Credit Schedule was generated via information from the Franchise Tax Board. The benefit schedules shown assume a single filer whose adjusted gross income is solely composed of wage income, as self-employment income was not made a part of the definition of earned income for the purpose of the CalEITC until 2017.

Figure A.6: Triple-difference estimate of the Effect of the CalEITC on the After-tax rate

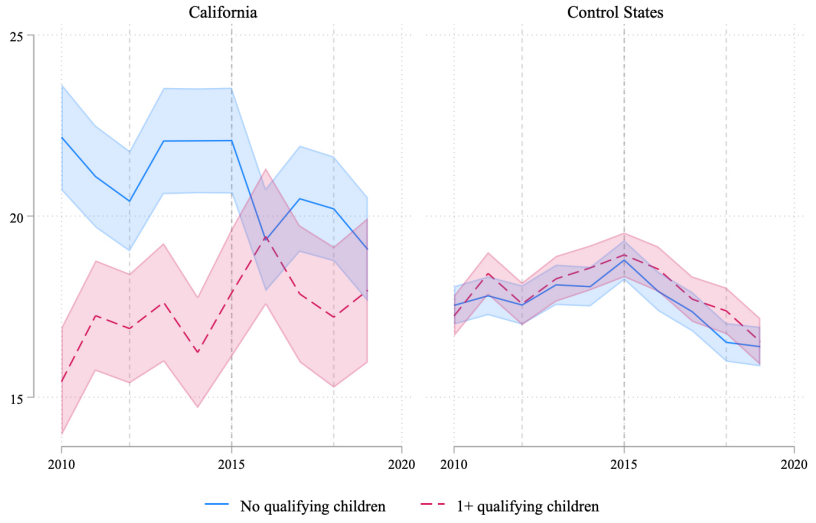


Note: This figure reports the results of Equation (2.3), estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, without a college degree, and not potential qualifying children as define in Appendix A.2.2, via the American Community Survey (ACS) from 2010-2017. The outcome variable is the average tax rate for a given individual, calculated using an individual's characteristics but assuming their earnings maximized the size of their CalEITC receipt were they to live in California. For 2010 through 2014, the 2015 CalEITC parameters are deflated. The different models shown reflect different choices of the pool of control states. Each model excludes both demographic and unemployment controls. Standard errors are clustered at the state-level.

Figure A.7: Annual full-time and part-time employment, via the CPS



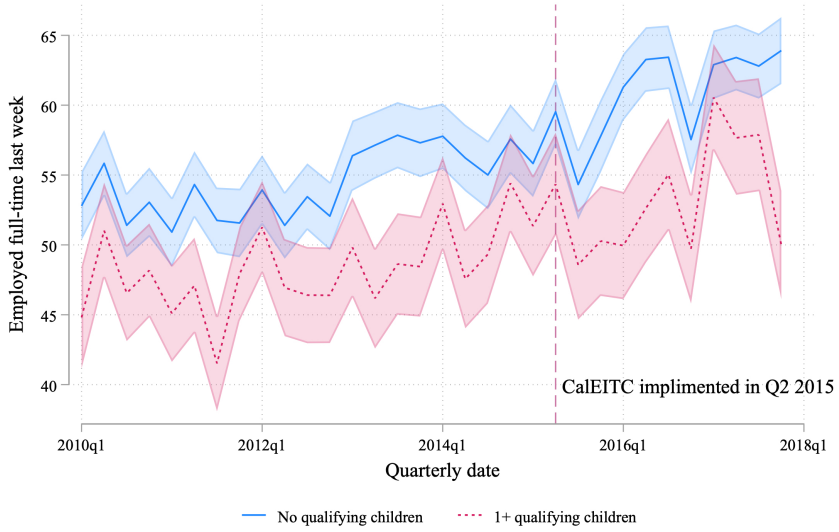
(a) Full-time employment



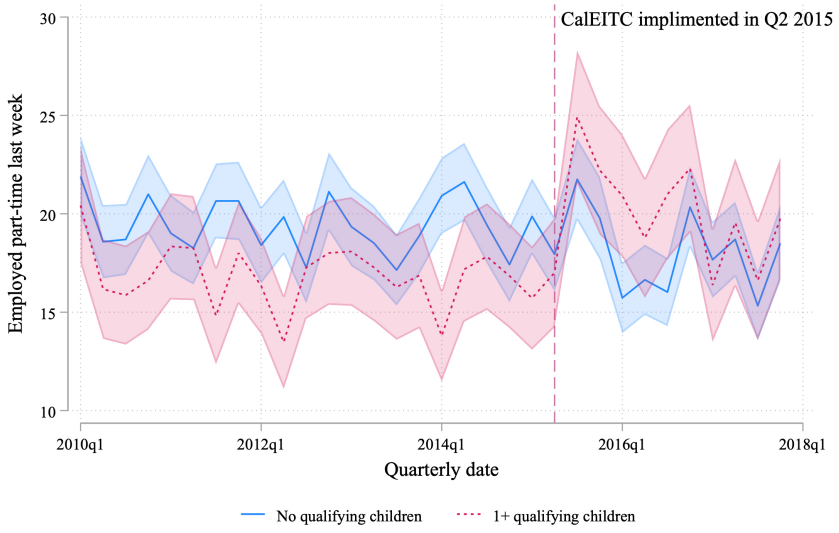
(b) Part-time employment

Note: Data for this figure are derived from the Current Population Survey (CPS), collapsed to produced annual averages by the presence of qualifying children, using individual-survey weights. The sample is limited to single women, aged 20-50, who are US citizens, not in the armed services, not in school, and not potential qualifying children as define in Appendix A.2.2. Full- and part-time employment is defined based on usual hours worked per week last year, with full-time work defined as working 35 or more hours per week, and part time as working fewer than 35 hours.

Figure A.8: Quarterly full- and part-time employment in California via the CPS



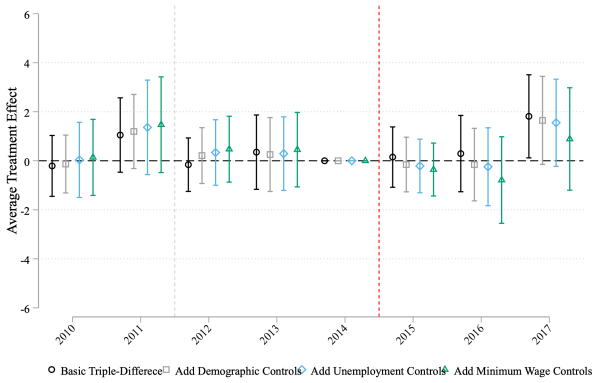
(a) Full-time employment



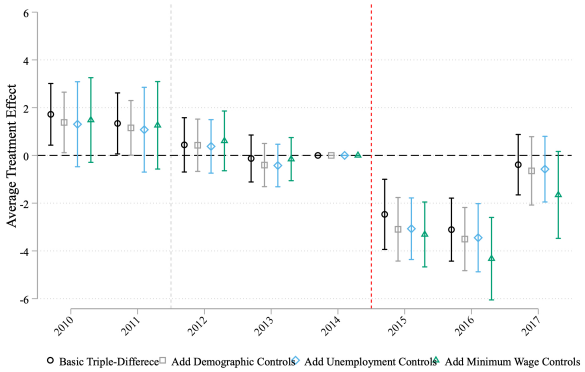
(b) Part-time employment

Note: Data for this figure are derived from the Current Population Survey (CPS), collapsed to produced quarterly averages (and 95-percent confidence intervals) by the presence of qualifying children, using individual-survey weights. The sample is limited to single women, aged 20-50, who are US citizens, not in the armed services, not in school, and not potential qualifying children as define in Appendix A.2.2. Full- and part-time employment is defined based on usual hours worked per week last year, with full-time work defined as working 35 or more hours per week, and part time as working fewer than 35 hours.

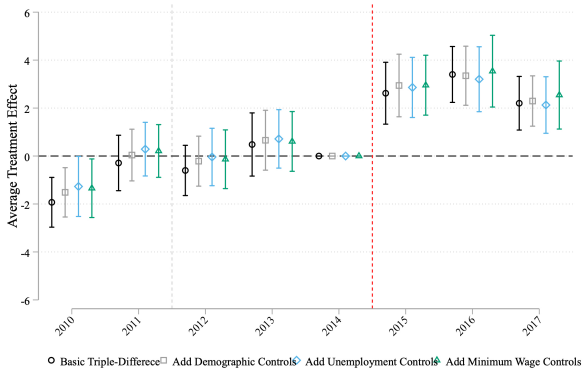
Figure A.9: Event-study estimates of the effect of the CalEITC on annual employment



(a) Employed in the last 12-months



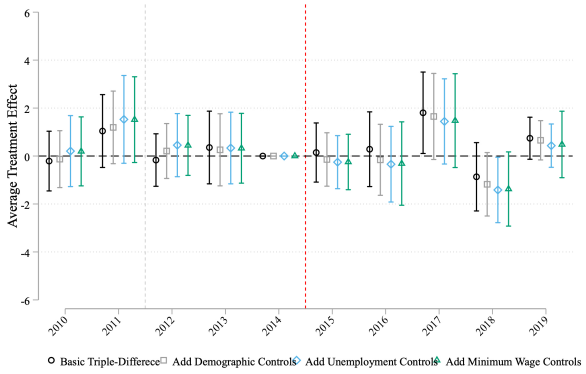
(b) Employed full-time in the last 12-months



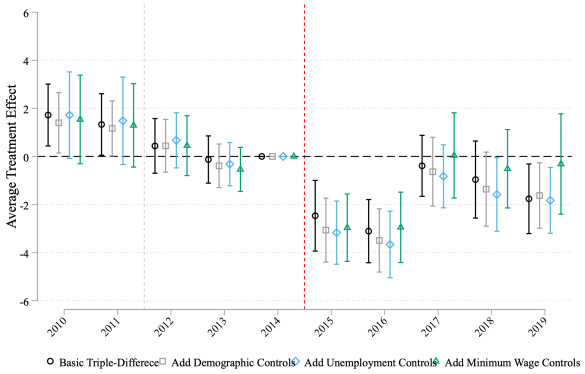
(c) Employed part-time in the last 12-months

Note: This figure reports the results of Equation (2.3), estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, not potential qualifying children as define in Appendix A.2.2, and who do not have a college degree, via the American Community Survey from 2010-2017. Each of the event-studies has a different dependent variable, defined as an indicator equal to 100 (for scaling purposes) if an individual is employed (using the relevant measure) in a given year. This means the point estimates can be interpreted as the percentage-point effect of the CalEITC on employment (however defined). Full time work is defined as working at least 35 hours a week, and part time work is defined as working under 35 hours a week. Each model is run with the full set of demographic controls, as well as controls for state unemployment and state minimum wage interacted with the number of qualifying children. Standard errors are clustered at the state-level.

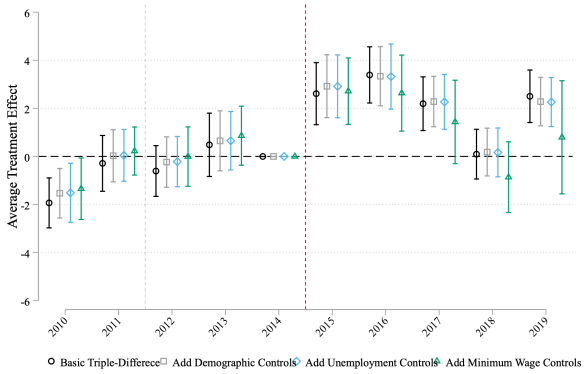
Figure A.10: Event-study estimates of the effect of the CalEITC on annual employment, 2012-2019



(a) Employed in the last 12-months



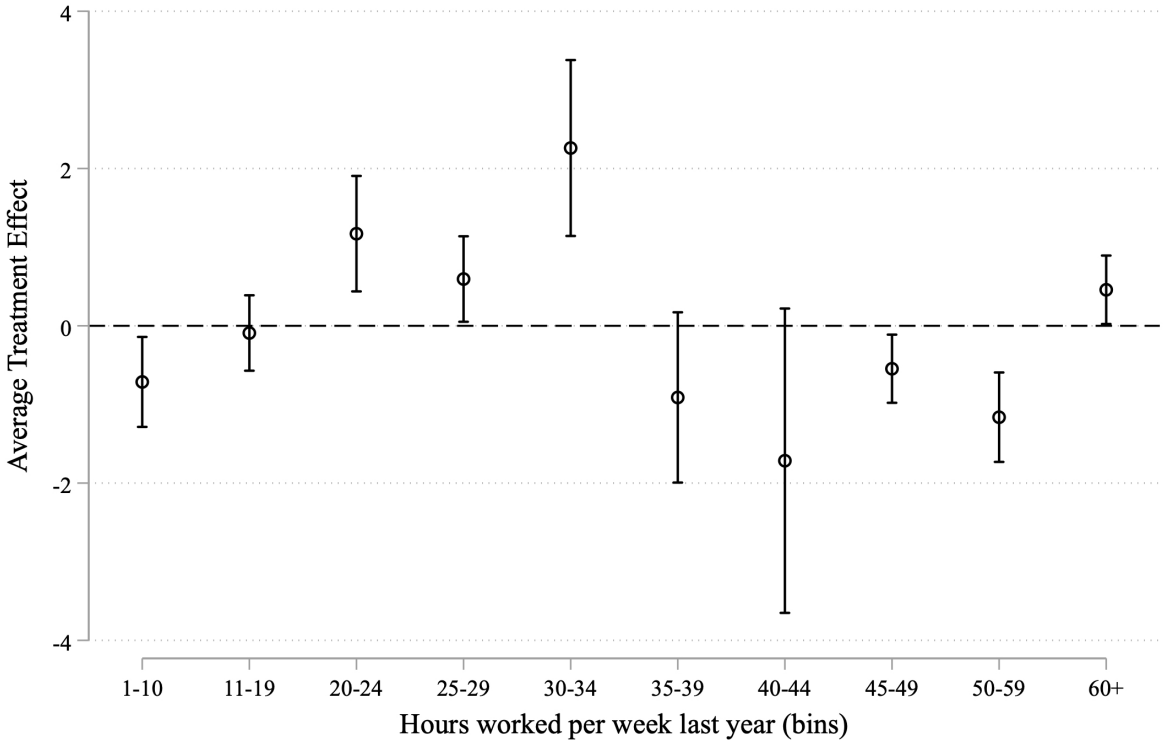
(b) Employed full-time in the last 12-months



(c) Employed part-time in the last 12-months

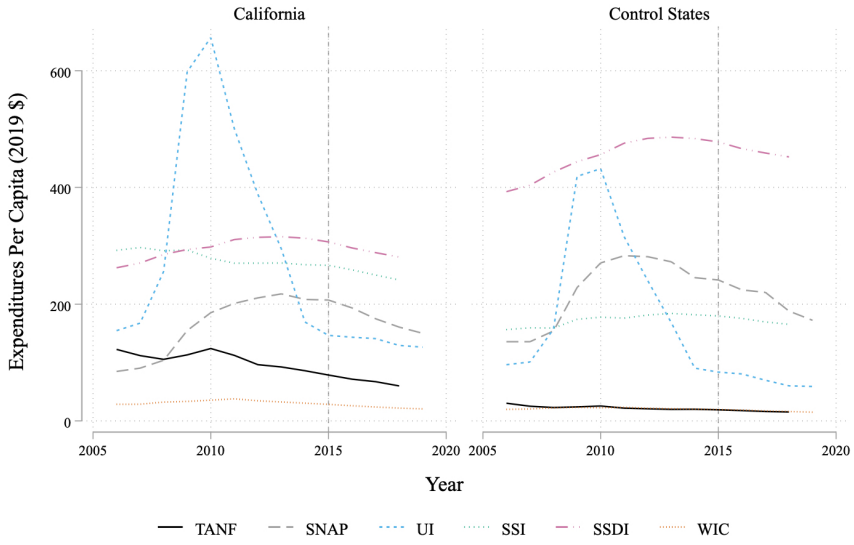
Note: This figure reports the results of Equation (2.3), estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, not potential qualifying children as define in Appendix A.2.2, and who do not have a college degree, via the American Community Survey from 2012-2019. Each of the event-studies has a different dependent variable, defined as an indicator equal to 100 (for scaling purposes) if an individual is employed (using the relevant measure) in a given year. This means the point estimates can be interpreted as the percentage-point effect of the CalEITC on employment (however defined). Full time work is defined as working at least 35 hours a week, and part time work is defined as working under 35 hours a week. Each model is run with the full set of demographic controls, as well as controls for state unemployment and state minimum wage interacted with the number of qualifying children. Standard errors are clustered at the state-level.

Figure A.11: Effect of the CalEITC on Employment, by hours worked per week

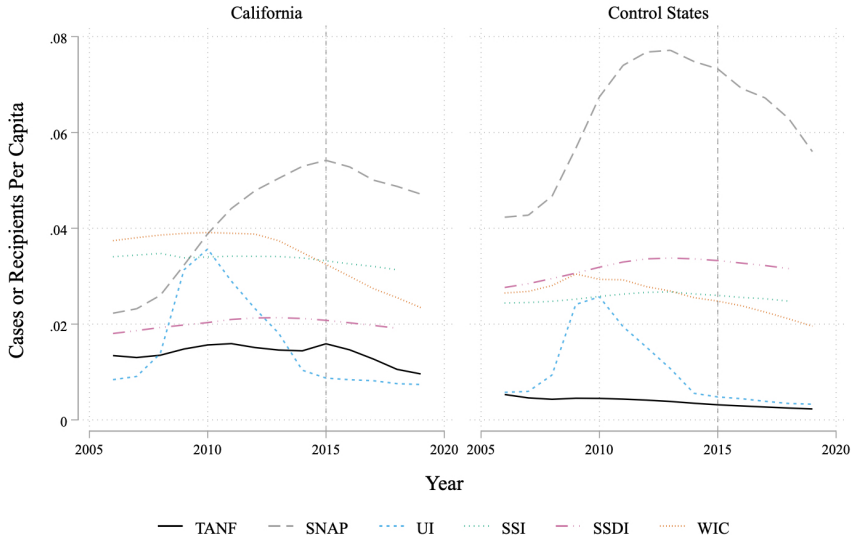


Note: This figure reports the results of Equation (2.1), estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, not potential qualifying children as define in Appendix A.2.2, and who do not have a college degree, via the American Community Survey from 2012-2017. Each of the coefficients has a different dependent variable, defined as an indicator equal to 100 (for scaling purposes) if an individual was employed in the last 12-months, with a given number usual hours worked per week. The scaling means that each coefficient can be interpreted as the percentage point change in employment within a given hours worked bin as a result of the CalEITC. Each model is run with the full set of demographic controls, as well as controls for state unemployment and state minimum wages interacted with the number of qualifying children. Standard errors are clustered at the state-level.

Figure A.12: Per capita social safety net expenditures and caseloads



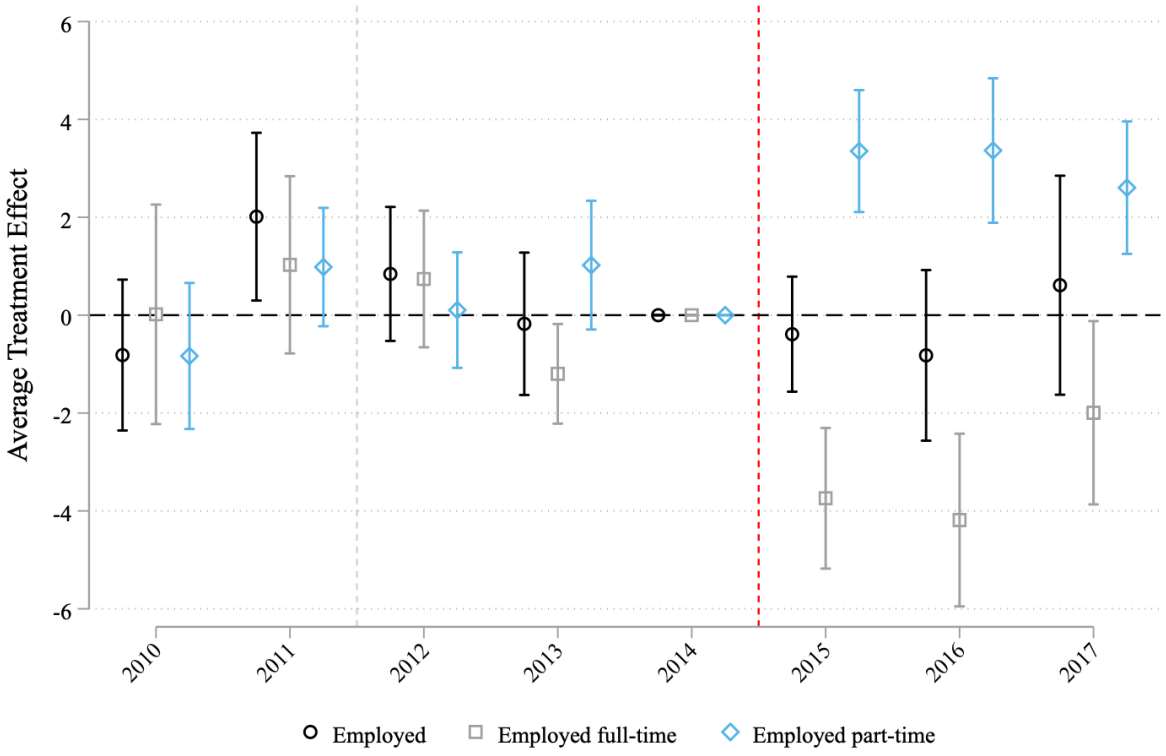
(a) Expenditures per capita



(b) Cases per capita

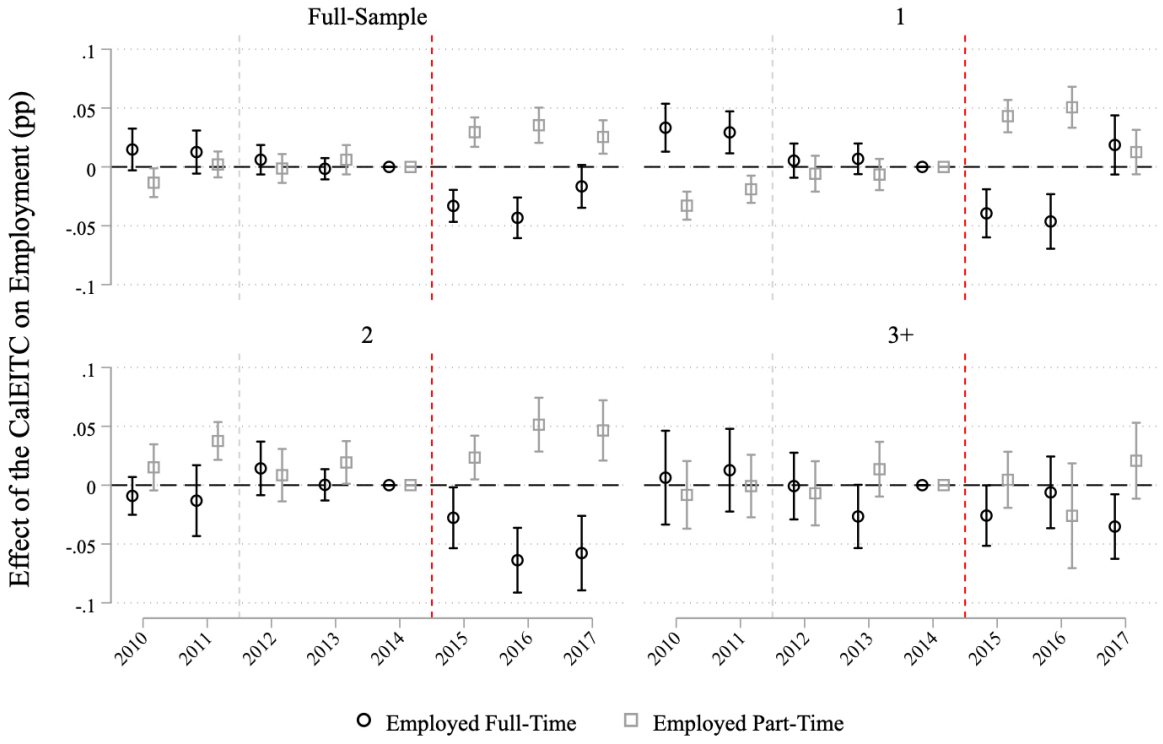
Note: This figure plots the per-capital expenditures and cases from social safety net programs in California and in the pool of untreated states, using data from Bitler et al. (2020). Programs plotted include Temporary Assistance for Needy Families (TANF), Supplemental Nutrition Assistance Program (previously known as food stamps) (SNAP), Unemployment Insurance (UI), Supplemental Security Income (SSI), Social Security Disability Insurance (SSDI), and Special Supplemental Nutrition Program for Women, Infants, and Children (WIC).

Figure A.13: Event-study estimates of the effect of the CalEITC on annual employment, using count of own children.



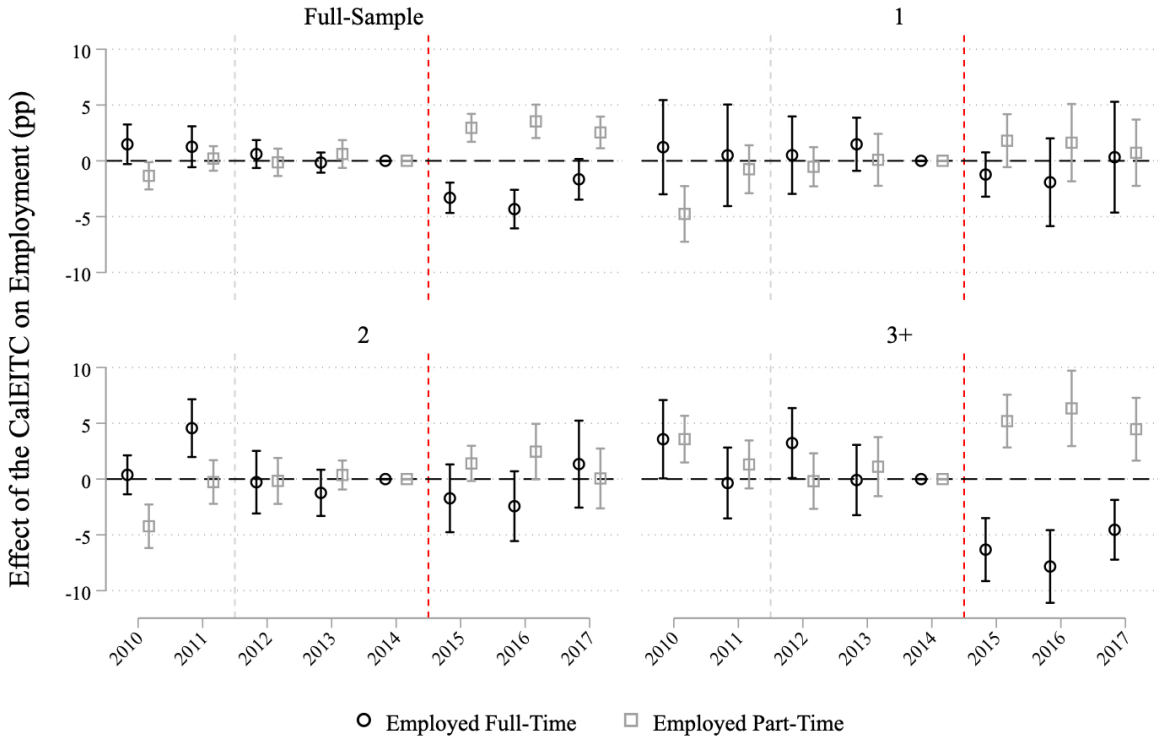
Note: This figure reports the results of Equation (2.3), estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, not potential qualifying children as define in Appendix A.2.2, and who do not have a college degree, via the American Community Survey from 2010-2017. The dashed red line denotes the split between the pre- and post-treatment periods, and the light dashed gray line represents the beginning of the pre-treatment period for my main results: 2012. The treatment variable is defined as a function of number of own-children in the household, rather than an imputation of the number of qualifying children assigned to a given individual. Each of the event-studies has a different dependent variable, defined as an indicator equal to 100 (for scaling purposes) if an individual is employed (using the relevant measure) in a given year. This means the point estimates can be interpreted as the percentage-point effect of the CalEITC on employment (however defined). Full time work is defined as working at least 35 hours a week, and part time work is defined as working under 35 hours a week. Each model is run with the full set of demographic controls, as well as controls for state unemployment and state minimum wage interacted with the number of own-children in the household. Standard errors are clustered at the state-level.

Figure A.14: Event-study estimates of the effect of the CalEITC on annual employment, by count of qualifying children



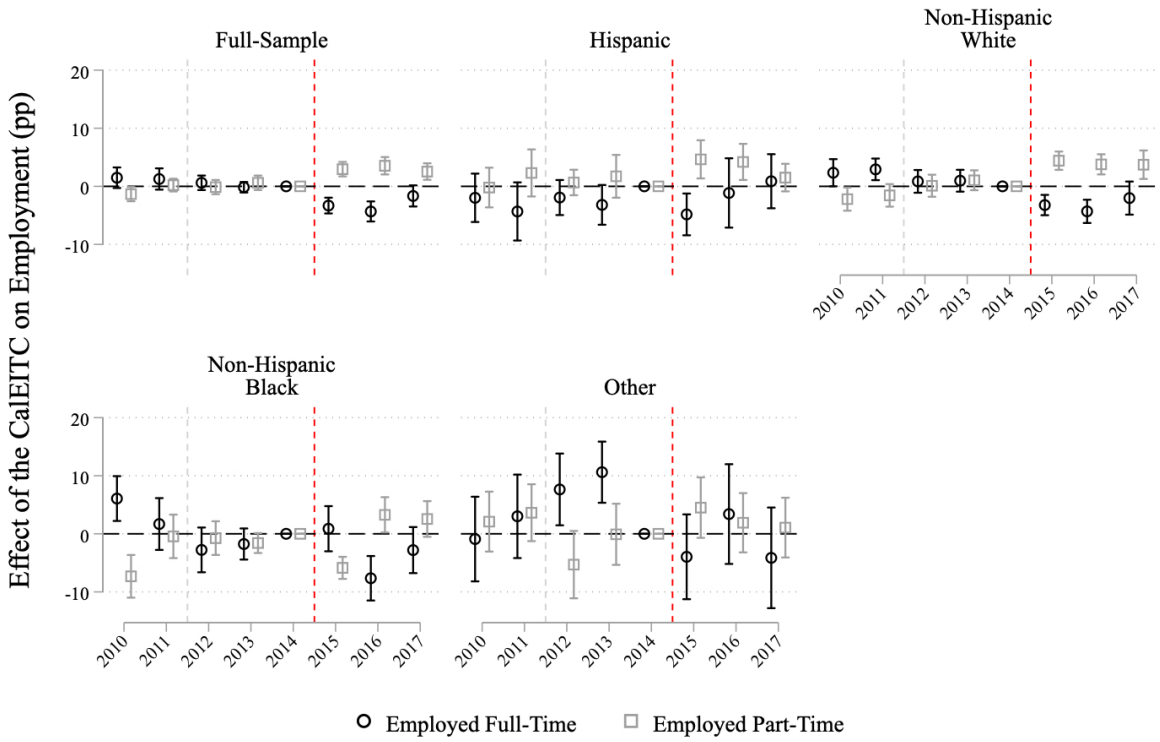
Note: This figure reports the results of Equation (2.3), estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, not potential qualifying children as define in Appendix A.2.2, and who do not have a college degree, via the American Community Survey from 2010-2017. Each panel is restricted to a different sample based on the number of qualifying children assigned to each mother in the sample (the sample of single women without kids remains uncharged). Each of the event-studies has a different dependent variable, defined as an indicator equal to 100 (for y-axis scaling) if an individual is employed (using the relevant measure) in a given year. This means the point estimates can be interpreted as the percentage-point effect of the CalEITC on employment (however defined). Full time work is defined as working at least 35 hours a week, and part time work is defined as working under 35 hours a week. Each model is run with the full set of demographic controls, as well as controls for state unemployment and state minimum wage interacted with the number of qualifying children. Standard errors are clustered at the state-level.

Figure A.15: Event-study estimates of the effect of the CalEITC on annual employment, by number of adults in the household



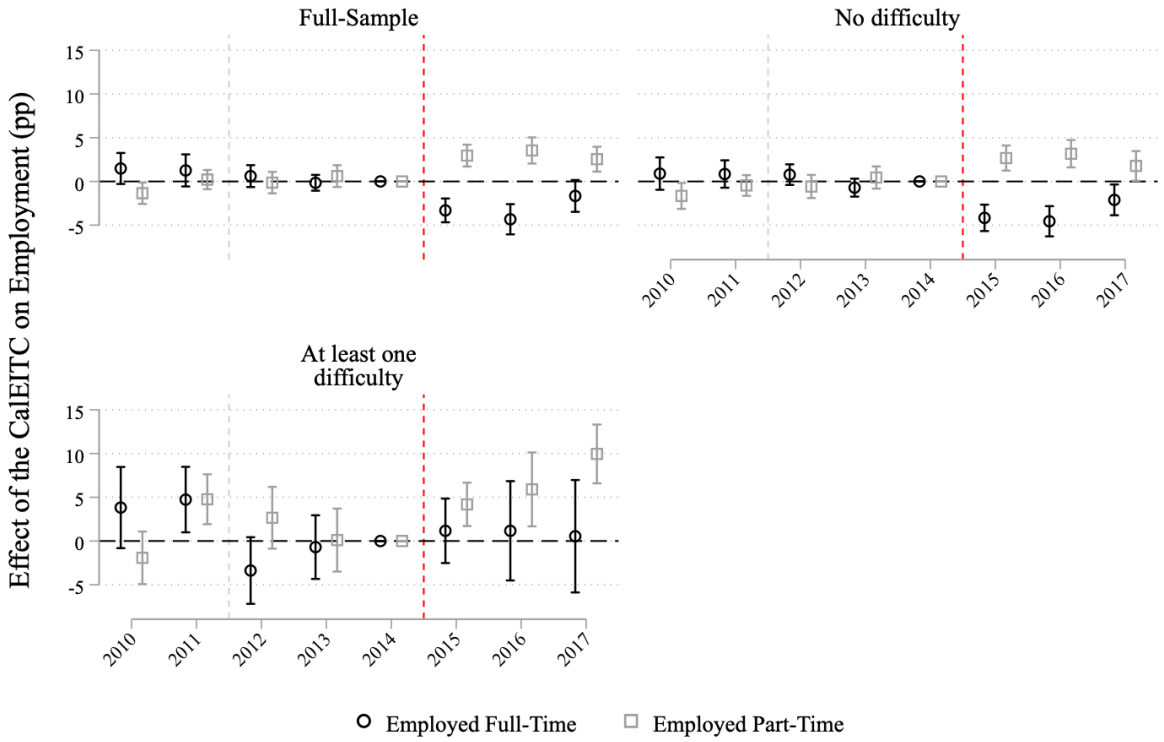
Note: This figure reports the results of Equation (2.3), estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, not potential qualifying children as define in Appendix A.2.2, and who do not have a college degree, via the American Community Survey from 2010-2017. Each panel is restricted to a different sample based on the number of adults in the household (inclusive of the single women, and exclusive of any qualifying children who are 18 years or older). Each of the event-studies has a different dependent variable, defined as an indicator equal to 100 (for y-axis scaling) if an individual is employed (using the relevant measure) in a given year. This means the point estimates can be interpreted as the percentage-point effect of the CalEITC on employment (however defined). Full time work is defined as working at least 35 hours a week, and part time work is defined as working under 35 hours a week. Each model is run with the full set of demographic controls, as well as controls for state unemployment and state minimum wage interacted with the number of qualifying children. Standard errors are clustered at the state-level.

Figure A.16: Event-study estimates of the effect of the CalEITC on annual employment, by Race and Hispanic Origin



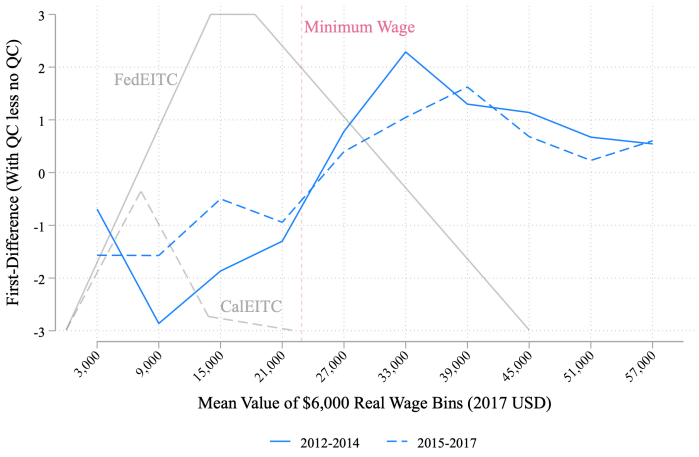
Note: This figure reports the results of Equation (2.3), estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, not potential qualifying children as define in Appendix A.2.2, and who do not have a college degree, via the American Community Survey from 2010-2017. Each panel is restricted to a different sample based on the self-reported Race or Hispanic Origin of the women in the sample. Each of the event-studies has a different dependent variable, defined as an indicator equal to 100 (for y-axis scaling) if an individual is employed (using the relevant measure) in a given year. This means the point estimates can be interpreted as the percentage-point effect of the CalEITC on employment (however defined). Full time work is defined as working at least 35 hours a week, and part time work is defined as working under 35 hours a week. Each model is run with the full set of demographic controls, as well as controls for state unemployment and state minimum wage interacted with the number of qualifying children. Standard errors are clustered at the state-level.

Figure A.17: Event-study estimates of the effect of the CalEITC on annual employment, by disability status

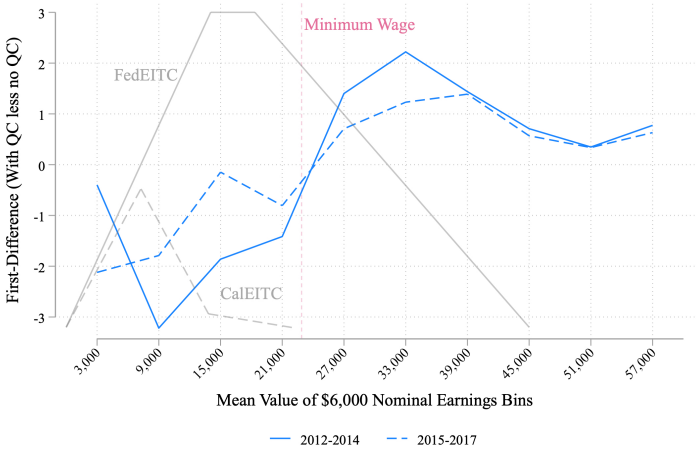


Note: This figure reports the results of Equation (2.3), estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, not potential qualifying children as define in Appendix A.2.2, and who do not have a college degree, via the American Community Survey from 2010-2017. Each panel is restricted to a different sample based on the self-reported work-related disabilities of the women in the sample. Each of the event-studies has a different dependent variable, defined as an indicator equal to 100 (for y-axis scaling) if an individual is employed (using the relevant measure) in a given year. This means the point estimates can be interpreted as the percentage-point effect of the CalEITC on employment (however defined). Full time work is defined as working at least 35 hours a week, and part time work is defined as working under 35 hours a week. Each model is run with the full set of demographic controls, as well as controls for state unemployment and state minimum wage interacted with the number of qualifying children. Standard errors are clustered at the state-level.

Figure A.18: Changes in the earnings distribution over time, alternative California distributions



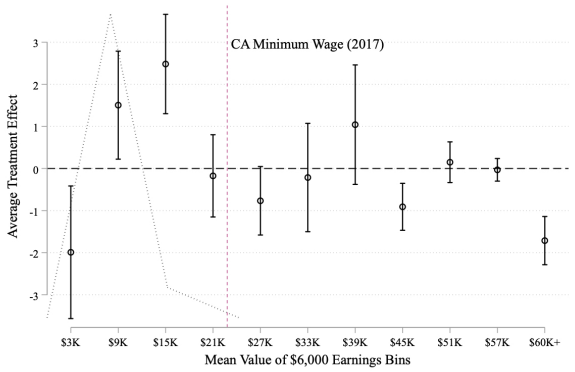
(a) Using distribution of real wages



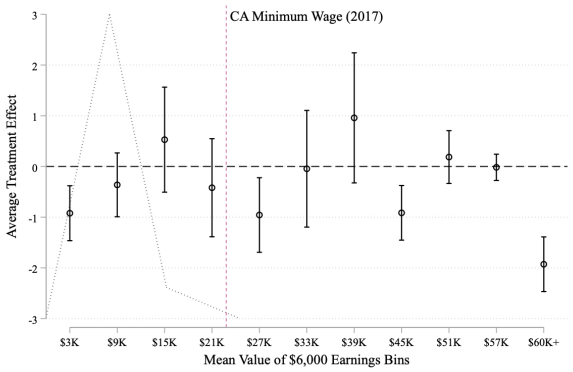
(b) Using distribution of nominal earnings

Note: These figures show the difference in the distribution of earned income (defined differently across panels) between individuals with and without qualifying children over two periods: 2012-2014 (solid line) and 2015-2017 (dashed line) in California. Figure A.18a shows the results when we plot the distribution of nominal wages, and Figure A.18b shows the results for the distribution of real earnings (2017 USD). The distribution is for the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, not potential qualifying children as define in Appendix A.2.2, and without a college degree. Both panels also plots the federal and California EITC schedule for a single parent with two qualifying children in 2017, as well as the level of earnings a full-time worker (40 hours a week, 52 weeks a year) would make at the California minimum wage in 2017.

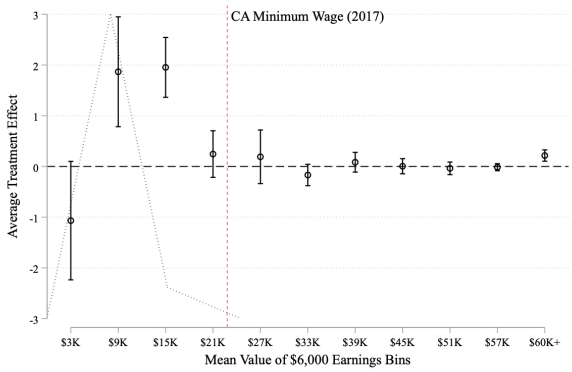
Figure A.19: Effect of the CalEITC on employment by earnings bins.



(a) Employed



(b) Full-time



(c) Part-time

Note: This figure reports the results of Equation (2.1), estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, not potential qualifying children as define in Appendix A.2.2, and who do not have a college degree, via the American Community Survey (ACS) from 2012-2017. The dependent variable is an indicator equal to 100 (for scaling of the y-axis) if the individual is employed (at all, or full- or part-time) and has earned income within a given real $6,000 earnings bin. Full time work is defined as working at least 35 hours a week, and part time work is defined as working under 35 hours a week. Each model is run with the full set of demographic, unemployment, and minimum wage controls. The dashed line represents the annual earnings of a single mother, working 52 weeks per year, 40 hours a week, at the binding minimum wage in California in 2017. Standard errors are clustered at the state-level.

Table A.1: State-level population and sample summary statistics

	Population Statistics			Samples		
	Mean	Min	Max	Main	Medicaid	No EITC
California	1,487,888	1,409,704	1,562,307	1	1	1
Alabama	263,404	249,129	276,250	1	0	1
Alaska	37,028	31,889	42,896	0	0	0
Arizona	316,689	279,694	341,074	1	1	1
Arkansas	146,827	141,210	154,019	1	1	1
Colorado	191,722	178,915	202,360	0	0	0
Connecticut	149,698	144,182	159,236	0	0	0
Delaware	47,236	44,016	50,201	1	1	0
District of Columbia	34,687	31,565	37,359	0	0	0
Florida	875,775	795,745	920,047	1	0	1
Georgia	506,628	471,773	528,786	1	0	1
Hawaii	52,291	45,565	56,013	0	0	0
Idaho	66,066	58,219	70,043	1	0	1
Illinois	553,604	539,302	563,838	0	0	0
Indiana	323,327	313,001	334,213	1	0	0
Iowa	111,295	106,037	116,586	0	0	0
Kansas	108,795	104,396	114,563	0	0	0
Kentucky	212,344	203,731	220,829	1	1	1
Louisiana	294,538	283,563	302,659	0	0	0
Maine	61,982	55,260	68,738	0	0	0
Maryland	263,486	256,773	275,934	0	0	0
Massachusetts	249,660	238,583	261,639	0	0	0
Michigan	494,819	473,307	509,259	0	0	0
Minnesota	191,881	185,389	195,677	0	0	0
Mississippi	171,947	163,830	180,228	1	0	1
Missouri	283,480	270,573	294,048	1	0	1
Montana	43,333	35,449	48,155	0	0	0
Nebraska	67,275	62,964	72,375	1	0	0
Nevada	151,376	140,105	163,838	1	1	1
New Hampshire	53,345	47,844	56,918	1	1	1
New Jersey	332,191	313,976	341,969	0	0	0
New Mexico	106,418	98,806	118,494	0	0	0
New York	754,176	734,028	775,590	1	1	0
North Carolina	434,132	412,159	449,745	1	0	1
North Dakota	26,990	22,787	32,258	1	1	1
Ohio	578,046	550,148	599,415	0	0	0
Oklahoma	183,212	164,035	193,152	1	0	0
Oregon	173,370	159,616	190,061	0	0	0
Pennsylvania	581,805	567,886	597,320	1	0	1
Rhode Island	49,113	45,722	51,474	0	0	0
South Carolina	256,611	237,632	276,166	0	0	0
South Dakota	34,632	31,559	37,298	1	0	1
Tennessee	339,937	326,004	350,350	1	0	1
Texas	1,168,805	1,088,195	1,232,414	1	0	1
Utah	101,934	89,304	117,716	1	0	1
Vermont	26,087	19,799	29,361	0	0	0
Virginia	323,266	305,794	338,719	1	0	0
Washington	285,679	279,473	297,210	1	1	1
West Virginia	95,120	89,248	104,245	1	1	1
Wisconsin	243,655	237,874	251,454	0	0	0
Wyoming	22,715	19,942	26,519	1	0	1

Note: This table reports summary statistics regarding the individual weighted state-level annual population of single women, aged 20-50, who are US citizens, not in the armed services, not in school, not potential qualifying children as define in Appendix A.2.2, and who do not have a college degree via the American Community Survey (ACS) from 2010-2017. The table also details some information about the relevant policy characteristics of each state: whether they experienced no change in their state EITC policy between 2008-2019 (my main sample), whether the expanded Medicaid in early 2014, and whether they have no state EITC. As a note, relevant to the discussion in Appendix A.5, California is the largest state by approximately 300,000 women, relative the second-largest (Texas), and by more than 500,000 women relative to any other state.

Table A.2: Triple-difference balance test

	(1) Hispanic	(2) NH White	(3) NH Black	(4) Other race	(5) No HS	(6) HS Only	(7) Some College	(8) 3+ adult HH	(9) Age
Treatment	-0.022***	0.023***	-0.006*	0.005***	-0.015***	0.009**	0.005	-0.005	-0.046
	(0.003)	(0.004)	(0.003)	(0.002)	(0.002)	(0.004)	(0.003)	(0.004)	(0.060)
Unadjusted p-values	0.000	0.000	0.065	0.003	0.000	0.016	0.128	0.196	0.452
Romano Wolf p-values	0.020	0.030	0.683	0.327	0.020	0.475	0.723	0.723	0.723
BKY (2006) sharpened q-values	0.001	0.001	0.046	0.005	0.001	0.017	0.079	0.108	0.178
Observations	588,737	588,737	588,737	588,737	588,737	588,737	588,737	588,737	588,737
R-Squared	0.168	0.110	0.113	0.036	0.011	0.008	0.009	0.075	0.011

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: This table is derived from the American Community Survey (ACS) from 2010-2017, limited to the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, do not have a college degree, and who are not potential qualifying children as define in Appendix A.2.2. The coefficient on treatment is the result of running 2.1 with no demographic or unemployment controls, with the dependent variable varying across columns. The point estimates imply that the share of the population that is Hispanic is falling in the treated group and the share without a high-school degree is rising, both by approximately 1 to 2 percentage points. I also report p-values obtained via the *rwolf* package in Stata, which implements Romano-Wolf multiple hypothesis correction. Standard errors are clustered at the state level * 0.10 ** 0.05 *** 0.01.

Table A.3: Triple-difference estimates, annual verses weekly measures of employment

	(1) Basic Triple Difference	(2) Add Demographic Controls	(3) Add Unemployment Controls	(4) Add Minimum Wage Controls
Panel A: Employed in the last 12-months				
ATE	0.7*	0.3	0.1	-0.7
	(0.4)	(0.4)	(0.5)	(0.5)
Observations	441,172	441,172	441,172	441,172
Adj. R-Square	0.009	0.076	0.076	0.076
Treated group mean in pre-period	71.6	71.6	71.6	71.6
Implied employment effect	11,718	5,209	1,646	-11,737
Panel B: Employed last week				
ATE	0.8	0.4	0.4	-0.2
	(0.5)	(0.5)	(0.7)	(0.6)
Observations	441,172	441,172	441,172	441,172
Adj. R-Square	0.009	0.074	0.074	0.074
Treated group mean in pre-period	64.0	64.0	64.0	64.0
Implied employment effect	13,575	6,323	6,443	-3,407
Triple-Difference	Yes	Yes	Yes	Yes
Add Demographic Controls	No	Yes	Yes	Yes
Add Unemployment Controls	No	No	Yes	Yes
Add Minimum Wage Controls	No	No	No	Yes

Note: This table reports the results of Equation 2.1, estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, not potential qualifying children as define in Appendix A.2.2, and who do not have a college degree, via the American Community Survey (ACS) from 2012-2017. Panel's A and B vary only in the dependent variables, which are indicators for any employment any point in the last 12-months and employment last week, respectively. In columns 2, 3, and 4, demographic controls include indicators for 5-year age bins, race and Hispanic origin groups, educational attainment, bins for the age of the individual's youngest qualifying child, and the number of adults in a individual's household. Standard errors clustered at the state level. * 0.10 ** 0.05 *** 0.01

Table A.4: Triple-difference estimates of the effect of the CalEITC on annual employment, 2012-2019

	(1) Basic Triple Difference	(2) Add Demographic Controls	(3) Add Unemployment Controls	(4) Add Minimum Wage Controls
Panel A: Employed in the last 12-months				
ATE	0.4	0.0	-0.3	-0.4
	(0.3)	(0.3)	(0.4)	(0.5)
Observations	585,888	585,888	585,888	585,888
Adj. R-Square	0.009	0.075	0.075	0.075
Treated group mean in pre-period	71.6	71.6	71.6	71.6
Implied employment effect	10,019	468	-9,415	-9,894
Panel B: Employed full-time in the last 12-months				
ATE	-1.8***	-2.1***	-2.9***	-2.3***
	(0.4)	(0.4)	(0.4)	(0.7)
Observations	585,888	585,888	585,888	585,888
Adj. R-Square	0.008	0.066	0.066	0.066
Treated group mean in pre-period	51.0	51.0	51.0	51.0
Implied employment effect	-51,624	-57,588	-81,278	-63,226
Panel C: Employed part-time in the last 12-months				
ATE	2.2***	2.1***	2.6***	1.9**
	(0.3)	(0.3)	(0.4)	(0.8)
Observations	585,888	585,888	585,888	585,888
Adj. R-Square	0.003	0.027	0.027	0.027
Treated group mean in pre-period	20.6	20.6	20.6	20.6
Implied employment effect	61,643	58,055	71,862	53,332
Triple-Difference	Yes	Yes	Yes	Yes
Add Demographic Controls	Yes	Yes	Yes	Yes
Add Unemployment Controls	Yes	Yes	Yes	Yes
Add Minimum Wage Controls	Yes	Yes	Yes	Yes

Note: This table reports the results of Equation (2.1), estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, not potential qualifying children as define in Appendix A.2.2, and without a college degree, via the American Community Survey (ACS) from 2012-2019. Panel's A, B, and C vary only in the dependent variables, which are indicators for any employment, full-time employment, and part-time employment at any point in the last 12-months, respectively. Full time work is defined as working at least 35 hours a week, and part time work is defined as working under 35 hours a week. In columns 2, 3, and 4, demographic controls include indicators for 5-year age bins, race and Hispanic origin groups, educational attainment, bins for the age of the individual's youngest qualifying child, and the number of adults in a individual's household. Standard errors clustered at the state level. * 0.10 ** 0.05 *** 0.01

Table A.5: Synthetic Difference-in-Difference Estimates, Weighted

	(1)	(2)	(3)	(4)
	Cross-State		Triple Difference	
	No Controls	With Controls	No Controls	With Controls
Panel A: Employed in the last 12-months				
ATT	1.33	-1.38	0.34	-1.17
SE	0.37	0.38	0.59	0.60
Panel B: Employed full-time in the last 12-months				
ATT	-1.34	-3.89	-2.60	-3.66
SE	0.62	0.59	0.73	1.01
Panel C: Employed part-time in the last 12-months				
ATT	2.94	2.58	3.55	2.87
SE	0.42	0.48	0.73	0.79

Note: This table reports the effect of CalEITC on employment, estimated using the Synthetic Difference-in-Differences approach created by (Arkhangelsky et al., 2021), implemented via the *sdid* command in Stata (Clarke et al., 2023). The model is run over a county-year panel, constructed as the weighted average of the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, not potential qualifying children as define in Appendix A.2.2, and without a college degree, via the American Community Survey (ACS) from 2010-2017. All individuals in a given state who are not assigned to a specific county are group in a state-specific "other" county. Panel's A, B, and C vary only in the dependent variables, which are indicators for any employment, full-time employment, and part-time employment at any point in the last 12-months, respectively. Full time work is defined as working at least 35 hours a week, and part time work is defined as working under 35 hours a week. Columns 1 and 2 use as the dependent variable mean share of single women with qualifying children the who are employed (however measured) among, while Columns 3 and 4 use as the dependent variable the difference within a given county in the employment rate between single women with and without qualifying children. Demographic and economic controls, included in columns 2 and 4 include controls for county-level race, Hispanic origin, age, and education characteristics, as well as state unemployment rates. All columns report the weighted average of the treatment effects estimated on each of the counties individually. Bootstrap standard errors are reported below ATT.

Table A.6: Triple-difference estimates of the effect of the CalEITC on annual employment, simulated instrument approach

	(1) Basic Triple Difference	(2) Add Demographic Controls	(3) Add Unemployment Controls	(4) Add Minimum Wage Controls
Panel A: Employed in the last 12-months				
Simulated CalEITC (1,000s USD)	3.4	0.5	2.6	-6.3
	(3.0)	(3.0)	(4.0)	(3.8)
Observations	441,172	441,172	441,172	441,172
Adj. R-Square	0.009	0.076	0.076	0.076
Treated group mean in pre-period	71.6	71.6	71.6	71.6
Implied employment effect	58,138	7,807	44,886	-106,991
Panel B: Employed full-time in the last 12-months				
Simulated CalEITC (1,000s USD)	-14.3***	-15.9***	-13.8***	-22.5***
	(2.9)	(3.1)	(4.9)	(5.6)
Observations	441,172	441,172	441,172	441,172
Adj. R-Square	0.008	0.068	0.068	0.068
Treated group mean in pre-period	51.0	51.0	51.0	51.0
Implied employment effect	-243,042	-270,638	-234,877	-383,150
Panel C: Employed part-time in the last 12-months				
Simulated CalEITC (1,000s USD)	17.7***	16.3***	16.4***	16.2***
	(1.8)	(1.8)	(3.5)	(4.6)
Observations	441,172	441,172	441,172	441,172
Adj. R-Square	0.003	0.028	0.028	0.028
Treated group mean in pre-period	20.6	20.6	20.6	20.6
Implied employment effect	301,180	278,444	279,763	276,159
Triple-Difference	Yes	Yes	Yes	Yes
Add Demographic Controls	No	Yes	Yes	Yes
Add Unemployment Controls	No	No	Yes	Yes
Add Minimum Wage Controls	No	No	No	Yes

Note: This table reports the results of Equation (2.1), where treatment is defined as the average CalEITC amount for a given year, state, and number of qualifying children, estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, do not have a college degree, and not potential qualifying children as define in Appendix A.2.2, via the American Community Survey (ACS) from 2012-2017. Panel's A, B, and C vary only in the dependent variables, which are indicators for any employment, full-time employment, and part-time employment at any point in the last 12-months, respectively. Full time work is defined as working at least 35 hours a week, and part time work is defined as working under 35 hours a week. In columns 2, 3, and 4, demographic controls include indicators for 5-year age bins, race and Hispanic origin groups, educational attainment, bins for the age of the individual's youngest qualifying child, and the number of adults in a individual's household. Standard errors clustered at the state level. * 0.10 ** 0.05 *** 0.01

Table A.7: Placebo test using Longitudinal CPS, conditioning on prior-year employment

	(1)	(2)	(3)	(4)	(5)
	Full Sample	Longitudinal Sample			
	All	All	Full-time last year	Part-time last year	Not employed last year
Panel A: Employed last week					
1+ QC X post-2015	0.8	-0.1	1.1	-2.5	-3.4
	(1.4)	(2.3)	(1.9)	(3.1)	(2.6)
Observations	188,038	53,912	24,006	10,200	19,706
Adj. R-Square	0.068	0.066	0.019	0.027	0.061
Treated group mean in pre-period	64.6	64.1	90.5	83.3	25.1
Panel B: Employed full-time last week					
1+ QC X post-2015	0.9	0.1	0.7	-0.4	-2.0
	(1.3)	(2.1)	(3.1)	(4.3)	(2.1)
Observations	188,038	53,912	24,006	10,200	19,706
Adj. R-Square	0.059	0.059	0.026	0.019	0.028
Treated group mean in pre-period	46.5	45.7	77.3	41.3	12.9
Panel C: Employed part-time last week					
1+ QC X post-2015	-0.2	-0.2	0.4	-2.1	-1.4
	(0.7)	(1.7)	(2.1)	(3.3)	(1.8)
Observations	188,038	53,912	24,006	10,200	19,706
Adj. R-Square	0.014	0.019	0.012	0.016	0.032
Treated group mean in pre-period	18.1	18.4	13.2	42.0	12.2
Basic Triple-Difference	Yes	Yes	Yes	Yes	Yes
Add Demographic Controls	Yes	Yes	Yes	Yes	Yes
Add Unemployment Controls	Yes	Yes	Yes	Yes	Yes
Add Minimum Wage Controls	Yes	Yes	Yes	Yes	Yes

Note: This table reports the results of a placebo test, estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, not potential qualifying children as define in Appendix A.2.2, without a college degree, not in California, via the longitudinal Monthly Current Population Survey (CPS) sample, from 2012-2017, omitting the first two quarters of 2015. The placebo test is a simple difference-in-difference, comparing employment for single women with and without qualifying children before and after 2015. Panel's A, B, and C vary only in the dependent variables, which are indicators for any employment, full-time employment, and part-time employment last week, respectively. Full time work is defined as working at least 35 hours a week, and part time work is defined as working under 35 hours a week. Column 1 shows the results of the model run over the full sample, while Columns 2 through 5 report the results of the full longitudinal sample, with columns 3 through 5 restricting the sample further conditional on an individual's employment status one year ago. Standard errors clustered at the state and individual level for the full-sample and at the state-level for the longitudinal sample. * 0.10 ** 0.05 *** 0.01

Table A.8: Triple-difference estimates of the effect of the CalEITC on weekly employment via the CPS, 2012 through 2017, including the first two quarters of 2015

	(1)	(2)	(3)	(4)	(5)
	Full Sample	Longitudinal Sample			
	All	All	Full-time last year	Part-time last year	Not employed last year
Panel A: Employed last week					
ATE	3.9*	4.0	-3.2	5.0	3.1
	(1.9)	(3.7)	(2.2)	(5.5)	(3.2)
Observations	232,382	66,687	29,460	12,760	24,467
Adj. R-Square	0.069	0.069	0.025	0.037	0.061
Treated group mean in pre-period	59.9	59.2	59.2	59.2	59.2
Panel B: Employed full-time last week					
ATE	2.0	2.7	-6.0*	4.5	3.4
	(2.3)	(4.2)	(3.2)	(4.2)	(3.5)
Observations	232,382	66,687	29,460	12,760	24,467
Adj. R-Square	0.059	0.061	0.030	0.020	0.030
Treated group mean in pre-period	43.0	42.9	42.9	42.9	42.9
Panel C: Employed part-time last week					
ATE	1.9**	1.3	2.8	0.5	-0.3
	(0.9)	(1.6)	(1.8)	(5.7)	(2.4)
Observations	232,382	66,687	29,460	12,760	24,467
Adj. R-Square	0.016	0.021	0.014	0.019	0.031
Treated group mean in pre-period	16.9	16.3	16.3	16.3	16.3
Basic Triple-Difference	Yes	Yes	Yes	Yes	Yes
Add Demographic Controls	Yes	Yes	Yes	Yes	Yes
Add Unemployment Controls	Yes	Yes	Yes	Yes	Yes

Note: This table reports the results of Equation (2.1), estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, who do not have a college degree, and who are not potential qualifying children as define in Appendix A.2.2, via the monthly Current Population Survey (CPS) from 2012-2017, including the first two quarters of 2015. Panel's A, B, and C vary only in the dependent variables, which are indicators for any employment, full-time employment, and part-time employment last week, respectively. Full time work is defined as working at least 35 hours a week, and part time work is defined as working under 35 hours a week. Column 1 shows the results of the model run over the full sample, while Columns 2 through 5 report the results of the full longitudinal sample, with columns 3 through 5 restricting the sample further conditional on an individual's employment status one year ago. All specifications include demographic controls, including indicators for 5-year age bins, race and Hispanic origin groups, educational attainment, bins for the age of the individual's youngest qualifying child, and the number of adults in a individual's household, as well as controls for state unemployment rates interacted with the number of qualifying children. Standard errors clustered at the state and individual level for the full-sample and at the state-level for the longitudinal sample. * 0.10 ** 0.05 *** 0.01

Table A.9: Triple-difference estimates of the effect of the CalEITC on annual employment, by race and Hispanic origin

	(1) Full sample	(2) Hispanic	(3) Non-Hispanic White	(4) Non-Hispanic Black	(5) Non-Hispanic Other
Panel A: Employed in the last 12-months					
ATE	-0.7	1.1	-0.7	-1.2	-2.9
	(0.5)	(1.5)	(0.8)	(1.4)	(2.3)
Observations	441,172	85,403	234,569	91,239	29,950
Adj. R-Square	0.076	0.064	0.090	0.079	0.077
Treated group mean in pre-period	71.6	72.7	72.3	63.9	74.9
Implied employment effect	-11,737	9,770	-2,922	-2,669	-4,684
Panel B: Employed full-time in the last 12-months					
ATE	-3.9***	-3.1**	-3.7***	-2.0	-6.5
	(0.6)	(1.3)	(1.1)	(1.7)	(4.3)
Observations	441,172	85,403	234,569	91,239	29,950
Adj. R-Square	0.068	0.062	0.077	0.075	0.066
Treated group mean in pre-period	51.0	53.4	49.9	42.6	54.2
Implied employment effect	-66,628	-27,145	-15,920	-4,637	-10,590
Panel C: Employed part-time in the last 12-months					
ATE	3.2***	4.2***	3.0***	0.9	3.6
	(0.6)	(0.8)	(0.9)	(2.0)	(2.8)
Observations	441,172	85,403	234,569	91,239	29,950
Adj. R-Square	0.028	0.024	0.034	0.029	0.030
Treated group mean in pre-period	20.6	19.3	22.5	21.3	20.7
Implied employment effect	54,890	36,915	12,997	1,967	5,906
Basic Triple-Difference	Yes	Yes	Yes	Yes	Yes
Add Demographic Controls	Yes	Yes	Yes	Yes	Yes
Add Unemployment Controls	Yes	Yes	Yes	Yes	Yes
Add Minimum Wage Controls	Yes	Yes	Yes	Yes	Yes

Note: This table reports the results of Equation (2.1), estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, who do not have a college degree, and who are not potential qualifying children as define in Appendix A.2.2, via the American Community Survey (ACS) from 2012-2017. Panels A, B, and C vary only in the dependent variables, which are indicators for any employment, full-time employment, and part-time employment at any point in the last 12-months, respectively. Column 1 reports the results from the full sample, while columns 2 through 5 restrict the sample by the reported Race and Hispanic Origin of the women. All models include demographic, state unemployment, and minimum wage controls. Standard errors clustered at the state level. * 0.10 ** 0.05 *** 0.01

Table A.10: Triple-difference estimates of the effect of the CalEITC on annual employment, by disability status

	(1) Full sample	(2) No reported disability	(3) One or more reported disability
Panel A: Employed in the last 12-months			
ATE	-0.7	-1.0	4.5**
	(0.5)	(0.7)	(2.0)
Observations	441,172	387,259	53,913
Adj. R-Square	0.076	0.065	0.068
Treated group mean in pre-period	71.6	74.3	35.9
Implied employment effect	-11,737	-15,785	5,239
Panel B: Employed full-time in the last 12-months			
ATE	-3.9***	-4.0***	-0.6
	(0.6)	(0.8)	(2.1)
Observations	441,172	387,259	53,913
Adj. R-Square	0.068	0.071	0.058
Treated group mean in pre-period	51.0	53.2	21.8
Implied employment effect	-66,628	-64,023	-692
Panel C: Employed part-time in the last 12-months			
ATE	3.2***	3.0***	5.0**
	(0.6)	(0.7)	(2.0)
Observations	441,172	387,259	53,913
Adj. R-Square	0.028	0.028	0.025
Treated group mean in pre-period	20.6	21.1	14.0
Implied employment effect	54,890	48,238	5,931
Basic Triple-Difference	Yes	Yes	Yes
Add Demographic Controls	Yes	Yes	Yes
Add Unemployment Controls	Yes	Yes	Yes
Add Minimum Wage Controls	Yes	Yes	Yes

Note: This table reports the results of Equation (2.1), estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, who do not have a college degree, and who are not potential qualifying children as define in Appendix A.2.2, via the American Community Survey (ACS) from 2012-2017. Panels A, B, and C vary only in the dependent variables, which are indicators for any employment, full-time employment, and part-time employment at any point in the last 12-months, respectively. Column 1 reports the results from the full sample, while columns 2 and 3 restrict the sample by the number of reported disabilities in the ACS. All models include demographic, state unemployment, and minimum wage controls. Standard errors clustered at the state level. * 0.10 ** 0.05 *** 0.01

Table A.11: Triple-difference estimates, effect of the CalEITC on annual earnings

	(1) Basic Triple Difference	(2) Add Demographic Controls	(3) Add Unemployment Controls	(4) Add Minimum Wage Controls
Panel A: OLS - No restrictions				
ATE	-419.0***	-688.3***	-954.3***	-1332.8***
	(150.9)	(173.7)	(293.8)	(361.7)
Observations	323,550	323,550	323,550	323,550
Adj. R-Square	0.013	0.125	0.125	0.125
Treated group mean in pre-period	19,915	19,915	19,915	19,915
Panel B: OLS - positive income only				
ATE	-1127.5***	-1197.1***	-1614.2***	-1961.2***
	(217.5)	(232.8)	(341.7)	(520.7)
Observations	225,459	225,459	225,459	225,459
Adj. R-Square	0.014	0.137	0.137	0.137
Treated group mean in pre-period	30,117	30,117	30,117	30,117
Panel C: Poisson				
ATE	-0.03***	-0.04***	-0.06***	-0.08***
	(0.01)	(0.01)	(0.02)	(0.02)
Observations	323,550	323,550	323,550	323,550
Pseudo-R-squared	0.018	0.167	0.167	0.167
Treated group mean in pre-period	19,915	19,915	19,915	19,915
Effect in USD	-503	-775	-1,030	-1,353
Triple-Difference	Yes	Yes	Yes	Yes
Add Demographic Controls	No	Yes	Yes	Yes
Add Unemployment Controls	No	No	Yes	Yes
Add Minimum Wage Controls	No	No	No	Yes

Note: This table reports the results of 2.1, estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, not potential qualifying children as define in Appendix A.2.2, and who do not have a college degree via the American Community Survey (ACS) from 2010-2017. The dependent variable is annual earnings, defined as wages plus self-employment income, in real 2019 dollars. I drop all observations whose earned income is recorded as "allocated", meaning the earnings value is imputed via a hot-deck allocation method. Panels A estimates an OLS model, and includes individuals with no earned income. Panel B estimates the same model excluding non-workers. Panel C estimates a Poisson model via the Stata command *ppmlhdfe*, including those with no earned income. In columns 2, 3, and 4, demographic controls include indicators for 5-year age bins, race and Hispanic origin groups, educational attainment, bins for the age of the individual's youngest qualifying child, and the number of adults in a individual's household. Standard errors clustered at the state level. * 0.10 ** 0.05 *** 0.01

Table A.12: Hours spent on different activities, by employment status, based on data from Bastian and Lochner (2022)

	(1) All	(2) Non-workers	(3) Part-time workers	(4) Full-time workers
Work hours (CPS)	18.93	4.237	21.10	33.67
	(0.376)	(0.360)	(0.659)	(0.468)
Home production	42.58	50.23	40.67	35.31
	(0.464)	(0.801)	(0.954)	(0.592)
With children	20.57	27.09	18.14	14.80
	(0.419)	(0.755)	(0.818)	(0.517)
Without children	22.01	23.14	22.53	20.51
	(0.345)	(0.597)	(0.825)	(0.457)
Leisure	34.58	43.88	30.29	26.77
	(0.479)	(0.787)	(0.967)	(0.644)
With children	17.41	23.55	13.80	12.66
	(0.405)	(0.707)	(0.760)	(0.556)
Without children	17.17	20.33	16.50	14.12
	(0.423)	(0.783)	(0.902)	(0.518)
Total hours with children	39.23	52.00	33.51	28.44
	(0.652)	(1.076)	(1.275)	(0.885)
Investment hours with children	5.539	8.045	4.492	3.378
	(0.205)	(0.390)	(0.397)	(0.235)
N	2,528	1,014	510	1,004

Note: This table reports a series of descriptive statistics, based on American Time Use Survey (ATUS) data obtained from the replication files of Bastian and Lochner (2022). The sample is limited to single women, aged 20-to-50, without a college degree, with children, during the period of 2010 through 2015. The results, broken out by employment status at the time of the ATUS interview, show the hours per week spent on each given activity. The results are weighted using individual-level ATUS weights. Unfortunately, I am unable to directly estimate the effect of the CalEITC on parental time spent with their children on "investment" activities, due to the sample size of the ATUS. These results should not be taken as a causal estimate of the effect of a particular employment status on hours spent on a given activity; rather, they are illustrative how single mothers during this time-period spend their time across employment statuses. The results indicate that single women induced to work part-time rather than full-time by the CalEITC would spend approximately one more hour a week on investment time with their children, a marginal increases in parental investment time.

A.2 Data

Data for this project comes primarily from the American Community Survey (ACS) and the Current Population Survey (CPS), downloaded via IPUMS (Ruggles et al., 2023, Flood et al., 2022). I obtain state- and county-level unemployment data from the U.S. Bureau of Labor Statistic's Local Area Unemployment Statistics (LAUS) database. Specifically, I download the text files "la.data.3.AllStatesS" and "la.data.64.County" via https://download.bls.gov/pub/time.series/la/. Tax parameters are obtained via a combination of the Urban-Brookings Tax Policy Center Statistics, the federal Schedule EIC, and the California FTB 3514 form. I obtain data on state- and local-level minimum wage policies from Vaghul and Zipperer (2022).

A.2.1 Sample construction

I limit my sample to individuals in either California or one of the 27 states in my control pool, as detailed in Table A.1. I restrict attention to single women, aged 20-50, who are not currently in school, not in group quarters, not in the armed services, who are not potentially qualifying children themselves, who are not married but living separately from their partner, who do not have a college degree, and who are US citizens. I exclude those in the armed services or in group quarters due to concerns about my ability to link parents to children. The restrictions on students stems from the fact that students age 18 through 23 can be claimed as qualifying children by their parents (or other relation), complicating treatment assignment. I exclude those who are married but living separately from their partner because tax units who file a married filing separately return are ineligible for the EITC. I exclude single women who are not US citizens due to complexity regarding the eligibility of these workers for the EITC and potential confounding

due to the implementation of the Deferred Action for Childhood Arrivals policy during my pre-period (Pope, 2016).

A.2.2 Imputation of qualifying children

A qualifying child (QC) for the purposes of claiming either the federal EITC or the CalEITC must meet four tests: the relationship test, the age test, the residency test, and the joint return test, plus they must have a valid social security number. These tests dictate that a QC is a relative of the taxpayer who is younger than 19 (or in school and younger than 24), living with the taxpayer for more than half the year, did not file their own joint return, and who is a US citizen[1]. Other relevant restrictions are on the citizenship of the tax filers (they must be either a US citizen or a resident alien), their filing status (taxpayers who filed a married filing separately return are ineligible for the EITC), and their age (filers must be between 25 and 64 if they have no qualifying children).

In the ACS and the CPS, I can impute whether or not a given child is likely to meet many of these criteria, and which adults in their household may be able to claim the child on their tax returns. I begin with the pool of children younger than 18, and add to this the pool of adults aged 18 to 23 who are in school (the age test). I then remove from the pool any children who are married (the joint return test) and those who are not US citizens. This creates my pool of qualifying children who could be claimed by an adult. I assume that children can only be claimed by adults in their household, reflecting the residency test. I first match each potential

[1]The relationship test states that a QC must be the child, stepchild, or foster child (or any of their decedents, i.e. grandchild) of at least one of the filers. The age test states that the QC must be, as of the end of the tax year, under the age of 19 or under the age of 24 if enrolled in school. Further, they must be younger than at least one of the tax filers. The residency test states that the QC must have lived with you in the United States for more than half of the year, excluding limited circumstances (US military, births or deaths of children, etc...). The joint return test states that a QC cannot file a joint return.

qualifying child to their parents, as identified by the IPUMS variables *momloc* and *poploc*. These variables primarily reflect the stated relationships between each individual in the household and the household head. For example, in a three person household with a single married couple and their child, the *relate* variable will identify one of the adults as the head (101), another as their spouse (201) and their child as their child (301). IPUMS is able to use that information to match the child to both parents (child to head, and child to spouse through head). There are cases where the *relate* variable does not provide enough information to match children to a specific adult in their household. In these cases, children are matched first to married couples, then to single adults, conditional on a set of age restrictions on both the adults and the child, reflecting the increased chance of incorrect matches. After using information to identify matches to parents, I take the remaining pool of unmatched children and assign them, using the *relate* variable, to grandparents, foster parents, and siblings (in that order).

A.2.3 Predicted Earnings

As discussed in the prior literature, a individuals with lower earning capacities are more likely to be treated by the EITC (or by EITC eligibility) (Eissa and Liebman, 1996, Eissa and Hoynes, 2004). The previous literature has generally relied on education to proxy for earnings capacity, but several researchers have instead predicted earnings directly using a regression model (Eissa and Liebman, 1996, Kleven, 2024). I follow Kleven (2024) and predict real 2019 earned income using the following equation:

$$y_{it} = \mathrm{E}_{it} + \mathrm{A}_{it} + \mathrm{C}_{it} + \mathrm{QC}_{it} + \mathrm{R}_{it} + \mathrm{H}_{it} + \mathrm{IN}_{it} + \mathrm{S}_{it} + \delta_s + \gamma_t + \epsilon_{it} \tag{A.1}$$

Where y_{it} is the earnings of a given individual in real 2019 dollars, E_{it} is a categorical variable for educational attainment, A_{it} is a categorical variable for the individual's age, C_{it} is a categorical variable for the age of the individual's youngest child, QC_{it} is a categorical variable for the number of qualifying children an individual might claim, R_{it} and H_{it} are categorical variables for an individual's Race and Hispanic origin status, IN is an indicator for whether an individual is in school, S_{it} is an indicator for an individual's sex, and δ_s and γ_t are state and year fixed effects. The model is estimated over a sample of single individuals aged 20 to 50, from 2006 through 2019, who reported positive earned income and who reported working in the last year. After running the regression model, I predict earned income for the entire sample, including those who did not work and reported no (or negative) earned income. I drop from my low-predicted earnings sample the 94,178 observations (across 2006-2019, men and women, married and unmarried) with predicted earnings of less than $0 .

A.3 Earni

As discussed in Section 2.2, the CalEITC broke with the federal credit by excluding–at least for the first two years of the program–self-employment income from the definition of earned income. Self-employment income, or income earned by individuals who work for themselves as a sole proprietor or independent contractor, has long been associated with the EITC. Work by Saez (2010) found that bunching around EITC kink points was driven solely by individuals reporting self-employment income, following a simple model of tax evasion. More recent work by Garin et al. (2024) built on this literature, finding that increases in reporting of self-employment income driven by EITC incentives is responsible for rise of reported self-employment in IRS data.

I first verify that the main employment response I observe is not driven solely by a self-employment response. I re-estimate my main model, re-defining my outcome variables to count as employed (however defined) only workers who report some wage or salary income in the past 12 months. The results, reported in Table A.13, show that in my preferred model, the full-time effect remains unchanged, while the part-time effect falls by 0.8 percentage points: from 3.1 to 2.3. This result highlights three points. First, the majority of the part-time result is driven by workers with wage income, as anticipated given the design of the CalEITC. Second, approximately a quarter of the overall part-time increase is driven by workers who report only self-employment income. The event-study results, reported in Figure A.21, show that the largest gap between my main estimates and these which restrict to part-time work with wage income are in 2015, after which the share of the response explained by individuals without wage earnings individuals shrinks. The increase in employment among workers with only self-employment income could be driven by the ambiguous nature of wage verses self-employment income in survey data (Abraham et al., 2021, 2024a,b) or a a lack of awareness among taxpayers of the unique earnings definition of the CalEITC.

Second, I estimate the effect of the CalEITC on self-employment directly, defined as either reporting any self-employment income last year, reporting more than $1,000 in self-employment income last year, and reporting being self-employed last week. In Table A.14 I report my results, which imply that self-employment increases by between 1.0 and 1.4 percentage points in my preferred specification. This is of a similar scale to the 0.8 percentage point gap between my full part-time estimates and those estimates restricted to wage earnings. Further, when considering the event-study results in Figure A.20, I observe a jump in self-employment across all measures immediately post reform, which decreases over time.

Table A.13: Triple-difference estimates of the effect of the CalEITC on employment, conditional on reporting wage income

	(1) Basic Triple Difference	(2) Add Demographic Controls	(3) Add Unemployment Controls	(4) Add Minimum Wage Controls
Panel A: Employed in the last 12 months				
ATE	-0.0	-0.4	-1.0*	-1.8***
	(0.4)	(0.4)	(0.5)	(0.6)
Observations	441,172	441,172	441,172	441,172
Adj. R-Square	0.008	0.070	0.070	0.070
Treated group mean in pre-period	68.4	68.4	68.4	68.4
Implied employment effect	-829	-6,753	-16,218	-30,358
Panel B: Employed full-time in the last 12-months				
ATE	-2.2***	-2.5***	-3.2***	-4.1***
	(0.4)	(0.4)	(0.6)	(0.6)
Observations	441,172	441,172	441,172	441,172
Adj. R-Square	0.007	0.064	0.064	0.064
Treated group mean in pre-period	49.4	49.4	49.4	49.4
Implied employment effect	-38,272	-43,379	-54,013	-70,314
Panel C: Self-employed last week				
ATE	2.2***	2.1***	2.2***	2.3***
	(0.4)	(0.4)	(0.7)	(0.7)
Observations	441,172	441,172	441,172	441,172
Adj. R-Square	0.003	0.032	0.032	0.032
Treated group mean in pre-period	19.0	19.0	19.0	19.0
Implied employment effect	37,443	36,627	37,795	39,956
Triple-Difference	Yes	Yes	Yes	Yes
Add Demographic Controls	No	Yes	Yes	Yes
Add Unemployment Controls	No	No	Yes	Yes
Add Minimum Wage Controls	No	No	No	Yes

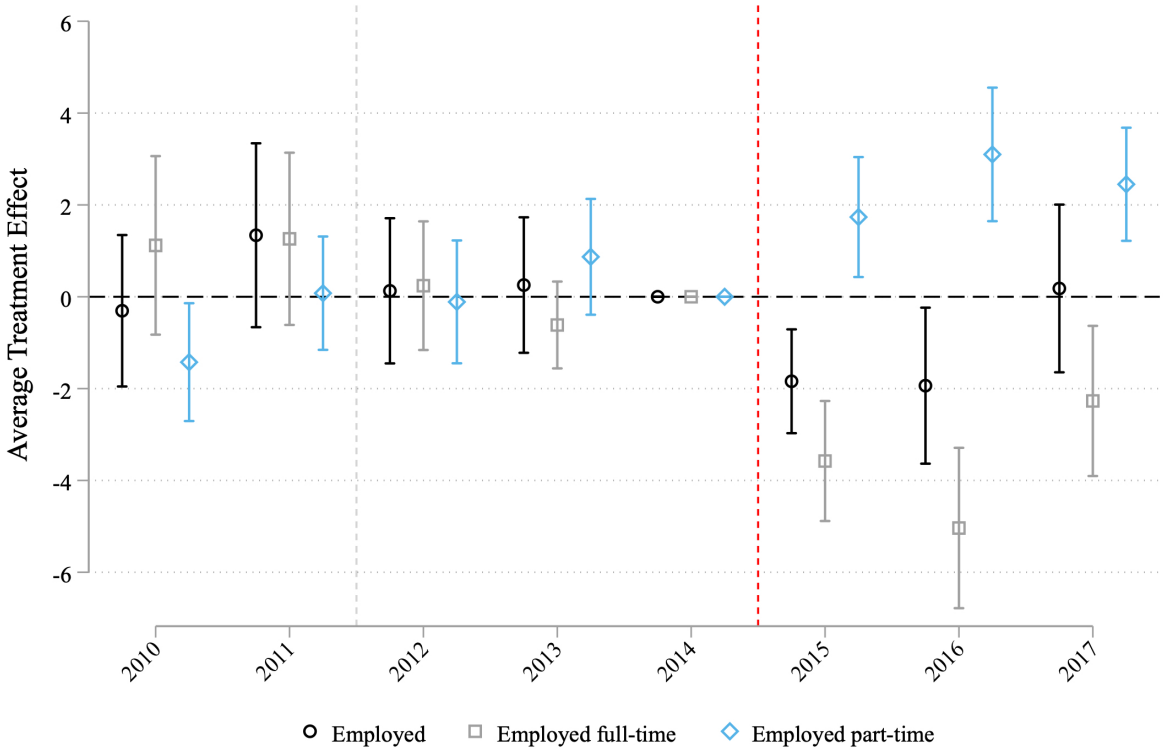
Note: This table reports the results of Equation (2.1), estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, who do not have a college degree, and who are not potential qualifying children as define in Appendix A.2.2, via the American Community Survey (ACS) from 2010-2017. Panels A, B, and C vary only in the dependent variables, which are indicators for any employment, full-time employment, and part-time employment at any point in the last 12-months, respectively, conditional on the individual reporting and wage income. Full time work is defined as working at least 35 hours a week, and part time work is defined as working under 35 hours a week. In columns 2, 3, and 4, demographic controls include indicators for 5-year age bins, race and Hispanic origin groups, educational attainment, bins for the age of the individual's youngest qualifying child, and the number of adults in a individual's household. Standard errors clustered at the state level. * 0.10 ** 0.05 *** 0.01

Table A.14: Triple-difference estimates of the effect of the CalEITC on self-employment

	(1) Basic Triple Difference	(2) Add Demographic Controls	(3) Add Unemployment Controls	(4) Add Minimum Wage Controls
Panel A: Self-employed last week				
ATE	1.1***	1.0***	1.3***	1.4***
	(0.1)	(0.1)	(0.2)	(0.2)
Observations	441,172	441,172	441,172	441,172
Adj. R-Square	0.002	0.010	0.010	0.010
Treated group mean in pre-period	4.5	4.5	4.5	4.5
Implied employment effect	18,651	17,795	22,998	24,419
Panel A: Any self-employment earnings in the last 12 months				
ATE	1.0***	1.0***	1.3***	1.5***
	(0.1)	(0.1)	(0.2)	(0.3)
Observations	441,172	441,172	441,172	441,172
Adj. R-Square	0.002	0.009	0.009	0.009
Treated group mean in pre-period	3.8	3.8	3.8	3.8
Implied employment effect	17,308	16,408	22,829	26,203
Panel B: More than $1,000 in self-employment earnings in the last 12-months				
ATE	0.7***	0.7***	1.0***	1.0***
	(0.1)	(0.1)	(0.2)	(0.3)
Observations	441,172	441,172	441,172	441,172
Adj. R-Square	0.002	0.009	0.009	0.009
Treated group mean in pre-period	3.5	3.5	3.5	3.5
Implied employment effect	12,363	11,565	16,319	17,495
Triple-Difference	Yes	Yes	Yes	Yes
Add Demographic Controls	No	Yes	Yes	Yes
Add Unemployment Controls	No	No	Yes	Yes
Add Minimum Wage Controls	No	No	No	Yes

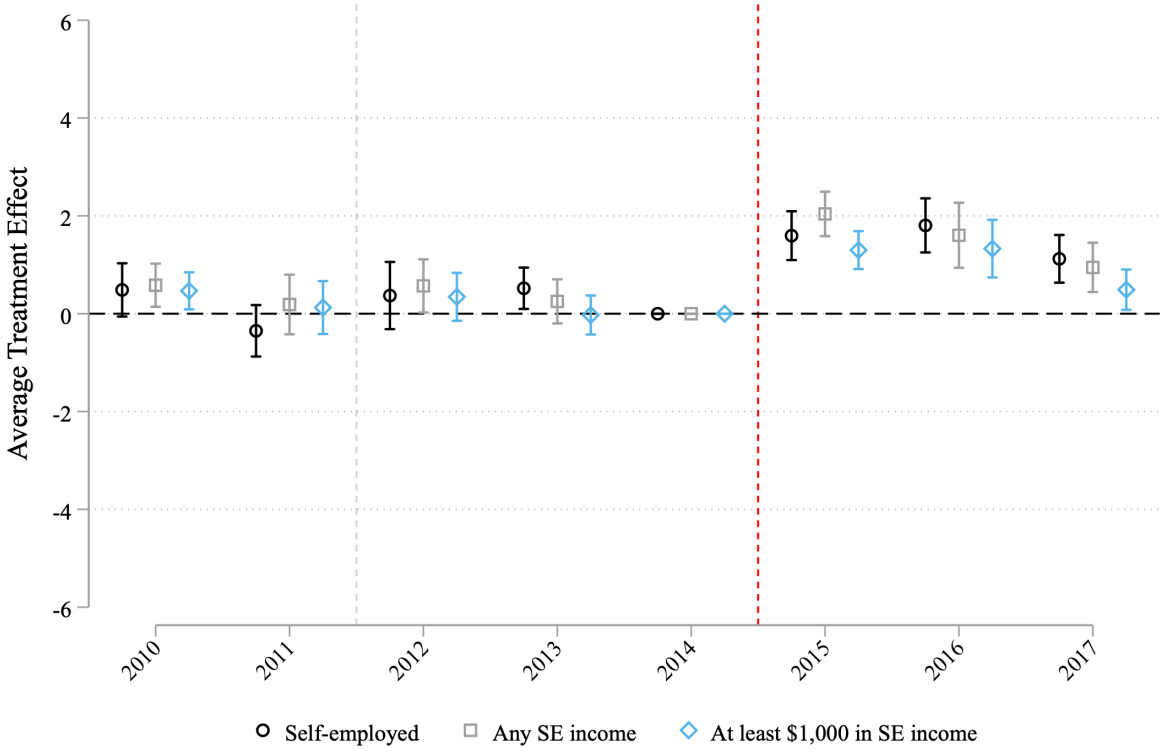
Note: This table reports the results of Equation (2.1), estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, not potential qualifying children as define in Appendix A.2.2, and without a college degree via the American Community Survey (ACS) from 2010-2017. Panel's A, B, and C vary only in the dependent variables, which are indicators for any self-employment income over the last 12 months, over $1,000 in self-employment income over the last 12 months, and whether an individual was self-employed last week, respectively. In columns 2, 3, and 4, demographic controls include indicators for 5-year age bins, race and Hispanic origin groups, educational attainment, bins for the age of the individual's youngest qualifying child, and the number of adults in a individual's household. Standard errors clustered at the state level. * 0.10 ** 0.05 *** 0.01

Figure A.20: Event-study estimates of the effect of the CalEITC on employment restricted to wage workers



Note: This figure reports the results of Equation (2.3), estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, not potential qualifying children as define in Appendix A.2.2, and who do not have a college degree, via the American Community Survey from 2010-2017. Each of the event-studies has a different dependent variable, which are indicators for any employment, full-time employment, and part-time employment at any point in the last 12-months, respectively, conditional on the individual reporting and wage income. Each model is run with the full set of demographic controls, as well as controls for state unemployment and state minimum wages interacted with the number of qualifying children. Standard errors are clustered at the state-level.

Figure A.21: Event-study estimates of the effect of the CalEITC on annual self-employment.



Note: This figure reports the results of Equation (2.3), estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, not potential qualifying children as define in Appendix A.2.2, and who do not have a college degree, via the American Community Survey from 2010-2017. Each of the event-studies has a different dependent variable, defined as an indicator equal to 100 (for scaling purposes) if an individual had any self-employment income over the last 12 months, over $1,000 in self-employment income over the last 12 months, and whether an individual was self-employed last week, respectively. Each model is run with the full set of demographic controls, as well as controls for state unemployment and state minimum wages interacted with the number of qualifying children. Standard errors are clustered at the state-level.

A.4 Elasticity Calculations

First, when estimating the extensive margin elasticity, I follow Kleven (2024) and estimate the following:

$$\zeta = \frac{\frac{\Delta E}{E}}{\frac{\Delta(1-\tau)}{(1-\tau)}} \tag{A.2}$$

Where ζ is the extensive margin elasticity, $\frac{\Delta E}{E}$ represents the percent change in employment as a result of the implementation of the CalEITC, and $\frac{\Delta(1-\tau)}{(1-\tau)}$ represents the analogous change in after-tax rates. I assume that the numerator of Equation (A.2) is the percent change in part-time employment as a result of the CalEITC. Therefore, E is the pre-period percent of the treated group that is employed part-time, and ΔE is the estimated effect of the CalEITC on part-time via Equation (2.1). The denominator is the percent change in the after-tax rate faced by the single mothers in California. Following Kleven (2024), I derive the net-of-tax rate using TAXSIM, assuming that each women has earned income that would maximize their CalEITC credit amount. For women prior to 2015, I take the level of annual earnings that maximized CalEITC receipt in 2015, and deflate it. As noted by Kleven (2024), this approach makes two assumptions: it assumes that women enter the labor market at the CalEITC kink point (the federal EITC kink point in Kleven's case), and that there is no effect on the net-of-tax rate of other non-tax-based social safety net programs. Both assumptions introduce some element of bias into the denominator of Equation (A.2). If some fraction of the part-time effect instead reflects labor market entry at a marginally higher level of earnings, as discussed in **??**, then I am overstating the change in $1-\tau$, meaning my elasticity estimate is biased upward. The exclusion of other non-tax programs has a similar effect, as it does not account for the loss of any program eligibility by entering employment. Therefore it is likely that my estimates of the participation elasticities should be seen as an upper bound.

To estimate the mobility elasticities, I follow closely the calculations in Saez (2002), Kreiner et al. (2015), and Kosonen and Matikka (2023). Starting with formulation in Saez (2002):

$$\eta = \frac{\frac{\Delta E_{j,j+1}}{E_j}}{\frac{\Delta((1-\tau_j)-(1-\tau j+1)}{(1-\tau_j)-(1-\tau j+1}} \tag{A.3}$$

Where η is the mobility elasticity. The numerator of this elasticity represents the percent change in employment at a given annual income level j due to a shift from a different point in the earnings distribution, $j+1$. This exercise is based in part on the estimates of the elasticity of taxable income as shown in Kreiner et al. (2015) and Kosonen and Matikka (2023). However, to maintain comparability with the literature on labor supply responses to tax changes, I focus on the change in employment in response to a change in tax rates. The denominator is the percent change in the difference in the after-tax rate at the same two point, j and $j+1$, on the annual earnings distribution, due to the CalEITC. I assume that j corresponds with a level of earnings that maximizes the CalEITC, which varies by year and count of qualifying children. I present two alternatives for the earning level of $j+1$. First, I assume $j+1$ reflects a full-time worker employed at the minimum wage as of 2017, with annual income equal to $21,840. Second, I use the estimated point on the earnings distribution where we observe a substantial reduction in full-time work, as observed in Figure A.19, namely earnings equal to $27,000. I approximate this elasticity dividing the change in the log share of part-time employment by the log change in the difference between net of tax earnings, following Chetty et al. (2013), Bastian (2020), and Kosonen and Matikka (2023).[2] I calculate this elasticity based on tax rates across 1, 2, and 3+ qualifying children, and present the average (weighted by the share of my sample with the appropriate number of QC).

[2]Specifically, my numerator is $\log(P+\Delta P) - \log(P)$, where P is the share of the population that is employed and ΔP is the estimated part-time effect of the CalEITC. My denominator is $\log(\text{abs}(n_{2017,j} - n_{2017,j+1})) - \log(\text{abs}(n_{2014,j} - n_{2014,j+1}))$, where $n_{y,j}$ is the net-of-tax earnings for a given year and point j.

A.5 Inference

As noted in the seminal paper by Bertrand et al. (2004), inference in difference-in-difference models is made difficult by the potential that errors might be correlated within groups and across time, which would lead to an underestimation of the standard errors. The most common approach in the literature is to use the cluster-robust variance estimator (CRVE), also referred to as clustering standard errors at the group-level (see Abadie et al. (2022) and MacKinnon et al. (2023) for recent reviews). This approach allows unrestricted heteroskedasticity and/or dependence within a given cluster (which can be a state, county, country, classroom, etc..., depending on the context), but assumes that errors are independent across clusters. When there are a large number of clusters that are reasonably homogeneous, with a mix of treated and control clusters, the cluster-robust standard errors generated using this procedure are asymptotically valid. However, there are a number of cases where this does not hold for finite samples. I highlight three of them here that are relevant in my empirical setting: when there are few clusters, when there is heterogeneity across clusters in the size of each cluster, and when the number of treated clusters is small. In each of these cases, CRVE-based inference risks over-rejecting the null hypothesis that there is no effect. Unsurprisingly, these issues interact with each other, making many of the proposed solutions for one case untenable in settings where multiple issues are present.

In my main sample, I have 28 clusters, with one treated cluster, and substantial heterogeneity across clusters. Taking the leading case of heterogeneity discussed in the literature, that of cluster size, there is a wide range in the number of single women in each state and year. Table A.1 shows the mean individual-weighted sample size for each state, averaging over 2010 through 2017. California, the only treated unit, has a mean population of 1.5 million single women per

year, while the next largest state–Texas–has 1.2 million, followed by Florida and New York at 875,000 and 754,000 respectively. All other states have an annual sample of less than 600,000, less than half of that of California. In this setting, therefore, it is important to consider whether my results are robust to alternative inference procedures.

As a brief overview, when there are relatively few clusters overall, as noted by Bertrand et al. (2004) and Donald and Lang (2007), among others, t-tests based on CRVE estimates tend to over-reject. This issue is exacerbated by heterogeneity across clusters, with the leading example being heterogeneity in cluster size (MacKinnon et al., 2023, MacKinnon and Webb, 2017). In some cases, even with a large number of clusters, having clusters of very different sizes can induce severe over-rejection. The wild-cluster bootstrap, proposed by Cameron et al. (2008), was shown by MacKinnon and Webb (2017) to preform fairly well in this setting, providing reliable inference.

A further issue emerges when there is only one treated cluster. Even with many homogeneous clusters, if only one of them is treated, standard CRVE has the potential to over-reject. Consider, following closely the set-up in Ferman and Pinto (2019), the simple case with a state-by-year panel, where a researcher wants to estimate the following difference-in-difference model, estimating the effect of some treatment, impacting some group of N states, who are all treated after some time t^* (with T denoting the last period):

$$y_{st} = \alpha + \beta D_{st} + \lambda_s + \gamma_t + \epsilon_{st}$$

Where y_{st} is our outcome of interest, λ_s and γ_t represent state and year fixed effects, and the treatment D_{st} is an indicator equal to 100 (for scaling purposes) if a state is among a pool

of treated states (denoted as states 1 through N_1) in the period after treatment begins ($t > t^*$). The pool of untreated states are denoted as states $N_1 + 1$ through N (with $N_0 = N - N_1$). I am therefore assuming there are no states that are always treated, and no staggered treatment, avoiding the further complications of variation in treatment time and heterogeneous treatment effects. Inference in this setting is discussed in a recent working paper by Alvarez and Ferman (2023).

The difference-in-difference estimator is written as:

$$\hat{\beta} = \frac{1}{N_1} \sum_{j=1}^{N_1} \left[\frac{1}{T - t^*} \sum_{t=t^*+1}^{T} y_{jt} - \frac{1}{t^*} \sum_{t=1}^{t^*} y_{jt} \right] - \frac{1}{N_0} \sum_{j=N_1+1}^{N} \left[\frac{1}{T - t^*} \sum_{t=t^*+1}^{T} y_{jt} - \frac{1}{t^*} \sum_{t=1}^{t^*} y_{jt} \right]$$

Now, define W_j as the pre-post difference in average errors for each group j:

$$W_j = \frac{1}{T - t^*} \sum_{t=t^*+1}^{T} \epsilon_{jt} - \frac{1}{t^*} \sum_{t=1}^{t^*} \epsilon_{jt}$$

Which allows us to rewrite our estimator $\hat{\beta}$ as:

$$\hat{\beta} = \beta + \frac{1}{N_1} \sum_{j=1}^{N_1} W_j - \frac{1}{N_0} \sum_{j=N_1+1}^{N} W_j$$

The variance of the estimator $\hat{\beta}$ is given by:

$$\text{var}(\hat{\beta}) = \left[\frac{1}{N_1}\right]^2 \sum_{j=1}^{N_1} \text{var}(W_j) + \left[\frac{1}{N_0}\right]^2 \sum_{j=N_1+1}^{N} \text{var}(W_j)$$

Under the standard assumptions, the ϵ_{st} in the model above are allowed to be correlated within a state s over time, but should be uncorrelated across different states. Given this, the estimator of the variance is given by:

$$\widehat{\text{var}(\hat{\beta})}_{cluster} = \left[\frac{1}{N_1}\right]^2 \sum_{j=1}^{N_1} \widehat{W_j}^2 + \left[\frac{1}{N_0}\right]^2 \sum_{j=N_1+1}^{N} \widehat{W_j}^2$$

Where, given the DiD residuals $\hat{\eta}_{jt}$:

$$\widehat{W_j} = \frac{1}{T - t^*} \sum_{t=t^*+1}^{T} \hat{\eta}_{jt} - \frac{1}{t^*} \sum_{t=1}^{t^*} \hat{\eta}_{jt}$$

Note that in the estimate of the variance above, the variance is calculated separately for the treated and untreated units, and then summed. This means that, even with many untreated units, with few (or one) treated units (or vice versa) one component of the variance trends towards 0, as noted by MacKinnon and Webb (2017). In other words, with few or one treated unit, there is not enough information on the errors of the treated group. This has lead econometricians to ask how to utilize information gleaned from the untreated group to fill in this gap, so to speak. Under a certain set of assumptions, namely that the ϵ_{st} are normal, homoskedastic, and serially uncorrelated, appropriate test statistics can be retrieved, as shown by Donald and Lang (2007), using any number of procedures. These include the block bootstrap proposed for the difference-

in-difference setting by Bertrand et al. (2004), the residual-based approach proposed by Conley and Taber (2011), and the coefficient-based randomization inference, as proposed by MacKinnon and Webb (2020), which is identical to the permutation-based tests proposed by Abadie et al. (2015) in cases with one treated unit. All these use information from the control groups (normally either their residuals or estimated coefficients under the null hypothesis) to create a new estimate of the variance.

It now appears that there are a range of procedures that can apply in settings with heterogeniety across clusters (wild-cluster bootstrap), and in settings with one treated unit (variants on block bootstrap procedures and randomization inference). However, each of these approaches runs into problems when both issues are present. In the case of the wild-cluster bootstrap, MacKinnon and Webb (2017) provided simulation evidence that, with one treated unit, the wild cluster bootstrap can severely under-reject (see their Figure 3 for rejection rates by the number of treated clusters), which is the inverse of the problem for CRVE. For the procedures proposed to handle cases with one treated unit, there are similar issues when the assumption of homoskedasticity above does not hold. Ferman and Pinto (2019) established that if there is heterogeneity across group size and if the treated group is large relative to the control group, then methods proposed by Bertrand et al. (2004) and Conley and Taber (2011) will under-reject the null. This concern also exists when using either the randomization-inference or its extension, the bootstrap-based randomization inference procedure proposed by MacKinnon and Webb (2019): in both settings, with one treated unit that is larger than the pool of control states the procedures under-rejects substantially (see figure 3 in MacKinnon and Webb (2020) for graphical evidence of this feature).

In response to these concerns, Ferman and Pinto (2019) developed a procedure that explicitly adjusts for heterogeniety of some known type, with the main case being the size of the

each cluster. The procedure, building on the block-bootstrap proposed by Bertrand et al. (2004), assumes that the sources of the heterogeneity is know; for example, that the heterogeneity comes from variation in the number of individuals in each cluster. The authors propose explicitly adjusting the residuals of the control groups based on estimates of the observed heterogeneity, in essence making the residuals of the control groups appear more similar to that of the treated group. This procedure requires both knowing and observing the sources of the heterogeneity, a strong assumption. A recent working paper by Hagemann (2024) proposes an alternative method, referred to as a rearrangement test, which allows for the form of the heterogeneity to be unknown. I have not yet implemented this methodology.

I examine the robustness of my results to all of these inference procedures, calculating a set of p-values for my estimates in Table 2.2, which I report in Table A.15. I exclude the results for overall employment, but the estimated effects remain statistically insignificant across all inference procedures. First, let us consider the WCBS procedure. As would be expected, given the concern about over-rejection from CRVE, the p-values are larger for both full-time and part-time, but the results remain marginally statistically significant for all models, and statistically significant at the five percent level when I include my full set of controls. This is notable, given the predictions from MacKinnon and Webb (2017) that the WCBS procedure likely under-rejects in cases with one treated unit. I cannot reject the null of no effect for either outcome when using the bootstrap-based randomization inference procedure (RIWB), irrespective of whether I consider a test based on the estimated coefficients or the t-statistics, as suggested by MacKinnon and Webb (2020). This is not surprising, given the strong results of under-rejection for these methods under the assumption that the sole treated cluster is larger that the pool of control clusters. One question that this raises, which I will not address here, is why WCBS and RIWB procedures, both of

which are expected to under-reject, do so to different degrees. Similarly, when using the block bootstrap (BB) procedure of Bertrand et al. (2004), I cannot reject the null of no effect for any model, as once again there is likely to be severe under-rejection in this case. However, I can strongly reject the null for full-time work, and reject the null at the five or ten percent level for part-time work, using the corrected BB procedure described by Ferman and Pinto (2019). For the specification with my full set of controls my results are robust at the five percent level for both full and part-time work using the Ferman and Pinto (2019). One concern with this procedure is that there is heterogeneity separate from that driven by the size of the sample, meaning that the correction does not fully address threat of under-rejection.

Taken together, these results suggest two takeaways. First, inference in this setting is difficult, and as such, any results should be viewed with some level of healthy skepticism. Second, however, is that, even given predictions of under-rejection across a range of models, several, notably the WCBS procedure and the corrected BB procedure, indicate statistical significance at some level for these results.

Table A.15: Triple-difference estimates of the effect of the CalEITC on annual employment, low-education sample, with different inference procedures

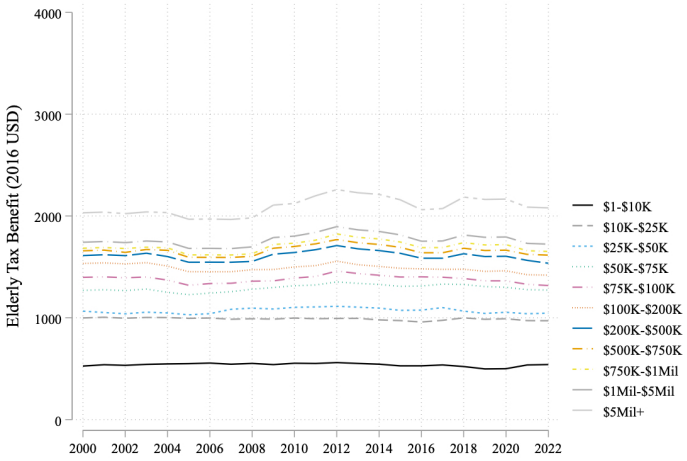
	(1) Basic Triple Difference	(2) Add Demographic Controls	(3) Add Unemployment Controls	(4) Add Minimum Wage Controls
Panel A: Employed full-time in the last 12-months				
ATE	-2.6***	-2.9***	-2.7***	-3.6***
	(0.3)	(0.3)	(0.5)	(0.6)
Observations	588,737	588,737	588,737	588,737
CRVE P-Value	0.000	0.000	0.000	0.000
WCBS P-Value	0.102	0.052	0.018	0.054
RIWB-t P-Value	0.342	0.287	0.274	0.207
RIWB-b P-Value	0.301	0.236	0.397	0.326
BB P-Value	0.389	0.322	0.366	0.216
Corrected BB P-Value	0.026	0.010	0.001	0.000
Panel B: Employed part-time in the last 12-months				
ATE	3.2***	3.1***	2.7***	3.1***
	(0.3)	(0.3)	(0.4)	(0.5)
Observations	588,737	588,737	588,737	588,737
CRVE P-Value	0.000	0.000	0.000	0.000
WCBS P-Value	0.050	0.068	0.018	0.046
RIWB-t P-Value	0.128	0.125	0.204	0.139
RIWB-b P-Value	0.081	0.078	0.280	0.242
BB P-Value	0.138	0.155	0.230	0.151
Corrected BB P-Value	0.089	0.111	0.068	0.038
Triple-Difference	Yes	Yes	Yes	Yes
Add Demographic Controls	No	Yes	Yes	Yes
Add Unemployment Controls	No	No	Yes	Yes
Add Minimum Wage Controls	No	No	No	Yes

Note: This table reports the results of Equation (2.1), estimated on the sample of single women, aged 20-50, who are US citizens, not in the armed services, not in school, who do not have a college degree, and who are not potential qualifying children as define in Appendix A.2.2, via the American Community Survey (ACS) from 2010-2017. Panels A, B, and C vary only in the dependent variables, which are indicators for any full-time employment and part-time employment at any point in the last 12-months. Column 1 reports the results from the full sample, column 2 adds demographic controls, column 3 adds the interactions between state unemployment and the number of qualifying children, and column 4 adds the interaction between the binding state minimum wage and the number of qualifying children. I report a series of p-values associated with each estimate. First, I report the P-value derived from Cluster-Robust Variance Estimator (CRVE), which correspond to the standard errors under the estimates. Second, I report the Wild-Cluster Bootstrap (WCBS) P-values, calculated using the *wildbootstrap* command in Stata. Third, I report two P-values, both constructed using the Randomization Inference Wild-Cluster Bootstrap (RIWB) procedure proposed by MacKinnon and Webb (2019). The first is based on the estimated betas (b) and the second on the estimated t-tests (t). Fourth, I report two P-values derived from Block Boostrap (BB) procedures. The first is the adjusted procedure proposed by Bertrand et al. (2004) and the second uses the adjustment proposed by Ferman and Pinto (2019). * 0.10 ** 0.05 *** 0.01

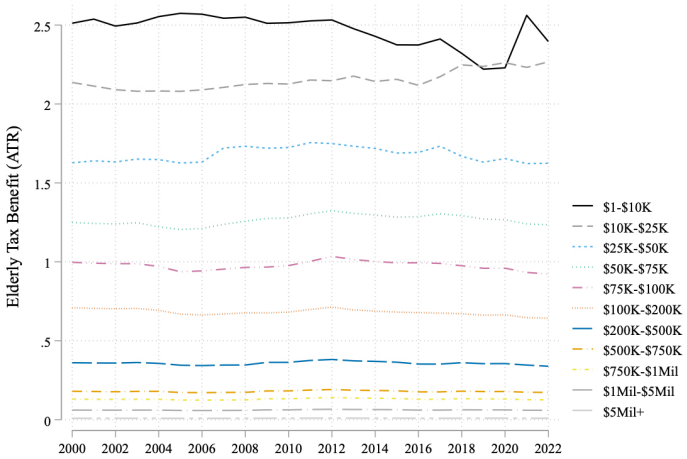
Appendix B: Chapter 3 Appendices

B.1 Additional Tables and Figures

Figure B.1: Average elderly tax benefit across AGI brackets for single returns.



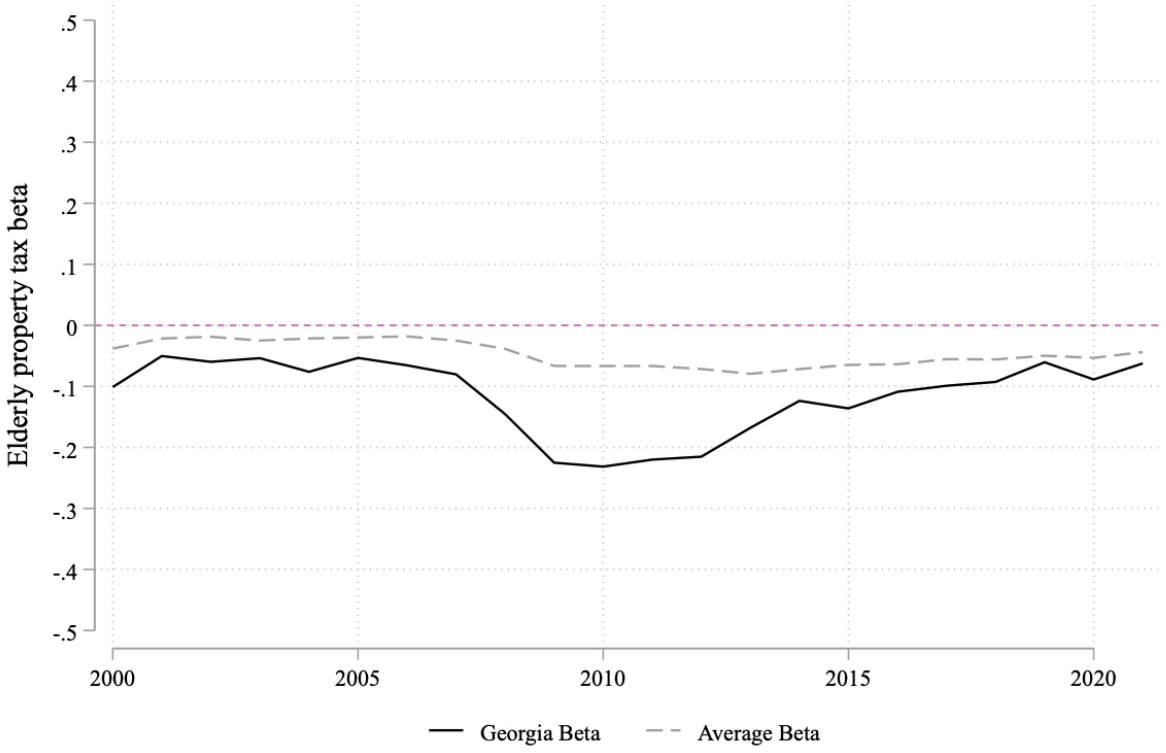
(a) Tax Liability (2019 USD)



(b) Average Tax Rate

Note: Data for these figures are derived from the SOI Databank, from 2000-2022. Each line is the average across all states with individual income taxes, calculated using the income profiles described in Section 3.3.1. The elderly benefit is the difference between the taxes owned (in terms of total liability or average tax rates) between our simulated elderly and non-elderly individual.

Figure B.2: Estimated Effect of Elderly HoH on Property Tax Liabilities in Georgia



Note: Data for these figures are derived from the American Community Survey (ACS) via Ruggles et al. (2023). Plotted are the betas from a series of year-state specific models, regressing property tax rates (property tax over property values) on property values, income, household characteristics, and an indicator for having an elderly head-of-household (HoH). The last of these is the reported beta, and represents the effect of having an elderly HoH on property tax rates. See Appendix B.2 for more details. The black line is the coefficients for Georgia, and the dashed grey line is the average across all states. We see that the elderly-specific generosity in GA became more generous over the period of 2008 through 2015.

Table B.1: Sample composition by age and AGI

Real AGI Brackets (2019 USD)	Young (25-54) Count (rounded to nearest 1,000)	Share	Near-Elderly (55-64) Count (rounded to nearest 1,000)	Share	Elderly (65+) Count (rounded to nearest 1,000)	Share	Total
$1-$10K	156,924,000	7.6%	40,836,000	7.1%	77,364,000	14.3%	275,120,000
$10K-$25K	428,500,000	20.7%	94,359,000	16.5%	136,021,000	25.2%	658,880,000
$25K-$50K	577,000,000	27.8%	142,971,000	25.0%	121,287,000	22.4%	841,260,000
$50K-$75K	369,800,000	17.8%	109,423,000	19.1%	82,116,000	15.2%	561,340,000
$75K-$100K	218,302,000	10.5%	68,760,000	12.0%	49,207,000	9.1%	336,270,000
$100K-$200K	242,751,000	11.7%	84,275,000	14.7%	54,109,000	10.0%	381,140,000
$200K-$500K	63,555,000	3.1%	24,317,000	4.3%	15,516,000	2.9%	103,390,000
$500K-$750K	7,576,000	0.4%	3,208,000	0.6%	2,037,000	0.4%	12,820,000
$750K-$1Mil	2,927,000	0.1%	1,322,000	0.2%	853,000	0.2%	5,100,000
$1Mil+	5,135,000	0.2%	2,577,000	0.5%	1,760,000	0.3%	9,470,000
Total	2,072,470,000	100.0%	572,048,000	100.0%	540,274,000	100.0%	3,184,792,000

Note: Data for this table are derived from the SOI Databank, from 2000 through 2019. We start with a 100 percent sample of all individuals 25 or older in the SOI Databank, and restrict to those individuals who, for a give year t filed taxes in year t and $t-1$, for whom we observe their corrected adjusted gross income (AGI) and who report positive AGI in year t, who did not file a Married Filing Separately return, and who is not matched to an individual associated with a fraud claim. The figures are all based on suppressed flows (no flow with fewer than 10 individuals is included) totals are rounded. AGI brackets are in 2019 USD.

Table B.2: Effect of after-tax rates on inter-state migration by age without controls, via Poisson Pseudo Maximum Likelihood

	(1) $1-$10K	(2) $10K-$25K	(3) $25K-$50K	(4) $50K-$75K	(5) $75K-$100K	(6) $100K-$200K	(7) $200K-$500K	(8) $500K-$750K	(9) $750K-$1M	(10) $1M-$5M	(11) $5M+
Panel A: Elderly (65+)											
Elderly 1-ATR Differential	-0.0024	-0.0125	-0.0020	-0.0173	-0.0086	0.0200	0.0431*	0.0527**	0.0468**	0.0350**	0.0464*
	(0.0180)	(0.0304)	(0.0345)	(0.0229)	(0.0203)	(0.0208)	(0.0204)	(0.0166)	(0.0169)	(0.0110)	(0.0203)
Observations	96311	96805	96729	96330	95380	95076	90003	62187	47215	57665	26790
Population in Sample	1494962	2528798	2268435	1516117	964631	1284561	442304	62306	26920	53375	8812
Percent effect	-.239	-1.24	-.202	-1.71	-.86	2.02	4.41	5.41	4.79	3.56	4.75
Panel B: Near-Elderly (55-64)											
Non-Elderly 1-ATR Differential	-0.0092	-0.0163	-0.0334	-0.0435	-0.0250	0.0082	0.0275*	0.0117	0.0326*	0.0377**	0.0580***
	(0.0118)	(0.0275)	(0.0229)	(0.0254)	(0.0226)	(0.0167)	(0.0134)	(0.0131)	(0.0141)	(0.0123)	(0.0176)
Observations	96425	96786	96824	96767	96140	96121	91371	69692	55404	64106	31008
Population in Sample	1287717	2454713	3024718	2048480	1306447	1888677	684425	93635	39966	78095	11780
Percent effect	-.914	-1.62	-3.28	-4.25	-2.47	.824	2.79	1.18	3.32	3.84	5.97
Panel C: Young (25-54)											
Non-Elderly 1-ATR Differential	0.0044	0.0125	-0.0029	-0.0107	-0.0052	0.0120	0.0215	0.0165	0.0170	0.0250*	0.0213
	(0.0074)	(0.0129)	(0.0134)	(0.0159)	(0.0176)	(0.0173)	(0.0146)	(0.0182)	(0.0157)	(0.0107)	(0.0153)
Observations	96900	96900	96900	96900	96881	96824	95133	78698	65683	72295	40812
Population in Sample	1.06e+07	2.10e+07	2.25e+07	1.16e+07	6225112	7306045	2075389	229200	88296	149150	19885
Percent effect	.436	1.25	-.285	-1.07	-.523	1.2	2.18	1.66	1.71	2.53	2.15

Note: This table reports the results of the Poisson Pseudo Maximum Likelihood estimation of the effect of changes in after-tax rates on inter-state migration, estimated via Equation (3.2) using the tax measure defined in Equation (3.3). Data for this table are derived from the SOI Databank, from 2000 through 2019. We start with a 100 percent sample of individuals 25 years or older who, for a give year t, filed taxes in year t and $t-1$, for whom we observe their corrected adjusted gross income (AGI) and who report positive AGI in year t, who did not file a married filing separately return, and who is not matched to an individual associated with a fraud claim. These data are collapsed to create a count of individuals in a given flow, defined by year, origin and destination state, AGI bracket, age group, and filing status (married or single, with other statuses included in single). The outcome variable is the count of individuals in a given if a given individual flow. Each panel presents the results for a different age group (65+, 55-64, and 25-54), and each row for different AGI groups (assigned via year $t-1$ AGI, with the AGI brackets in 2019 real USD). The conversion of the coefficients into percent terms ($\exp(\beta)-1$) are reported at the bottom of each panel. All models include flow by filing status fixed effects. Standard errors are clustered at the origin-year, destination-year, and flow level * p<0.05 ** p<0.01 *** p<0.001.

Table B.3: Effect of after-tax rates on inter-state migration by age, 2 year leads, via Poisson Pseudo Maximum Likelihood

	(1) $1-$10K	(2) $10K-$25K	(3) $25K-$50K	(4) $50K-$75K	(5) $75K-$100K	(6) $100K-$200K	(7) $200K-$500K	(8) $500K-$750K	(9) $750K-$1M	(10) $1M-$5M	(11) $5M+
Panel A: Elderly (65+)											
Elderly 1-ATR Differential	0.0001*	0.0001*	0.0001	-0.0000	0.0000	-0.0000	0.0009	0.0060	0.0154*	0.0148*	0.0269
	(0.0000)	(0.0001)	(0.0001)	(0.0002)	(0.0004)	(0.0005)	(0.0009)	(0.0032)	(0.0068)	(0.0072)	(0.0221)
Observations	101440	101940	101820	101420	100440	100100	94860	64411	47730	59469	24686
Population in Sample	1580974	2653050	2354129	1551623	979048	1301450	449461	63578	27526	54471	9002
Percent effect	.00567	.0137	.0134	-.00402	.0035	-.00248	.0854	.599	1.55	1.49	2.73
Panel B: Near-Elderly (55-64)											
Non-Elderly 1-ATR Differential	0.0001	0.0000	-0.0000	0.0000	-0.0000	-0.0001	0.0006	0.0039	0.0022	0.0235**	0.0327
	(0.0000)	(0.0001)	(0.0001)	(0.0002)	(0.0003)	(0.0005)	(0.0007)	(0.0032)	(0.0060)	(0.0079)	(0.0184)
Observations	101540	101900	101920	101860	101200	101200	96200	73280	57881	67622	30079
Population in Sample	1361540	2578757	3155765	2108684	1332785	1917551	694857	95500	40871	80002	12080
Percent effect	.00521	.00256	-.000148	.00479	-.00199	-.0141	.0618	.391	.219	2.37	3.32
Panel C: Young (25-54)											
Non-Elderly 1-ATR Differential	-0.0000	0.0000	0.0001	0.0000	-0.0001	-0.0000	0.0002	0.0028	0.0014	0.0070	0.0358**
	(0.0000)	(0.0001)	(0.0001)	(0.0002)	(0.0003)	(0.0005)	(0.0007)	(0.0019)	(0.0043)	(0.0052)	(0.0122)
Observations	102000	102000	102000	102000	101980	101920	100160	83076	69423	76615	42045
Population in Sample	1.14e+07	2.23e+07	2.36e+07	1.20e+07	6372861	7442881	2113447	234941	90819	154236	20583
Percent effect	-.000181	.00241	.00809	.00166	-.00873	-.00255	.0243	.284	.141	.699	3.64

Note: This table reports the results of the Poisson Pseudo Maximum Likelihood estimation of the effect of changes in after-tax rates on inter-state migration, estimated via Equation (3.2) using the tax measure defined in Equation (3.3), using a 2-year lead of the tax variable. Data for this table are derived from the SOI Databank, from 2000 through 2019. We start with a 100 percent sample of individuals 25 years or older who, for a give year t, filed taxes in year t and $t-1$, for whom we observe their corrected adjusted gross income (AGI) and who report positive AGI in year t, who did not file a married filing separately return, and who is not matched to an individual associated with a fraud claim. These data are collapsed to create a count of individuals in a given flow, defined by year, origin and destination state, AGI bracket, age group, and filing status (married or single, with other statuses included in single). The outcome variable is the count of individuals in a given if a given individual flow. Each panel presents the results for a different age group (65+, 55-64, and 25-54), and each row for different AGI groups (assigned via year $t-1$ AGI, with the AGI brackets in 2019 real USD). The conversion of the coefficients into percent terms ($\exp(\beta)-1$) are reported at the bottom of each panel. All models include flow by filing status fixed effects, origin-region by destination-regions by year fixed effects, year fixed effects, and controls for state-level population, policy, and economic characteristics, expressed as the difference between the destination and origin states. Standard errors are clustered at the origin-year, destination-year, and flow level * p<0.05 ** p<0.01 *** p<0.001.

Table B.4: Effect of elderly and non-elderly after-tax rates on inter-state migration by age without controls, via Poisson Pseudo Maximum Likelihood

	(1) $1-$10K	(2) $10K-$25K	(3) $25K-$50K	(4) $50K-$75K	(5) $75K-$100K	(6) $100K-$200K	(7) $200K-$500K	(8) $500K-$750K	(9) $750K-$1M	(10) $1M-$5M	(11) $5M+
Panel A: Elderly (65+)											
ETB	-0.0009	-0.0117	0.0043	-0.0172	-0.0246	-0.0119	-0.0158	0.0114	0.0764	0.0002	0.6325
	(0.0264)	(0.0341)	(0.0427)	(0.0300)	(0.0231)	(0.0295)	(0.0478)	(0.0994)	(0.1466)	(0.0217)	(0.4264)
NET	-0.0080	-0.0134	-0.0089	-0.0174	0.0016	0.0327	0.0442*	0.0510**	0.0477**	0.0388***	0.0465*
	(0.0361)	(0.0320)	(0.0315)	(0.0257)	(0.0246)	(0.0219)	(0.0202)	(0.0175)	(0.0173)	(0.0115)	(0.0204)
Observations	96311	96805	96729	96330	95380	95076	90003	62187	47215	57665	26790
Population in Sample	1494962	2528798	2268435	1516117	964631	1284561	442304	62306	26920	53375	8812
Percent effect - ETB	-.0918	-1.16	.434	-1.7	-2.43	-1.18	-1.57	1.15	7.94	.023	88.2
Percent effect - NET	-.792	-1.34	-.883	-1.73	.16	3.33	4.52	5.23	4.88	3.96	4.76
Panel B: Near-Elderly (55-64)											
ETB	0.0017	0.0077	0.0340	0.0022	-0.0119	-0.0121	-0.0099	0.0100	-0.0558	0.0153	-0.6609*
	(0.0181)	(0.0343)	(0.0396)	(0.0273)	(0.0223)	(0.0168)	(0.0249)	(0.0673)	(0.0874)	(0.0251)	(0.2915)
NET	-0.0137	-0.0177	-0.0143	-0.0342	-0.0289	0.0011	0.0251	0.0137	0.0326*	0.0395**	0.0554**
	(0.0212)	(0.0295)	(0.0298)	(0.0237)	(0.0221)	(0.0166)	(0.0136)	(0.0130)	(0.0149)	(0.0126)	(0.0170)
Observations	96425	96786	96824	96767	96140	96121	91371	69692	55404	64106	31008
Population in Sample	1287717	2454713	3024718	2048480	1306447	1888677	684425	93635	39966	78095	11780
Percent effect - ETB	.166	.776	3.46	.224	-1.18	-1.2	-.989	1	-5.43	1.55	-48.4
Percent effect - NET	-1.36	-1.75	-1.42	-3.37	-2.85	.105	2.54	1.38	3.32	4.02	5.69
Panel C: Young (25-54)											
ETB	0.0035	0.0082	0.0033	-0.0099	-0.0117	-0.0049	0.0149	0.0225	0.1088	0.0082	-0.2378
	(0.0116)	(0.0222)	(0.0215)	(0.0204)	(0.0158)	(0.0169)	(0.0228)	(0.0628)	(0.0733)	(0.0232)	(0.4944)
NET	0.0060	0.0113	-0.0011	-0.0136	-0.0123	0.0078	0.0237	0.0195	0.0229	0.0272*	0.0236
	(0.0123)	(0.0165)	(0.0167)	(0.0151)	(0.0164)	(0.0168)	(0.0152)	(0.0161)	(0.0164)	(0.0118)	(0.0154)
Observations	96900	96900	96900	96900	96881	96824	95133	78698	65683	72295	40812
Population in Sample	1.06e+07	2.10e+07	2.25e+07	1.16e+07	6225112	7306045	2075389	229200	88296	149150	19885
Percent effect - ETB	.347	.825	.327	-.985	-1.16	-.49	1.5	2.28	11.5	.82	-21.2
Percent effect - NET	.603	1.14	-.11	-1.35	-1.23	.786	2.39	1.97	2.31	2.76	2.39

Note: This table reports the results of the Poisson Pseudo Maximum Likelihood estimation of the effect of changes in after-tax rates on inter-state migration, estimated via Equation (3.6) using the 1-year lag of the tax measure defined in Equation (3.4) and (3.5). Data for this table are derived from the SOI Databank, from 2000 through 2019. We start with a 100 percent sample of individuals 25 years or older who, for a give year t, filed taxes in year t and $t-1$, for whom we observe their corrected adjusted gross income (AGI) and who report positive AGI in year t, who did not file a married filing separately return, and who is not matched to an individual associated with a fraud claim. These data are collapsed to create a count of individuals in a given flow, defined by year, origin and destination state, AGI bracket, age group, and filing status (married or single, with other statuses included in single). The outcome variable is the count of individuals in a given if a given individual flow. Each panel presents the results for a different age group (65+, 55-64, and 25-54), and each row for different AGI groups (assigned via year $t-1$ AGI, with the AGI brackets in 2019 real USD). The conversion of the coefficients into percent terms ($\exp(\beta)-1$) are reported at the bottom of each panel. All models include flow by filing status fixed effects. Standard errors are clustered at the origin-year, destination-year, and flow level * p<0.05 ** p<0.01 *** p<0.001.

Table B.5: Effect of after-tax rates on inter-state migration by age, 2 year leads, via Poisson Pseudo Maximum Likelihood

	(1) $1-$10K	(2) $10K-$25K	(3) $25K-$50K	(4) $50K-$75K	(5) $75K-$100K	(6) $100K-$200K	(7) $200K-$500K	(8) $500K-$750K	(9) $750K-$1M	(10) $1M-$5M	(11) $5M+
Panel A: Elderly (65+)											
ETB	-0.0001	0.0003	0.0066	0.0028	-0.0085	-0.0326	-0.0698	-0.1139	-0.1663	0.0197	0.5511
	(0.0002)	(0.0017)	(0.0111)	(0.0227)	(0.0287)	(0.0289)	(0.0417)	(0.0698)	(0.1318)	(0.0238)	(0.4305)
NET	0.0002	0.0001	-0.0002	-0.0001	0.0003	0.0006	0.0009	0.0038	0.0111	0.0154*	0.0294
	(0.0003)	(0.0008)	(0.0007)	(0.0009)	(0.0010)	(0.0008)	(0.0010)	(0.0033)	(0.0078)	(0.0074)	(0.0219)
Observations	101440	101940	101820	101420	100440	100100	94860	64411	47730	59469	24686
Population in Sample	1580974	2653050	2354129	1551623	979048	1301450	449461	63578	27526	54471	9002
Percent effect - ETB	-.00787	.0303	.663	.28	-.846	-3.21	-6.74	-10.8	-15.3	1.99	73.5
Percent effect - NET	.0229	.00586	-.0222	-.0126	.0306	.0551	.0931	.377	1.11	1.55	2.98
Panel B: Near-Elderly (55-64)											
ETB	-0.0001	0.0001	0.0101	0.0116	0.0026	-0.0098	-0.0575*	-0.0887	-0.2078*	0.0184	-0.1172
	(0.0003)	(0.0018)	(0.0190)	(0.0242)	(0.0236)	(0.0172)	(0.0277)	(0.0488)	(0.1053)	(0.0156)	(0.3650)
NET	0.0002	-0.0000	-0.0006	-0.0004	-0.0001	0.0000	0.0006	0.0017	-0.0044	0.0235**	0.0324
	(0.0004)	(0.0007)	(0.0011)	(0.0009)	(0.0008)	(0.0006)	(0.0007)	(0.0030)	(0.0054)	(0.0077)	(0.0180)
Observations	101540	101900	101920	101860	101200	101200	96200	73280	57881	67622	30079
Population in Sample	1361540	2578757	3155765	2108684	1332785	1917551	694857	95500	40871	80002	12080
Percent effect - ETB	-.012	.0141	1.01	1.16	.265	-.975	-5.59	-8.49	-18.8	1.85	-11.1
Percent effect - NET	.023	-.00226	-.0554	-.0363	-.00822	.000494	.061	.169	-.436	2.37	3.3
Panel C: Young (25-54)											
ETB	-0.0003	-0.0003	0.0001	0.0032	-0.0033	-0.0133	-0.0237	0.0094	-0.0990	0.0124	0.1379
	(0.0007)	(0.0032)	(0.0204)	(0.0227)	(0.0164)	(0.0167)	(0.0213)	(0.0391)	(0.0729)	(0.0104)	(0.2495)
NET	0.0003	0.0002	0.0001	-0.0001	-0.0000	0.0002	0.0002	0.0028	-0.0016	0.0076	0.0404**
	(0.0008)	(0.0013)	(0.0012)	(0.0009)	(0.0006)	(0.0006)	(0.0007)	(0.0020)	(0.0046)	(0.0051)	(0.0127)
Observations	102000	102000	102000	102000	101980	101920	100160	83076	69423	76615	42045
Population in Sample	1.14e+07	2.23e+07	2.36e+07	1.20e+07	6372861	7442881	2113447	234941	90819	154236	20583
Percent effect - ETB	-.0259	-.034	.00936	.316	-.326	-1.32	-2.35	.94	-9.43	1.25	14.8
Percent effect - NET	.0299	.0167	.0098	-.00923	-.000236	.019	.0237	.277	-.156	.758	4.12

Note: This table reports the results of the Poisson Pseudo Maximum Likelihood estimation of the effect of changes in after-tax rates on inter-state migration, estimated via Equation (3.6) using a two-year lead of the tax measures defined in Equation (3.4) and Equation (3.5). Data for this table are derived from the SOI Databank, from 2000 through 2019. We start with a 100 percent sample of individuals 25 years or older who, for a give year t, filed taxes in year t and $t-1$, for whom we observe their corrected adjusted gross income (AGI) and who report positive AGI in year t, who did not file a married filing separately return, and who is not matched to an individual associated with a fraud claim. These data are collapsed to create a count of individuals in a given flow, defined by year, origin and destination state, AGI bracket, age group, and filing status (married or single, with other statuses included in single). The outcome variable is the count of individuals in a given if a given individual flow. Each panel presents the results for a different age group (65+, 55-64, and 25-54), and each row for different AGI groups (assigned via year $t-1$ AGI, with the AGI brackets in 2019 real USD). The conversion of the coefficients into percent terms ($\exp(\beta)-1$) are reported at the bottom of each panel. All models include flow by filing status fixed effects, origin-region by destination-regions by year fixed effects, year fixed effects, and controls for state-level population, policy, and economic characteristics, expressed as the difference between the destination and origin states. Standard errors are clustered at the origin-year, destination-year, and flow level * p<0.05 ** p<0.01 *** p<0.001.

Table B.6: Effect of elderly-specific after-tax rates on inter-state migration across the elderly, 2-year leads, via Poisson Pseudo Maximum Likelihood

	(1) $1-$10K	(2) $10K-$25K	(3) $25K-$50K	(4) $50K-$75K	(5) $75K-$100K	(6) $100K-$200K	(7) $200K-$500K	(8) $500K-$750K	(9) $750K-$1M	(10) $1M-$5M	(11) $5M+
Panel A: Elderly, 65-74											
ETB	-0.0001	0.0003	0.0101	0.0076	-0.0063	-0.0329	-0.0848	-0.1875*	-0.2654	0.0266	0.7273
	(0.0003)	(0.0024)	(0.0149)	(0.0239)	(0.0291)	(0.0279)	(0.0448)	(0.0770)	(0.1481)	(0.0305)	(0.4594)
NET	0.0002	0.0001	-0.0005	-0.0002	0.0004	0.0004	0.0011	0.0021	0.0134	0.0210*	0.0389
	(0.0004)	(0.0011)	(0.0010)	(0.0010)	(0.0010)	(0.0008)	(0.0010)	(0.0032)	(0.0081)	(0.0094)	(0.0240)
Observations	100560	101500	101440	100760	98620	98180	90044	56843	40071	51016	19839
Population in Sample	894499	1498453	1421023	994384	661702	913040	311883	43330	18533	36740	6230
Percent effect - ETB	-.00968	.0301	1.02	.763	-.63	-3.24	-8.13	-17.1	-23.3	2.7	107
Percent effect - NET	.0216	.00547	-.0452	-.0249	.0378	.0365	.112	.206	1.35	2.12	3.97
Panel B: Elderly, 75+											
ETB	-0.0001	0.0000	0.0003	-0.0060	-0.0118	-0.0210	-0.0209	0.0741	0.1319	0.0006	0.2515
	(0.0005)	(0.0012)	(0.0081)	(0.0221)	(0.0299)	(0.0309)	(0.0384)	(0.1033)	(0.1963)	(0.0368)	(0.7357)
NET	0.0003	0.0002	0.0001	0.0001	0.0001	0.0009	0.0006	0.0091	0.0064	0.0051	0.0062
	(0.0007)	(0.0006)	(0.0005)	(0.0009)	(0.0010)	(0.0009)	(0.0012)	(0.0054)	(0.0103)	(0.0095)	(0.0319)
Observations	99320	101200	100660	99240	95320	95360	81975	42599	27092	38291	10122
Population in Sample	686475	1154597	933106	557239	317346	388410	137578	20248	8992	17731	2762
Percent effect - ETB	-.00761	.00203	.0282	-.596	-1.17	-2.08	-2.06	7.69	14.1	.0635	28.6
Percent effect - NET	.0252	.0151	.0126	.00785	.0115	.0874	.0555	.911	.639	.507	.626

Note: This table reports the results of the Poisson Pseudo Maximum Likelihood estimation of the effect of changes in after-tax rates on inter-state migration, estimated via Equation (3.6) using the two-year lead of the tax measure defined in Equations (3.4) and (3.5). Data for this table are derived from the SOI Databank, from 2000 through 2019. We start with a 100 percent sample of individuals 25 years or older who, for a give year t, filed taxes in year t and $t-1$, for whom we observe their corrected adjusted gross income (AGI) and who report positive AGI in year t, who did not file a married filing separately return, and who is not matched to an individual associated with a fraud claim. These data are collapsed to create a count of individuals in a given flow, defined by year, origin and destination state, AGI bracket, age group, and filing status (married or single, with other statuses included in single). The outcome variable is the count of individuals in a given if a given individual flow. Each panel presents the results for a different age group (65-74, and 75+), and each row for different AGI groups (assigned via year $t-1$ AGI, with the AGI brackets in 2019 real USD). The conversion of the coefficients into percent terms ($\exp(\beta)-1$) are reported at the bottom of each panel. All models include flow by filing status fixed effects, origin-region by destination-regions by year fixed effects, year fixed effects, and controls for state-level population, policy, and economic characteristics, expressed as the difference between the destination and origin states. Standard errors are clustered at the origin-year, destination-year, and flow level * p<0.05 ** p<0.01 *** p<0.001.

Table B.7: Effect of after-tax rates on inter-state migration by among Elderly, dropping no-IIT states, via Poisson Pseudo Maximum Likelihood

	(1) $1-$10K	(2) $10K-$25K	(3) $25K-$50K	(4) $50K-$75K	(5) $75K-$100K	(6) $100K-$200K	(7) $200K-$500K	(8) $500K-$750K	(9) $750K-$1M	(10) $1M-$5M	(11) $5M+
Panel A: Elderly (65+), Basic											
Elderly 1-ATR Differential	-0.0004	0.0072	0.0226	0.0095	0.0109	0.0036	0.0263*	0.0467**	0.0607***	-0.0064	-0.0077
	(0.0095)	(0.0183)	(0.0199)	(0.0180)	(0.0165)	(0.0150)	(0.0130)	(0.0154)	(0.0180)	(0.0123)	(0.0319)
Observations	68362	68590	68514	68286	67640	67507	63705	42897	30817	38604	13549
Population in Sample	939298	1520534	1332186	876394	551328	726945	238642	30465	12607	23898	3714
Percent effect - elderly	-.0381	.724	2.28	.95	1.09	.36	2.66	4.79	6.26	-.64	-.767
Panel B: Elderly (65+), ETB + NET											
ETB	-0.0067	-0.0025	0.0141	-0.0007	-0.0201	-0.0439*	-0.0423	-0.0058	0.0946	-0.0257	-0.5242
	(0.0098)	(0.0229)	(0.0253)	(0.0248)	(0.0219)	(0.0209)	(0.0302)	(0.0681)	(0.1463)	(0.0329)	(0.5090)
NET	0.0169	0.0210	0.0336	0.0190	0.0328	0.0223	0.0272*	0.0439**	0.0620***	-0.0047	-0.0072
	(0.0220)	(0.0213)	(0.0197)	(0.0211)	(0.0189)	(0.0161)	(0.0130)	(0.0148)	(0.0181)	(0.0149)	(0.0321)
Observations	68362	68590	68514	68286	67640	67507	63705	42897	30817	38604	13549
Population in Sample	939298	1520534	1332186	876394	551328	726945	238642	30465	12607	23898	3714
Percent effect - ETB	-.666	-.252	1.42	-.0749	-1.99	-4.29	-4.14	-.582	9.93	-2.53	-40.8
Percent effect - NET	1.7	2.12	3.42	1.92	3.33	2.26	2.76	4.49	6.4	-.464	-.714

Note: This table reports the results of the Poisson Pseudo Maximum Likelihood estimation of the effect of changes in after-tax rates on inter-state migration, estimated via Equation (3.2) using the tax measure defined in Equation (3.3), using a 1-year lag of the tax variable, with geographic data pulled from the F1040 Zipcodes. Data for this table are derived from the SOI Databank, from 2000 through 2019. We start with a 100 percent sample of individuals 25 years or older who, for a give year t, filed taxes in year t and $t-1$, for whom we observe their corrected adjusted gross income (AGI) and who report positive AGI in year t, who did not file a married filing separately return, and who is not matched to an individual associated with a fraud claim. These data are collapsed to create a count of individuals in a given flow, defined by year, origin and destination state, AGI bracket, age group, and filing status (married or single, with other statuses included in single). The outcome variable is the count of individuals in a given if a given individual flow. Each panel represents a different specification for the elderly, and each row for different AGI groups (assigned via year $t-1$ AGI, with the AGI brackets in 2019 real USD). The conversion of the coefficients into percent terms ($\exp(\beta)-1$) are reported at the bottom of each panel. All models include flow by filing status fixed effects, origin-region by destination-regions by year fixed effects, year fixed effects, and controls for state-level population, policy, and economic characteristics, expressed as the difference between the destination and origin states. Standard errors are clustered at the origin-year, destination-year, and flow level * p<0.05 ** p<0.01 *** p<0.001.

Table B.8: Effect of after-tax rates on inter-state migration by among Elderly, using F1040 Zip Codes, via Poisson Pseudo Maximum Likelihood

	(1) $1-$10K	(2) $10K-$25K	(3) $25K-$50K	(4) $50K-$75K	(5) $75K-$100K	(6) $100K-$200K	(7) $200K-$500K	(8) $500K-$750K	(9) $750K-$1M	(10) $1M-$5M	(11) $5M+
Panel A: Elderly (65+), Basic											
Elderly 1-ATR Differential	0.0235	0.0109	0.0163	0.0062	0.0024	0.0169	0.0366*	0.0406*	0.0034	0.0187	-0.0629
	(0.0138)	(0.0213)	(0.0141)	(0.0195)	(0.0144)	(0.0143)	(0.0158)	(0.0182)	(0.0169)	(0.0154)	(0.0349)
Observations	94715	96292	96444	95855	94031	94639	87287	51699	34579	44213	12265
Population in Sample	1049048	1842955	1700783	1147751	733713	1002944	338250	44301	18902	33979	4729
Percent effect	2.38	1.09	1.64	.625	.237	1.7	3.73	4.14	.336	1.89	-6.09
Panel B: Elderly (65+), ETB + NET											
ETB	0.0184	0.0024	0.0077	0.0034	-0.0165	-0.0231	-0.0271	-0.0334	-0.0104	0.0238	-1.1156*
	(0.0141)	(0.0255)	(0.0175)	(0.0297)	(0.0193)	(0.0193)	(0.0408)	(0.0885)	(0.2232)	(0.0383)	(0.4908)
NET	0.0466*	0.0233	0.0270	0.0090	0.0169	0.0392*	0.0415**	0.0395*	0.0053	0.0188	-0.0637
	(0.0186)	(0.0196)	(0.0145)	(0.0164)	(0.0147)	(0.0158)	(0.0151)	(0.0172)	(0.0182)	(0.0168)	(0.0362)
Observations	94715	96292	96444	95855	94031	94639	87287	51699	34579	44213	12265
Population in Sample	1049048	1842955	1700783	1147751	733713	1002944	338250	44301	18902	33979	4729
Percent effect - ETB	1.85	.245	.777	.336	-1.63	-2.29	-2.67	-3.29	-1.03	2.4	-67.2
Percent effect - NET	4.77	2.36	2.73	.902	1.7	3.99	4.23	4.03	.53	1.89	-6.17

Note: This table reports the results of the Poisson Pseudo Maximum Likelihood estimation of the effect of changes in after-tax rates on inter-state migration, estimated via Equation (3.2) using the tax measure defined in Equation (3.3), using a 1-year lag of the tax variable, with geographic data pulled from the F1040 Zipcodes. Data for this table are derived from the SOI Databank, from 2000 through 2019. We start with a 100 percent sample of individuals 25 years or older who, for a give year t, filed taxes in year t and $t-1$, for whom we observe their corrected adjusted gross income (AGI) and who report positive AGI in year t, who did not file a married filing separately return, and who is not matched to an individual associated with a fraud claim. These data are collapsed to create a count of individuals in a given flow, defined by year, origin and destination state, AGI bracket, age group, and filing status (married or single, with other statuses included in single). The outcome variable is the count of individuals in a given if a given individual flow. Each panel represents a different specification for the elderly, and each row for different AGI groups (assigned via year $t-1$ AGI, with the AGI brackets in 2019 real USD). The conversion of the coefficients into percent terms ($\exp(\beta)-1$) are reported at the bottom of each panel. All models include flow by filing status fixed effects, origin-region by destination-regions by year fixed effects, year fixed effects, and controls for state-level population, policy, and economic characteristics, expressed as the difference between the destination and origin states. Standard errors are clustered at the origin-year, destination-year, and flow level * p<0.05 ** p<0.01 *** p<0.001.

B.2 Dataset Construction

As discussed in Section 2.4.2, the data for this project are constructed using confidential IRS tax records, including both tax returns and information returns, linked to social security birth and death records. This appendix describes the process for constructing both the state-to-state flows and the individual-level sample used in this paper.

To construct our datasets, we loop over each year from 2000 through 2022. Table XX shows the sample size at each stage of this process for a selection of representative years. For each year (t), we preform the following process. First, we begin with the Statistics SOI Databank (Chetty et al., 2018), drawing the universe of individuals in year t who were (a) over 22, (b) who filed taxes, and (c) for whom the primary filer's Taxpayer Identification Number (TIN) was valid. From this dataset, we obtain information on the demographics of the taxpayer, including their gender, date of birth, and date of death, from the Social Security Administration, as well as select information from their F1040 and information returns. Second, we check to see that key variables (filing status, date of birth, etc...) are present, and make sure that we observe the correct number of filers based on a return's filing status. We drop individuals who are recorded as both primary filer and spouse on a return, or who are on a return with the incorrect number of filers. Third, given the need to observe location information over two years to measure migration, we draw the universe of individuals from year $t-1$, using the same criteria as in year t. We then match the year t and $t-1$ samples on masked TINs, and drop those individuals not observed in both years. This is the step that reduces the sample the most, by between XX and YY million individuals per year. Fourth, we merge in a computed AGI variable for year $t-1$ from a separate IRS database, which corrects several potential inaccuracies on F1040 forms, and use this variable when present. Fifth,

Table B.9: Elderly Benefit Cost Estimate for 2019

State	State IIT Revenue via DQS (Millions 2019 USD)	State IIT Revenue via TAXSIM (Millions 2019 USD)	Gap (Millions 2019 USD)	Gap (Percent of DQS)
Alaska	-	-	-	0%
Alabama	4,195	2,427	(1,768)	-42%
Arkansas	3,012	1,913	(1,099)	-36%
Arizona	5,357	3,875	(1,482)	-28%
California	100,080	84,078	(16,002)	-16%
Colorado	8,172	5,790	(2,382)	-29%
Connecticut	8,446	7,688	(758)	-9%
Delaware	1,741	887	(854)	-49%
District of Columbia	2,299	1,928	(371)	-16%
Florida	-	-	-	0%
Georgia	12,177	8,705	(3,472)	-29%
Hawaii	2,568	1,984	(584)	-23%
Iowa	4,098	2,556	(1,542)	-38%
Idaho	1,669	1,576	(93)	-6%
Illinois	18,762	14,322	(4,440)	-24%
Indiana	8,853	3,447	(5,406)	-61%
Kansas	3,778	2,607	(1,172)	-31%
Kentucky	4,629	2,885	(1,744)	-38%
Louisiana	3,833	2,427	(1,406)	-37%
Massachusetts	17,151	11,912	(5,239)	-31%
Maryland	10,067	6,883	(3,184)	-32%
Maine	1,709	1,233	(476)	-28%
Michigan	10,150	6,480	(3,670)	-36%
Minnesota	12,400	8,585	(3,815)	-31%
Missouri	6,588	4,688	(1,900)	-29%
Mississippi	1,968	1,223	(746)	-38%
Montana	1,413	1,023	(390)	-28%
North Carolina	13,259	8,595	(4,664)	-35%
North Dakota	415	253	(163)	-39%
Nebraska	2,546	1,800	(746)	-29%
New Hampshire	123	87	(35)	-29%
New Jersey	15,903	15,590	(313)	-2%
New Mexico	1,590	1,025	(565)	-36%
Nevada	-	-	-	0%
New York	48,087	36,428	(11,659)	-24%
Ohio	9,313	6,612	(2,701)	-29%
Oklahoma	3,565	2,373	(1,193)	-33%
Oregon	9,847	5,414	(4,433)	-45%
Pennsylvania	13,517	8,222	(5,295)	-39%
Rhode Island	1,376	976	(401)	-29%
South Carolina	4,759	3,933	(826)	-17%
South Dakota	-	-	-	0%
Tennessee	202	101	(101)	-50%
Texas	-	-	-	0%
Utah	4,979	2,955	(2,024)	-41%
Virginia	14,918	9,743	(5,174)	-35%
Vermont	861	623	(238)	-28%
Washington	-	-	-	0%
Wisconsin	8,760	6,372	(2,388)	-27%
West Virginia	2,097	1,130	(966)	-46%
Wyoming	-	-	-	0%
Total	411,235	303,356	(107,878)	-26%

Note: This table compares the estimates of total state individual income tax benefits for 2019, using data from the SOI Databank, other IRS sources, and the State and Local Finance (SLFI) Data Query System (DQS) (The Urban Institute and The Urban-Brookings Tax Policy Center, 2024). The methodology using the IRS data and TAXSIM under-estimates total individual income tax (IIT) revenues relative to the estimates from the US Census of Governments, collected and organized by SLFI DQS (reported in millions of real 2019 USD). This could be due to sample restrictions implemented within the IRS, under-claiming of state tax programs, or several other reasons. The gap is why we express our results in percent terms.

we drop individuals who are older than 120 as unrealistic, and younger than 25 as potentially impacted by post-college moves. Sixth, using information the IRS has on recorded instances of fraud, we drop any TIN that has been associated with a fraudulent tax return in any tax year.

With this sample, we begin the process of cleaning the observed ZIP codes. We have six ZIP codes at this point: for both the current year (t) and the prior year ($t-1$), we have the ZIP codes on the F1040, the most common ZIP code on all information returns, and the SUPER ZIP, which combines the previous two, preferencing the information return ZIP code. We mostly use the 5-digit ZIP code, and only using the 9-digit ZIPs when a 5-digit ZIP code straddles two states (this is true for 137 ZIP codes in 2020). For the 5-digit ZIP codes, we trim the variables to eliminate leading and lagging spaces, then make sure they have the correct number of digits, are not recorded as "00000", and do not contain non-numeric symbols. In the case of the incorrect number of digits, if the 5-digit ZIP is has too-many digits, we use the first five, but in all other cases we record the ZIPs as missing. We preform similar cleaning procedures on the 9-digit ZIPs.

At this stage, we have the baseline dataset of individuals that we use to either construct the state-to-state flows, the tax profiles, or the individual-level sample.

B.2.1 Flow-level data

For the flow-level dataset, we first restrict the sample to those individuals who are not missing their SUPER ZIP (which means we can observe at least one ZIP-code for the individual). We then proceed to loop over the six ZIP-codes mentioned above to determine the state of residence, using the 5-digit ZIP for those ZIP-codes that are completely within a given state, and the 9-digit ZIPs otherwise. If the first five digits of the 9-digit ZIP do not match 5-digit ZIP code, we assume

that the 9-digit ZIP-code is incorrect, and use the 5-digit ZIP code plus "0000". We then collapse the individual-level data to the state (t) by state ($t-1$) by filing status by year by age group by AGI group by homeowner level. We construct this flow file for our three separate ZIP-codes (SUPER, F1040, and information-returns), and for two different ways of identifying AGI. First, we use the unchanged AGI, and second, we add back in certain sources of eldelry-specific income that are untaxed and therefore not included in AGI; namely, income from social security, pensions, and IRA distributions. This creates six files.

For the poisson regression analysis, we also drop those individuals with negative or $\$0$ in prior-year AGI, those with married filing separately filing status, and restrict the data to 2000 through 2019 who moved between states in a given year.

B.2.2 Income profiles

To construct our income profiles, we start with the baseline dataset discussed above. For each year, we select a random sample of 1,000 tax units for every AGI group by age group (elderly and non-elderly) by filing status (single, married filing jointly, married filing separately) group. For these selected returns, we compile the required information to calculate their federal and state individual income tax liability, using the NBER TAXSIM model (V35). For most elements of required by the calculator, we can obtain the required information from the SOI Databank, but for several (childcare costs and capital gains) we pull information from other IRS databases. Below we detail a few of the assumptions made in this process.

First, we assume that elderly individuals have no dependents (which is on average true for our sample, and for non-elderly individuals we use information of claiming of qualified child care

expenses, personal exemptions, and EITC claiming to estimate the number and age of dependents (up to 3). Second, for wages and self-employment income, we apportion total amounts reported on the 1040 to each spouse for married returns proportional to income of each type reported on information returns (W2s, 1099 MISCs, etc...). Third, due to outliers in the tax data, we cap the maximum amount of taxable and gross social security income at the 99.9th percentile. Finally, we assume several TAXSIM variables are equal to 0 (other non-property income, rent paid, real estate taxes paid, active S-Corp income, and primary and secondary Qualified Business Income (QBI) and Specialized Service Trade / Business Income (SSTB)).

After constructing the income profiles, we take the average of the 1,000 returns to create a representative taxpayer for each year, AGI group, filing status, and age group. To abstract from noise in the annual measures of income in each sample, we take the average income values (after adjusted for inflation) for each marital status, age, and AGI group across the entire period (2000-2022). This creates a representative sample of taxpayers for each group across our sample. In order to provide a sense of the relative benefit the tax code of a given state might provide to the elderly, we create an alternative version of each of these taxpayers that is non-elderly, by assigning each elderly individual (and their spouse, in the case of the married filing jointly taxpayers) to be aged 45, and by reallocating any source of retirement income (pensions, social security, etc...) to be "other non-property income". There are a few isolated cases where NBER TAXSIM has coded a state-specific policy in some way that interacts with other non-property income in an unusual way. In these cases, we reassign retirement income to be interest income. These cases are Arkansas, Massachusetts, Missouri, Nevada, New Jersey, Pennsylvania, and Tennessee. We then run each representative tax unit through TAXSIM for each year of our sample. By comparing the tax liability of the elderly and non-elderly versions, we can determine the scale of the elderly tax

preference across each state over time, abstracting from changes in the composition and level of income.

B.2.3 Individual-level sample

For the difference-in-difference analysis, we start with the baseline dataset discussed above. We want to construct a stratified random sample that over-samples higher-income taxpayers. To do this, we pull a sample of every tax unit who, between 2000 and 2022, reported AGI (in real 2019 dollars) over $500,000. We then stratify the sample as follows. We take a 100 percent sample of individuals ever on a tax return reporting over $1,000,000 in real 2019 AGI, a 50 percent sample of individuals ever on a tax return reporting over $500,000 in real 2019 AGI (exclusive of the $1,000,000 sample), and a 20 percent sample of all other filers in the baseline dataset. We therefore weight our difference-in-difference models using simple weights of 1, 2, and 5 depending on the AGI groups.

B.2.4 Cost dataset

Importantly, for the calculation of the cost of elderly-specific tax policies, we want to have a sample of tax units without some of the required restrictions applied to the migration sample. Namely, we don't want to limit the dataset to individuals observed in both t and $t-1$. This requires re-creating the analysis dataset for a limit number of years (2017 through 2019) without this restriction. We follow the same process of pulling data from the SOI Databank, plus the procedure for collecting additional data for use in TAXSIM, but for the universe of tax records. We then limit the sample using the same stratified random sampling we used for the individual-

level sample, and weight the resutls accordingly.

B.2.5 Construction of controls

Our Poisson models use a range of control variables, encompassing demographic, economic, and policy characteristics that could potentially both impact migration decisions and be correlated with elderly-specific individual income tax policies. In this subsection, we discuss the sources and creation of these variables. First, we get data on state-by-year-by-age group population estimates from the Centers For Disease Control (CDC) WONDER Online database. We use the 2020 Vintage of the Bridged-Race Population Estimates, produced by the National Center for Health Statistics, based on the estimates of the July 1st resident population of the United States, accessed via: https://wonder.cdc.gov/bridged-race-v2020.html. We use only data on states and age groups. Second, we get data on new mortgages issued from 1998 through 2022 via the Federal Housing Finance Agency's National Mortgage Database. In particular, we download data on New Residential Mortgage Statistics for all markets at the state level via https://www.fhfa.gov/data/national-mortgage-database-aggregate-statistics. In our model, we convert this into per-capita terms, and control for the difference between origin and destination states. Third, we obtain data on state and local revenue and expenditures by state and source collected via the Census of Governments and organized by the Urban Institute's State and Local Finance Initiative https://state-local-finance-data.taxpolicycenter.org/pages.cfm. Here, we have data from 1995 through 2021. Again, these are converted into per-capita terms and included as controls as the difference between origin and destination states. Fourth, we pull data on state-level unemployment from the State Economic Monitor, created by the Urban Institute's

State and Local Finance Initiative, which itself pulls data from the Bureau of Labor Statistics https://apps.urban.org/features/state-economic-monitor/.

We also include two set of controls designed to account for state-level elderly-specific tax breaks. The first is a set of indicators for the presence of specific types of eldelry-specific property tax programs, derived from data collected by the Lincoln Institute of Land Policy as part of their Significant Features of the Property Tax program https://www.lincolninst.edu/data/significant-features-property-tax/access-database/residential-property-tax-relief-programs/. As these data begin in 2006, we assume that there are no changes in these indices between 2000 and 2006. Second, we construct a measure of the elderly-specific property tax benefit at the state level using American Community Survey (ACS) Data from 2001 through 2022. The ACS is designed as a 1 percent sample of the US population, and asks a range of questions regarding individual and household demographic and economic characteristics.[1] Importantly, this includes information about home ownership, property values, and property taxes paid. We construct a simple model where we predict the effect of being elderly on a household's property tax liability, controlling for property values and other characteristics.

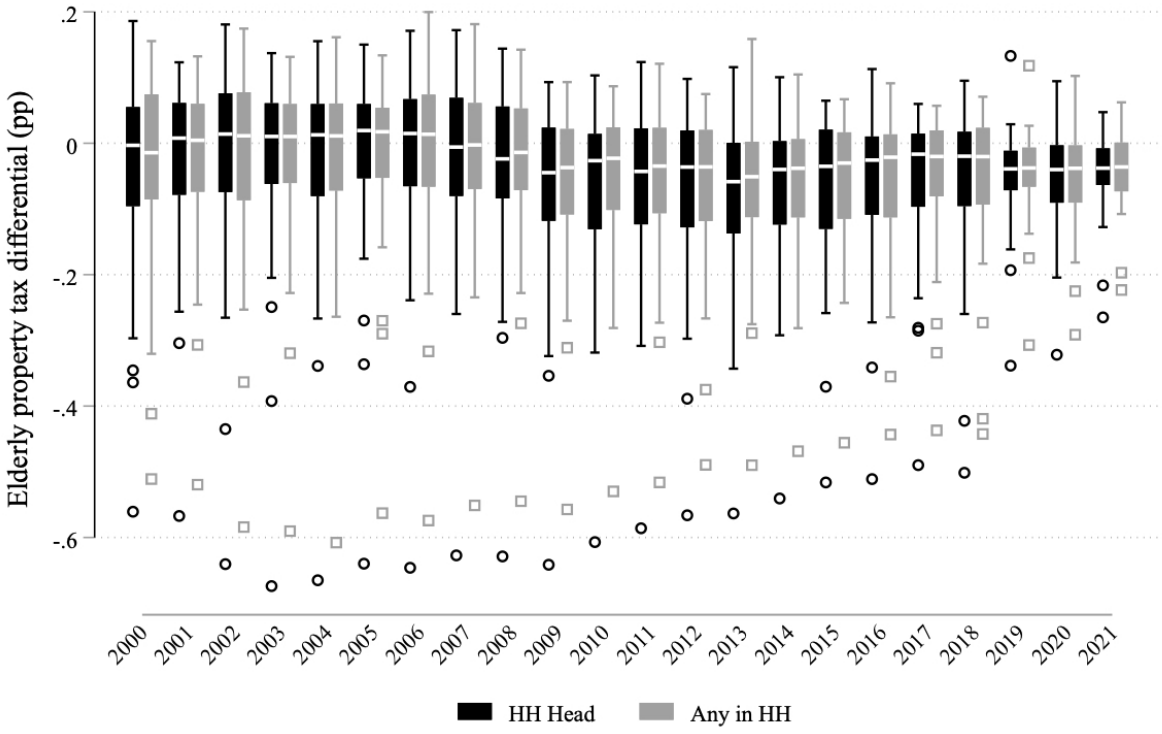
Beginning with ACS data from 2000-2022, obtained via IPUMS (Ruggles et al., 2023), we limit the sample to ACS households (i.e. we are using a household-level weighted sample) who owned their own home, were not in group quarters, who did not move out of state in a given year, who report positive home value and household income, and who do not have imputed property tax or household value characteristics. For each year and state, we separately estimate the following model:

[1]The ACS was rolled out between 2000 and 2005, and for this period had a smaller sample, but was still designed to be representative at the state level.

$$\tau_{hst} = \alpha + \beta E_{hst} + F(V_{hst}) + F(I_{hst}) + \text{Vet}_{hst} + \text{Dis}_{hst} + \epsilon_{hst}$$

Where τ_{hst} is the property tax rate for a given households, defined as the property tax divided by the property value, E_{hst} is an indicator equal to one if the household head is over 62 years old (the normal age cutoff for elderly-specific property tax breaks), $F(.)$ is a quadratic function applied to both household value h_{hst} and household income I_{hst}, and Vet_{hst} and Dis_{hst} are indicators for the household head being either a veteran or disabled, two other groups with targeted property tax breaks. The coefficient beta captures the degree to which households with an elderly head pays a higher or lower property tax rate. Figure B.3 plots the average across states of β under two different modeling approaches: using the characteristics of the household head to determine the elderly, veteran, and disabled indicators, and using the presence of any individuals in a household who are elderly, veterans, or disabled. The results are similar: there are a wide range of elderly indicators, though the average beta is negative (meaning elderly households pay lower property taxes) and the median value trends downwards over time. Importantly, there are some large outliers, like Texas and and New York, that appear in this figure with large (in magnitude) negative betas.

Figure B.3: Elderly Property Tax Differentials By Year and Method



Note: This figure reports the average across states of the betas on the elderly indicator in the regression run on the 2000-2021 ACS sample. The "HH Head" model bases the elderly, veteran, and disabled indicators on the characteristics of the household head, while the "Any in HH" model uses the presence of any individual in the household with those characteristics.

Appendix C: Chapter 4 Appendices

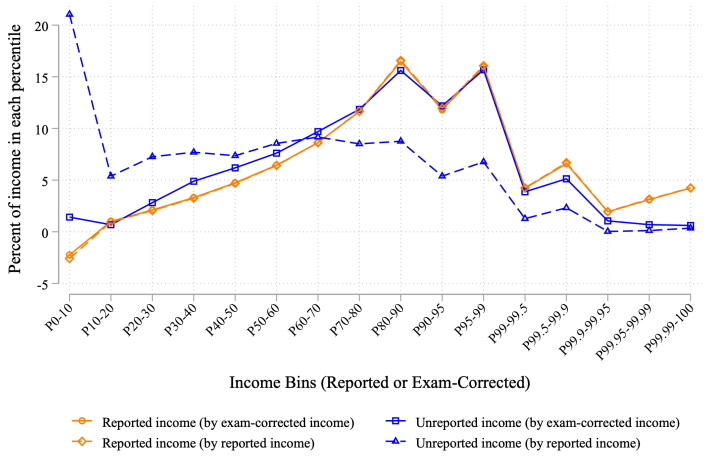
C.1 Additional Tables and Figures

Table C.1: Estimates of the top 1 percent share, varying R^c_{1t} and E^c_{1t}, 1962-2014

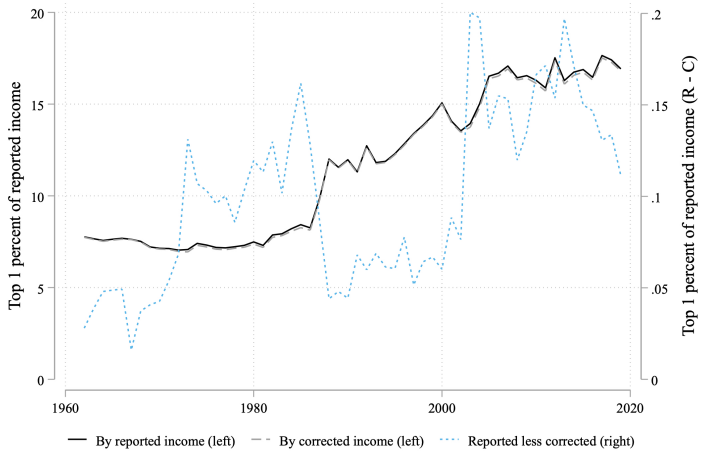
	Income Excl. Misreporting	Assn. 2.1 (AS)	Assn. 2.2 (PSZ)	Assn. 2.3 (Us)
		Reported Fiscal Income		
1962	8.3	8.9	9.0	8.9
2014	17.8	17.7	20.2	19.9
Change	9.5	8.8	11.3	11.0
		Pre-Tax National Income (AS)		
1962	10.8	11.1	11.2	11.2
2014	14.1	14.2	15.9	15.6
Change	3.3	3.0	4.7	4.5
		Pre-Tax National Income (PSZ)		
1962	12.5	12.8	12.9	12.8
2014	17.4	17.3	19.0	18.8
Change	4.8	4.5	6.1	5.9

Notes: This table reports the top 1 percent shares of income (Y^c_{1t}) for 1962 and 2014 (following Table 4 in Auten and Splinter (2024)), as well as the increase in the top 1 percent share (in percentage points) over this period. Each panel corresponds with a different estimate of R^c_{1t}, and each column with a different estimate of E^c_{1t}, or the pre-misreporting top 1 percent share of reported fiscal income or pre-tax National Income under the assumptions of AS and PSZ.

Figure C.1: Evaluating Assumption 1: $R_{1t}^R = R_{1t}^c$



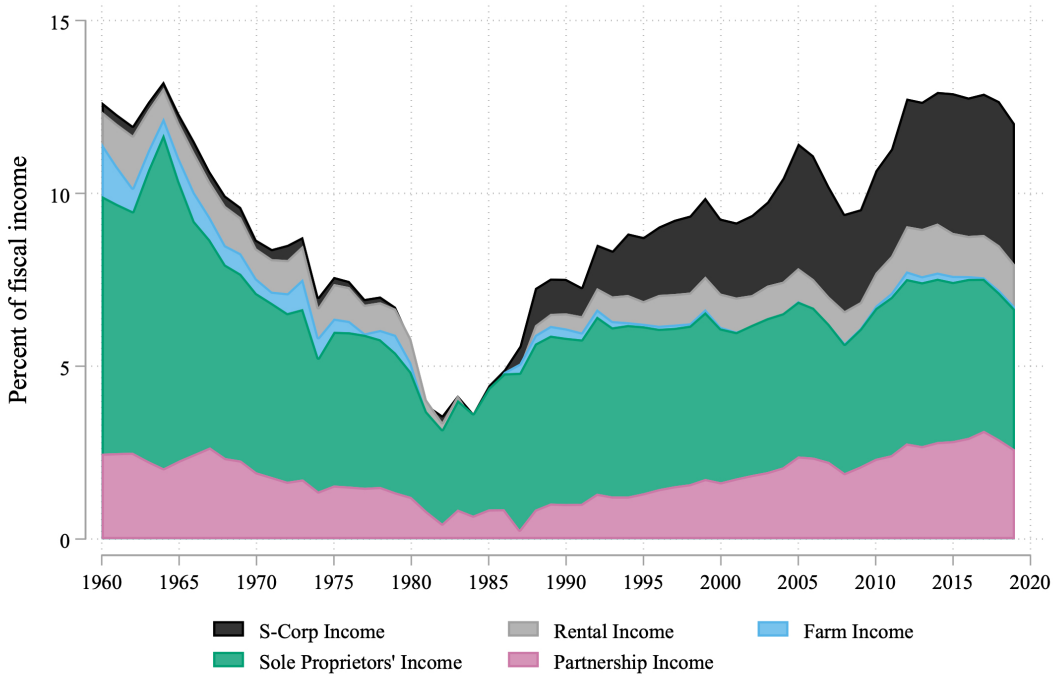
(a) Shares of reported and misreported income across income bins, ranked by either reported or exam-corrected income (2006-2013)



(b) Dynamics of top one percent of reported income ranked by corrected and reported income

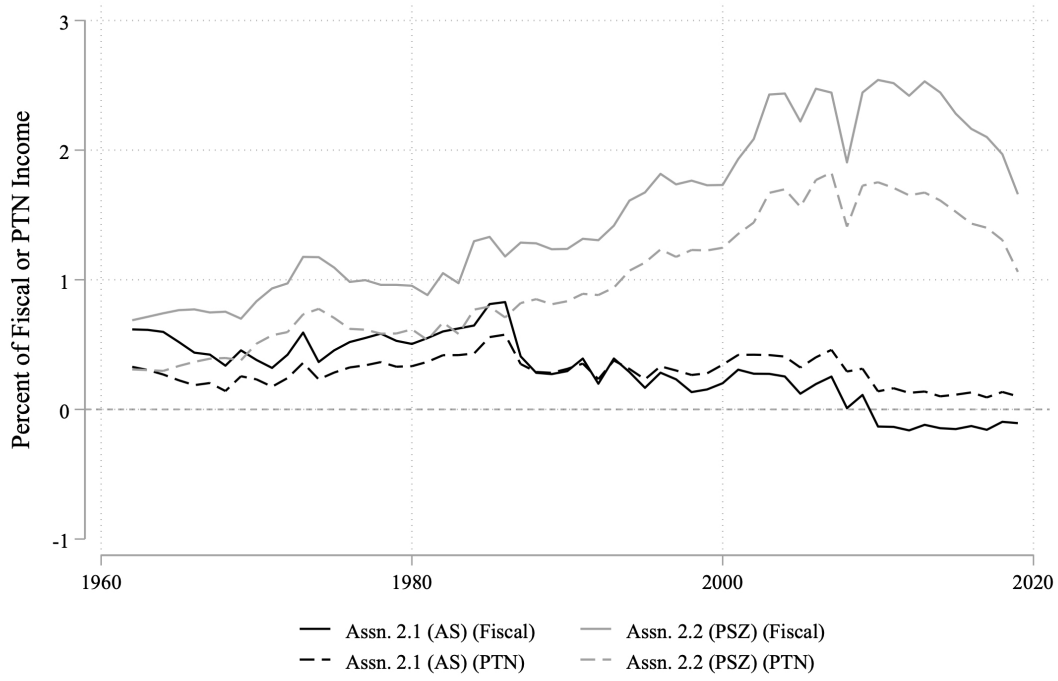
Note: This figure examines Assumption 1: that the top one percent share of reported income ranked by reported income is the same as the top 1 percent share of reported income when ranked by corrected income. The first panel plots the shares of reported and misreported income through the distribution, using NRP data for tax years 2006–2013 from (Guyton et al., 2023a). We include only audit-detected under-reporting (no Detection Controlled Estimation). The solid lines represent income measures sorted by exam-corrected income percentiles, while the dashed lines represent income measures sorted by reported-income percentiles. We observe that the concentration of reported income is unaffected by re-ranking, which suggests that our allocation does not create mechanical bias by imposing unrealistic structure on reranking. We also observed that the distribution of unreported income is impacted: with a larger portion of unreported income allocated further up the distribution when ranking by exam-corrected income. The second panel plots the top 1 percent share of reported income ranked by reported (solid black) or exam-corrected (gray) income, using DINA data. The right-hand-side axis plots the difference between the two. The gap between the two is small (less than 0.2 percentage points in all periods), and shows some variability over time.

Figure C.2: Reported non-wage income shares of total fiscal income



Notes: The data from these figures is derived from supplemental data provided by Auten and Splinter (2024) on NIPA totals by year and type of income. Overall, non-wage business income as a share of fiscal income fell between 1960 and 1980, driven mainly in a reduction in sole proprietors' income. Following that period, partnership and s-corporate income shares grow substantially.

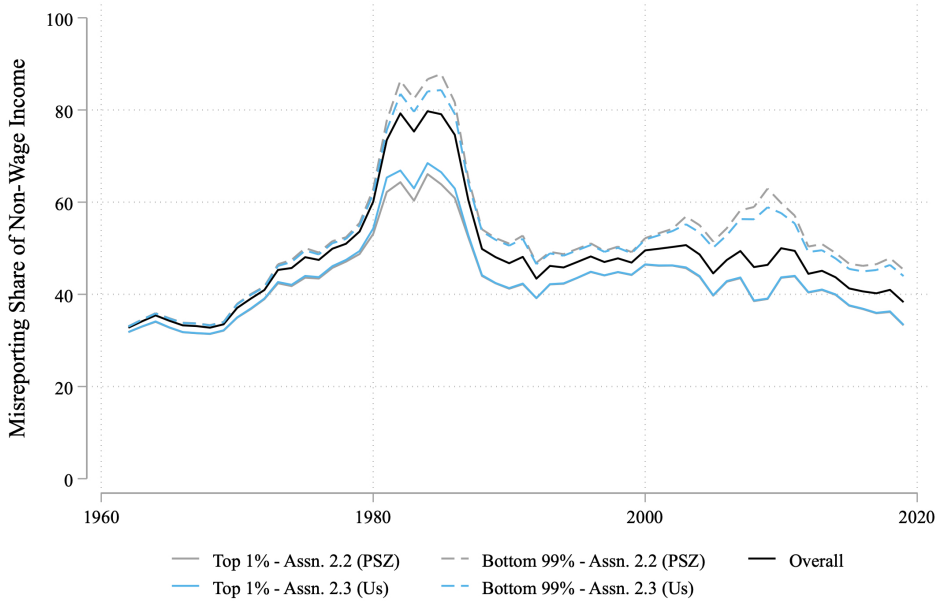
Figure C.3: Effect of misreporting on the top one percent share of fiscal income and pre-tax National Income



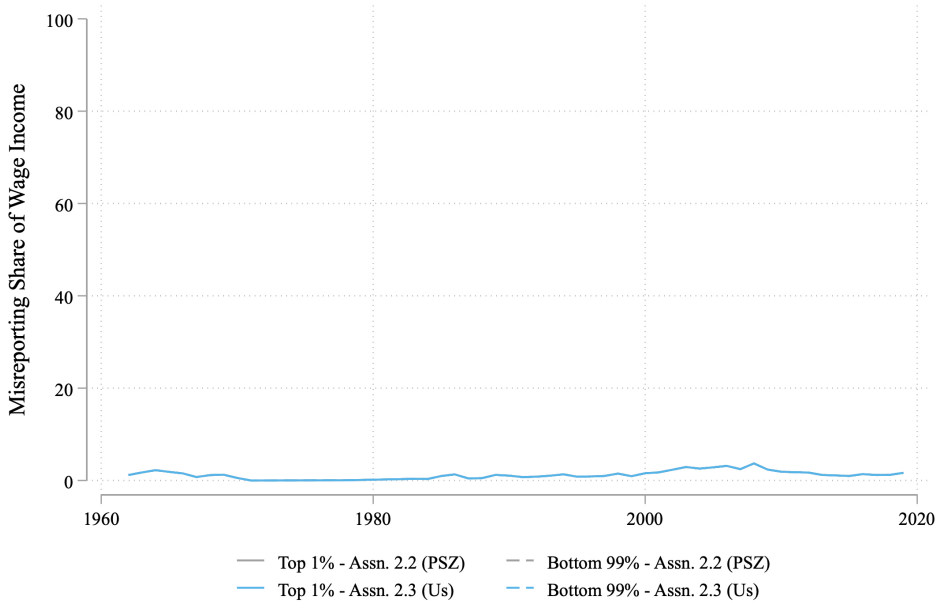
Notes: This figure plots the difference between the top 1 percent share of fiscal income (solid lines) or pre-tax National Income (dashed lines) excluding capital gains and the estimates of Y_{1t}^{c} derived from assumptions 2.1 (AS) and 2.2 (PSZ).

Figure C.4: Misreporting share of Income by Type, Assumption, and Place in Distribution

(a) Non-Wage Income



(b) Wage Income



Notes: This figure plots the misreporting share of wage and non-wage income for the top 1 percent and bottom 99 percent of the fiscal income distribution, under Assumption 2.2 and 2.3. We do not have separate estimates of E^c_{1jt} under Assumption 2.1.

C.2 Additional Details on Auten and Splinter's Methodology

Following PS, AS begins with a set of annual stratified random samples of tax returns, from which the distribution of income in tax data (referred to as fiscal income) is readily observable at the tax unit level. From this starting point, they employ a process of adjustments, additions, and subtractions to produce a measure of income for each household that is consistent with the income concepts reported in NIPA. The whole process employed by AS can be broken down into five stages, detailed in AS's appendix table B2.

Step 1: Improving fiscal income. This step involves adjustments to the IRS samples that can be completed using information from the IRS. Some are adjustments to the sample construction to account for the difference between tax units and households[1] (e.g. removing filers under 20 or who are dependents). Others add or remove types of income to and from Adjusted Gross Income (AGI) to bring fiscal income more in line with the BEA income concepts (e.g. adding tax exempt interest and excluded dividends, removing capital gains).

Step 2: Expansions to pre-tax income. Several components of pre-tax income are not directly observed on tax returns. The adjustments in this stage aim to allocate income from such components to individual tax returns based on related information reported on tax returns (for example, allocating imputed rent in proportion to deducted real-estate taxes). This subsection includes the allocation of under-reported income and excess depreciation/LEBI, which is the focus of this

[1] Tax units are groups of individuals who report their income on the same tax return, or who would have done so if they are nonfilers. Dependents living in separate households can be part of the same tax unit, while unmarried cohabiting adults can be in the same household and different tax units. We do not engage here with the question of accounting for the difference between tax units and households, which is the source of a quantitatively minor disagreement between the estimates.

comment.

Step 3: Adding transfers to pre-tax income. Adjustments in this stage allocate government transfers, such as social security benefits, unemployment insurance benefits, and the value of health insurance, to tax units based on either reported benefits or demographic characteristics.

Step 4: Calculating after-tax income. In this stage, the value of federal, state, and local taxes are subtracted from the income of tax units.

Step 5: Adding value from government sector to after-tax income. Finally, the value of government deficits and surpluses, as well as government consumption, is added to individual income.

We are interested in the expansion to pre-tax income in stage 2 from the inclusion of misreported income, which is income that should appear on a tax return but is not present. This can be due either to intentional evasion or to misunderstanding, ambiguities in tax rules, or other less nefarious causes. In determining how to account for this gap in the calculation of the distribution of income, there are two critical questions: how big is the gap, and how should the missing dollars be distributed among households?

AS start by calculating the gap between NIPA- and tax-reported total income for each of five different types of income: wages and salaries, rental income, S-corporation income, farm income, and nonfarm proprietor income. For income type j in tax year t, the gap is calculated as:

$$G_{jt} = Y_{jt}^{NIPA} - \sum_{i} y_{ijt},$$

where Y_{jt}^{NIPA} is the NIPA total in year t for income type j, and $\sum_i y_{ijt}$ is the total income amount

observed on tax returns (i) for the same income type and year, after adjustments made in stage 1 or 2.[2] With one exception we describe shortly, their approach assumes that these differences are entirely driven by misreporting, and that the gap between NIPA and tax-reported income is an accurate estimate of under-reporting.

Two complicating factors are worth noting. First, NIPA does not rely on tax data to estimate total income from rents, S corporations, or farms. Second, NIPA does include explicit adjustments for under-reported income for wages and salary income and non-farm proprietor income, using figures it obtains from the IRS based on official Tax Gap statistics (including DCE).[3] However, for non-farm proprietor income, the overall gap G_t is larger than the explicit adjustment for unreported income due to the additional adjustment NIPA makes for excess depreciation/capital consumption.

For most of the income groups (wages, rents, s-corporations, and farm income), the gap G_{jt} is assumed to be entirely composed of misreporting (M_{jt}). To address the second complication above, for non-farm proprietor income ($j = nfp$), the gap is broken down into three categories: misreporting, (M_{jt}^{NIPA}), adjustments for excess depreciation (D_{jt}^{NIPA}), both obtained directly

[2]Throughout this subsection, all values directly reported by NIPA are indicated with the superscript "NIPA;" other variables are calculated/estimated by AS.

[3]In the NRP era, DCE is based on the premise that the probability of observing a given dollar of income is less than one, but that the counterfactual where the probability is equal to one can be approximated using the variation in auditor effectiveness. If there are some group of auditors that are 100 percent effective in finding misreported income, and those auditors are assigned to tax units in some way uncorrelated with the probability of misreporting, then it is possible to construct the counterfactual where all misreported income is found. There are a number of critiques of this model - namely that auditors are not 100 percent effective and are not randomly assigned - that argue that the DCE methodology is either an under or over estimate of the misreporting gap. In the TCMP, methdods for estimating undetected under-reporting were more ad hoc, see Erard and Feinstein (2011).

from NIPA, and a residual factor. In other words:

$$G_{jt} = \begin{cases} M_{jt} & \text{if } j \neq nfp \\ M_{jt}^{NIPA} + D_{jt}^{NIPA} + R_{jt} & \text{if } j = nfp \end{cases}$$

We focus on how AS approach the distribution of the total misreporting gap,

$$M_t = \sum_j M_{jt},$$

where $M_{nfp,t}^{NIPA} = M_{nfp,t}$. This implies that the full gap between total NIPA total income and tax income is:

$$G_t = M_t + D_{nfp,t}^{NIPA} + R_{nfp,t}.$$

The following mainly deals with the estimation of M_t, while $D_{nfp,t}^{NIPA}$ and $R_{nfp,t}$ are discussed at the end of this subsection.

Given this estimate of the overall misreporting gap, AS decompose this gap into several components:

$$M_t = (M_t^D + M_t^U + M_t^{NF}) * S_t,$$

where $M_t^{NF} = 0.05 * M_t$ is the portion (five percent) of the gap allocated to non-filers, M_t^D is the portion of misreporting that can be detected via random audit data, M_t^U is the share that is undetected but inferred via DCE, and S is a scaling factor to preserve the equality between the left and right hand sides of the equation. The distributional implications of the misreporting gap is derived from estimates and assumptions about how these difference components are distributed.

To estimate both the size and distribution of detected misreporting (M_t^D), AS employed

the methodology proposed by (Auten and Langetieg, 2023, henceforth AL) that uses information from random audit studies – the TCMP through 1988 and the NRP from 2001 through 2013 – to produce estimates of detected misreporting by year and reported income. AL tabulate the ratio of corrected income to reported income by year, rank of reported income, and - importantly - ratio groups, which are bins of the ratio of corrected to reported income. For each year and reported income percentile bin, AL report the share of tax units in each ratio group and summary statistics of the tax-unit-level ratio within that group. For example, one of the ratio groups contains observations with a ratio of corrected income to reported income of between 2 and 4. In 1988, AL estimate that 0.52 percent of tax returns between the 40th and 60th percentile of reported income were in this ratio group. For these returns in this ratio group, the average ratio is 2.56, and the "standard error for ratio" is 0.027 – whether this refers to the standard error of their estimate of the within-cell mean ratio or the estimated within-cell standard deviation of the ratio is unclear; the next step requires the latter but we cannot rule out an error here. The AL estimates contain many cell-level estimates of this kind. These estimates are the parameters of the micro-simulation we turn to next.

The steps of the micro-simulation used to distribute M_{jt}^{D} begins with tax-return level data, and does the following:

1. For each year, rank observations by reported income, and classify observations by their binned income percentile (p), where the bins match those reported by AL.

2. Randomly allocate observations within income percentiles to ratio groups (r) in proportion to the share of observations in that ratio group estimated by AL in the most recent wave of random audit data. Now each observation has been assigned an income percentile and ratio

group.

3. Within each ratio group by income percentile by year cell, each observation is assigned a multiplier based on a random draw from that cell's ratio distribution, which is derived from the cell-level parameter estimates in AL.[4] ($m_{it}^D = m^D(p, r, \epsilon)$)

4. Each tax unit's detected-misreporting-corrected income is $y_{it} * m^D(p, r, \epsilon)$, so $M_{it}^D = y_{it} * (m^D(p, r, \epsilon) - 1)$

The total detected misreporting gap is the sum of the individual-level estimates of the difference between reported and corrected income:

$$M_t^D = \sum_i M_{it}^D.$$

After simulating values for detected misreporting, AS follow the process outlined in Auten and Splinter (2021) to simulate the size and distribution of DCE-identified undetected misreporting M_t^U. Their parametric specification, referred to as distributionally consistent gradient multipliers (DCGM), is proposed as a correction to the process used to distribute DCE adjustments in Johns and Slemrod (2010) and Mazur and Plumley (2007). This method entails multiplying detected under-reporting at the individual level M_{it}^D by a multiplier $m_{it}^{U,DCE}$ to allocate the additional under-reporting implied by DCE:

$$M_{it}^{U,DCE} = M_{it}^D * (m_{it}^{U,DCE} - 1) = y_{it} * (m^D(p, r, \epsilon) - 1) * (m^{U,DCE} - 1)$$

[4] We are unable to locate a description of the specification of the assumed distribution in AS, AL, or appendices, apart from the fact that the distribution of the ratio in each cell is "bounded by max and min ratios, and the top group with a maximum of 125 percent of the cell mean."

With the original approach of Johns and Slemrod (2010), the multiplier $m_{it}^{U,DCE}$ is essentially a constant calibrated from a different model in which auditor effects are actually estimated using 2001 NRP data.[5] With DCGM, AS allow the multiplier to vary with the ratio class (r) assigned above: $m^{U,DCE} = m^U(r)$.[6] They claim that the DCGM approach results in a more realistic distribution, because returns with large detected under-reporting would be more likely to have been assigned more effective auditors (so returns with higher r should get a lower multiplier $m^U(r)$). We are unable to ascertain exactly how AS specify the DSGM parameters ($m^U(r)$), but we discuss the way they validate their specified multipliers below. Unlike $m^D(p, r, e)$, $m^U(r)$ does not vary at the taxpayer level within a ratio class.

Using these ratio-specific multipliers, AS produce simulated values of each tax unit's undetected misreporting following the expression above. This is summed over all tax units to obtain the overall undetected share of misreporting.

$$M_t^U = \sum_i M_{it}^U$$

The last piece of this process is to add in the 5 percent of the gap assumed to be held by non-filers and scale the resulting estimate of the misreporting gap to match the observed gap between the tax and NIPA data.

Now we turn to the two additional components of the full NIPA gap and adjusted fiscal income, $D_{nfp,t}^{NIPA}$ and $R_{nfp,t}$. The latter is simply a residual, to make sure that the equation

[5]The Johns and Slemrod (2010) multiplier method, which was also used in the first draft of Guyton et al. (2023b), scales detected under-reporting by one of four constants depending on the type of income and whether the taxpayer was self-employed. See Guyton et al. (2023a) for further discussion; these authors describe this approach as DCE2001 because the multipliers are calibrated based on 2001 NRP data.

[6]To limit the effect of NOL carryovers, no multiplier is applied to tax units with negative AGI.

$G_{nfp,t} = M^{NIPA}_{nfp,t} + D^{NIPA}_{nfp,t} + R_{nfp,t}$ holds. The former is designed to account for the difference between economic depreciation and the tax deduction for depreciation. As discussed in the introduction, national accounts include an annual capital consumption adjustment for proprietor income to account for this difference; AS allocate the reported annual adjustment using two methodologies, depending on the period in question. For years prior to 1980, adjustments are allocated proportional to reported proprietor income. After 1980, first, 85 percent of reported expensing of sole proprietorships (and not partnerships) is added back to tax returns, then the remainder of the NIPA capital consumption adjustment is allocated proportionally to reported depreciation deductions of sole proprietorships (and not partnerships). The use of tax return variables pertaining to sole proprietor income to do this allocation rather than using sole proprietor and partnership income combined likely generates the bias described at the end of the introduction, which we do not discuss further in this comment.

C.3 Treatment of negative incomes and changes to the Tax System due to 1980s Reforms

C.3.1 Substitution Across Forms

Smith et al. (2019) estimate that about 30% of the increase in pass-through income over time is driven by shifting in corporate forms from C corporations to pass-through businesses, which became tax-preferred for many businesses after 1986 (see also Gordon and Slemrod (2002)). When we switch to a National Income concept in the the final subsection of the comment, our approach assumes C corporation misreporting is accurately accounted for in NIPA totals and accurately distributed by whichever method we use to distribute non-fiscal income (we consider

both the method of PSZ and AS). Under this assumption, if shifting away from C corporation forms decreases misreporting in C corporations, one should expect to find a corresponding increase in pass-through misreporting and therefore that misreporting as a share of fiscal income would increase, consistent with what we impose under Assumption 2.3. However, we acknowledge some quantitative uncertainty on this issue: it is not clear that Assumption 2.3 perfectly quantifies the increase in pass-through misreporting that we expect to see when some C corp misreporting becomes pass-through misreporting.This question contributes to the overall uncertainty involved in quantifying the distribution of misreporting with limited data, as discussed at the end of section 3.

C.3.2 Partnership Loss Shelters in the 1980s

Figure C.4 displays some odd dynamics in the mid-1980s and the same is true of some of the sensitivity analysis that is our focus in the next part of this Appendix. Here we briefly review what we know about the likely cause of this anomaly.

The early 1980s saw rapid growth of the use of partnership losses for tax sheltering that were shut down by the introduction of passive loss limitation rules, and perhaps a few other provisions that muted the benefits of the shelters, in the Tax Reform Act of 1986 (TRA86). See Samwick (1995) for a more detailed review of the strategies involved and how TRA86 shut them down.

The aggregate implications of these developments are that reported partnership losses grow substantially in the first part of the 1980s and then fall precipitously after 1986. What does this mean for our anlaysis? First, we note that these losses were not considered non-compliance

according to Tax Gap estimates (see e.g. the TCMP-based estimates for partnerships from 1985 and 1988 in Internal Revenue Service (1996)). The corresponding adjustments for misreporting in NIPA therefore do not include income that offsets these losses. When we estimate misreporting as a share of corrected income for non-wage business income in Figure C.4, therefore, the increase in reported losses shrinks the denominator of the statistic and increases the implied misreporting shares in the early 1980s. Then the misreporting shares fall when the losses disappear after 1986. As a result of this, how we treat negative incomes in the implementation of our distributional neutrality assumption (Assumption 2.3) matters significantly in these years, while, as we are about to show, it is a minor issue in most other years.

The implications of these mid-1980s partnership losses for the estimation of the distribution of National Income generally is murky. The discussion of the relevant tax sheltering strategies in Samwick (1995) suggests that these losses mainly reflect the re-timing of income to exploit deferral: the losses were typically investment expenses that would be offset by gains, and taxpayers abusing the shelters were accelerating losses and deferring gains to later years. In that case, it is unclear how we should treat these losses in estimating annual economic income per NIPA totals, let alone distributional figures. To our knowledge, there is no specific adjustment in NIPA totals to address that these partnership losses may not reflect real economic losses in the year in which they are reported on tax returns.

The fact that negative business incomes are often associated with tax sheltering, especially in the mid-1980s, is closely related to the rationale PSZ provide for excluding reported business losses from their definition of income entirely. This is a somewhat ad hoc approach, but it prevents factors like the explosion of partnership losses in the mid-1980s from exerting much influence on estimated inequality, and this is closely related to why, as we see below, estimates

using positive income per Assumption 2.3 appear appear to be better behaved in the 1980s than estimates using perfect distributional neutrality by type of income. At the same time, we recommend caution in interpreting all authors' estimates of the Distribution of National Income around 1986 in light of this issue. The sharp changes in time series based on reported incomes after 1986 is often attributed to income-shifting between C corporation and pass-through business forms, as discussed in the previous subsection. Such income-shifting also drives some minor differences between PSZ and AS unrelated to misreporting. But to some extent the sharp changes in time series we observe after 1986 are also driven by the disappearance of these losses claimed by partnerships. We acknowledge some quantitative uncertainty around what this should imply for estimates of National Income and its distribution in the mid-1980s. Fortunately, this does not seem to be relevant for the main argument of our comment.

C.3.3 Sensitivity Analysis for Assumption 2.3

Here, we review some alternative approaches to the the treatment of negative business income in the allocation of misreporting.

Figure C.5a illustrates the three series from Figure 3.1a alongside some alternatives in which we implement an alternative version of Assumption 2.3, which recall is the following:

Assumption 2.3 (Our Preferred Approach)

$$E^c_{1jt} = \frac{\sum_{i \in \text{[Top 1 percent by r]}} \max[r_{ijt}, 0]}{\sum_i \max[r_{ijt}, 0]}.$$

We first note that if r_{ijt} were everywhere positive, the above would be equivalent to $E^c_{1jt} = R^r_{1jt}$, and under Assumption 1 this further implies $E^c_{1jt} = R^c_{1jt}$. This is the condition we called

"perfect distributional neutrality by type of income" in the main text: that the top 1% share of income of type j should be completely unaffected by adding misreporting to income. If we assume perfect distributional neutrality by type of income when r_{ijt} can be negative, however, then whenever we find an increase in negative income in some part of the distribution, we implicitly add some negative misreporting to that point in the distribution with a proportional increase in positive misreporting elsewhere.

The series in Figure C.5 labeled "including negatives" illustrates the top 1% share of misreporting E^c_{1j} if we assume perfect distributional neutrality despite the fact that reported income can be negative. As the preceding logic suggests, we find a larger increase in the top 1% share of misreporting under perfect distributional neutrality, which will lead to a larger increase in the top 1% income share overall. More subtly, we note that the increase in reported partnership losses before 1986 discussed in the previous subsection causes a spike in the top 1% share of misreporting prior to 1986 in the this series.

Our approach to handling negative incomes is thus more conservative for the increase in the top 1% income share over time than what we would find under perfect distributional neutrality by type of income. But this does not eliminate all potential concerns with Assumption 2.3, especially if we extend our approach to examine the full distribution of income instead of just focusing on the top 1% versus the bottom 99%. To appreciate this, consider that in the bottom part of the distribution, income is negative in the aggregate. Under perfect distributional neutrality, we would allocate negative misreporting to this part of the distribution, and under Assumption 2.3, we would allocate zero misreporting. The random audit data suggest that both of these would allocate too little under-reporting to those with negative (corrected) incomes. Guyton et al. (2023b) Table A2 reports that about 1.4% of audit-detected mis-reporting belongs to the bottom 10% of

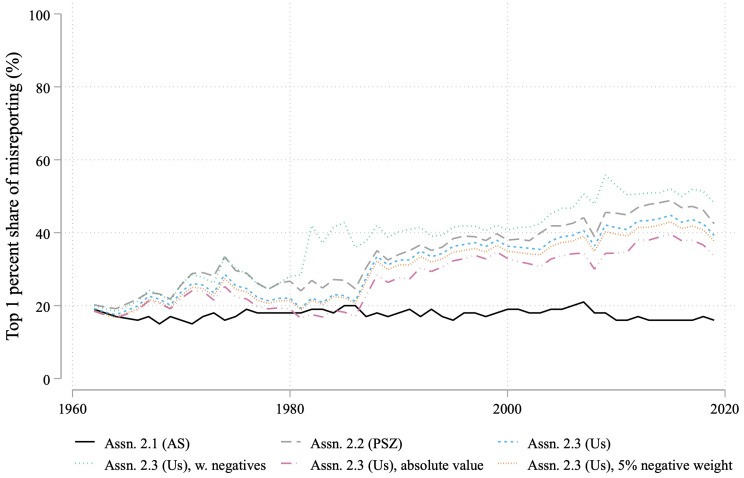
the corrected fiscal income distribution, where income is negative in the aggregate. We therefore illustrate in our sensitivity analysis two more specifications where we allocate some misreporting to those with negative incomes. In one alternative, we use the absolute value of reported income $|r_{ijt}|$ in place of positive reported income in Assumption 2.3. In the other, we allocate 5% of misreporting to the those with negative reported income of a given type (a figure conservatively calibrated to the 1.4% figure alluded to above) and allocate the rest in proportion to positive income as in Assumption 2.3. In these specifications, the top 1% share of misreported income decreases, because we are allocating more misreporting to the bottom of the corrected income distribution in either case. But the quantiative effects of these alterations to the specification are very small compared to the difference between Assumption 2.3 and AS' approach in Assumption 2.1. Moreover, with the exception of the pre-1986 issue arising under full distributional neutrality, we observe very similar dynamics in all of the alternatives we consider here, suggesting this question about how we should address negative business incomes will mainly matter for the overall level of the top 1% income share, rather than how it changes over time.

This sensitivity analysis is closely related to a point AS make in criticism of PSZ's method and an earlier version of this comment. Based on the sensitivity analysis here, we argue that the way we decided to handle negative incomes in Assumption 2.3 was reasonably conservative and the remaining uncertainty is small for the central question here about how the top 1% share changes over time. AS observe that in random audit data, substantial under-reporting of business income is attributable to those with negative reported incomes (see e.g. Auten and Langetieg, 2023). The main reason for this is that the largest corrections to business income in the random audit data reflect disallowed losses due to loss limitation rules (an individual rather than entity-level correction for pass-throughs). When such losses are disallowed, importantly, the taxpayer's

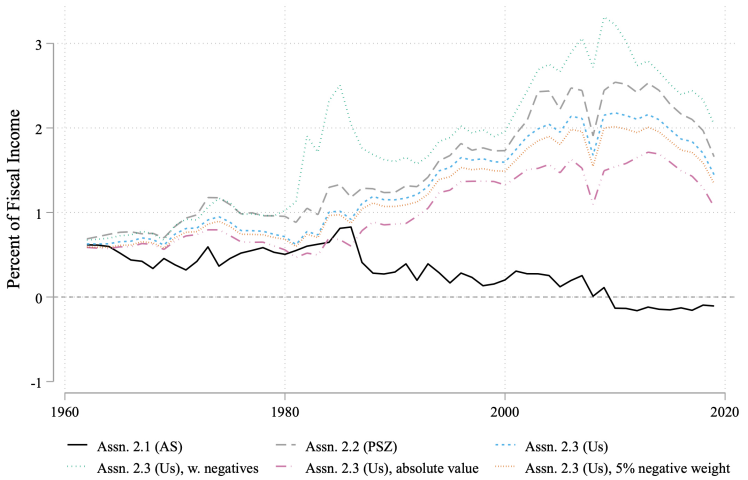
corrected income is typically *positive*, and disallowing losses can cause many taxpayers with negative reported incomes to re-rank all the way into the top 1% by corrected income (see also a related illustration in Guyton et al., 2023b, Figure A11). This explains why a large share of under-reporting belongs to those with negative reported incomes in Auten and Langetieg (2023), but a very small share belongs to those with negative corrected incomes in the numbers from Guyton et al. (2023b) referenced above. In Table A2 of Guyton et al. (2023b), we observe that there is also an effect due to those with negative misreporting moving downward in the distribution when re-ranking from reported to corrected income. Naturally, how this information in these authors' Table A2 would change after we account for undetected misreporting is unclear. But in the end, the available evidence suggests allocating little to no under-reporting to those with negative corrected incomes, which is what we require under Assumption 2.3, is much more realistic than assigning no under-reporting to those with negative reported incomes, which is what AS suggested that we do in their response to the first draft of this comment. As such, we argue that the way we handle negative incomes in implementing our method is a quantitatively minor source of bias/uncertainty. By very similar logic, our analysis here suggests that AS' criticism of PSZ's method overstates the bias from allocating no under-reporting to those with reported losses and thereby overstates the quantitative importance of the excess re-ranking effect discussed in the main text.

Figure C.5: Sensitively to treatment of negatives incomes

(a) Share of misreporting allocated to the top 1 percent



(b) Effect of misreporting on the top 1% share of fiscal income



Notes: Panel 1 re-creates Figure 3.1a, adding several robustness checks regarding the treatment of negative non-wage income. Panel 2 re-creates Figure 3.1b, adding the same series of robustness checks. The solid black line plots the share of total (detected and un-detected) misreporting allocated to the top 1 percent of the true income distribution in Auten and Splinter (2025a). The dashed gray line reflects the actual assumptions made by PSZ: namely, allocating misreporting proportional to reported positive income then re-ranking by corrected fiscal income (Piketty et al., 2018). The dashed blue line shows the effect of eliminating the re-ranking effect, allocating misreporting proportional to reported positive income without re-ranking. The faint dotted green line shows the effect of allocating misreporting proportional to reported total (positive and negative income without re-ranking. It rises due to the increase in negative non-wage income in the early 1980s, which artificially increases the amount of misreporting allocated to the top 1 percent. The dot-dash purple line shows effect of allocating misreporting proportional to the absolute value of income without re-ranking. Finally, dotted orange line allocates 95 percent of misreporting proportional to reported positive income and 5 percent to reported negative incomes (by type of income).

Bibliography

Jacob Bastian and Lance Lochner. The Earned Income Tax Credit and Maternal Time Use: More Time Working and Less Time with Kids? *Journal of Labor Economics*, 40(3):573–611, July 2022. ISSN 0734-306X, 1537-5307. doi: 10.1086/717729. URL https://www.journals.uchicago.edu/doi/10.1086/717729.

James Mirrlees. An Exploration in the Theory of Optimum Income Taxation. *The Review of Economic Studies*, 38(2):175–208, 1971. doi: https://doi.org/10.2307/2296779. URL https://www.jstor.org/stable/2296779.

Emmanuel Saez. Optimal Income Transfer Programs: Intensive verses Extensive Labor Supply Responses. *Quarterly Journal of Economics*, 117(3):1039–1073, 2002.

Emmanuel Saez, Joel Slemrod, and Seth H Giertz. The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review. *Journal of Economic Literature*, 50(1):3–50, March 2012. ISSN 0022-0515. doi: 10.1257/jel.50.1.3. URL https://pubs.aeaweb.org/doi/10.1257/jel.50.1.3.

Nada Eissa and Jeffrey B. Liebman. Labor Supply Response to the Earned Income Tax Credit. *Quarterly Journal of Economics*, 111(2):605–637, May 1996. URL https://www.jstor.org/stable/2946689.

Bruce D. Meyer and Dan T. Rosenbaum. Welfare, the Earned Income Tax Credit, and the Labor Supply of Single Mothers. *Quarterly Journal of Economics*, 116(3):1063–1114, 2001. URL https://www.jstor.org/stable/2696426.

Hilary W. Hoynes and Ankur J. Patel. Effective Policy for Reducing Poverty and Inequality?: The Earned Income Tax Credit and the Distribution of Income. *Journal of Human Resources*, 53 (4):859–890, 2018. ISSN 0022-166X, 1548-8004. doi: 10.3368/jhr.53.4.1115.7494R1. URL http://jhr.uwpress.org/lookup/doi/10.3368/jhr.53.4.1115.7494R1.

Jacob E. Bastian and Maggie R. Jones. Do EITC expansions pay for themselves? Effects on tax revenue and government transfers. *Journal of Public Economics*, 196:104355, April 2021. ISSN 00472727. doi: 10.1016/j.jpubeco.2020.104355. URL https://linkinghub.elsevier.com/retrieve/pii/S004727272030219X.

Diane Whitmore Schanzenbach and Michael R. Strain. Employment Effects of the Earned Income Tax Credit: Taking the Long View. *Tax Policy and the Economy*, 35:87–129, 2021. URL https://doi.org/10.1086/713494.

Brandon Enriquez, Damon Jones, and Ernie Tedeschi. Short-Term Labor Supply Response to the Expanded Child Tax Credit. *AEA Papers and Proceedings*, 113:401–405, May 2023. ISSN 2574-0768, 2574-0776. doi: 10.1257/pandp.20231087. URL https://pubs.aeaweb.org/doi/10.1257/pandp.20231087.

Jessica Pac and Lawrence M. Berger. Quasi-experimental evidence on the employment effects of the 2021 fully refundable monthly child tax credit. *Journal of Policy Analysis and Management*, 43(1):192–213, January 2024. ISSN 0276-8739, 1520-6688. doi: 10.1002/pam.22528. URL https://onlinelibrary.wiley.com/doi/10.1002/pam.22528.

Elizabeth Ananat, Benjamin Glasner, Christal Hamilton, Zachary Parolin, and Clemente Pignatti. Effects of the expanded Child Tax Credit on employment outcomes. *Journal of Public Economics*, 238:105168, October 2024. ISSN 00472727. doi: 10.1016/j.jpubeco.2024.105168. URL https://linkinghub.elsevier.com/retrieve/pii/S004727272400104X.

Eva Vivalt, Elizabeth Rhodes, Alexander W. Bartik, David E. Broockman, and Sarah Miller. The Employment Effects of a Guaranteed Income: Experimental Evidence from Two US States. Technical Report w32719, National Bureau of Economic Research, Cambridge, MA, July 2024. URL https://www.nber.org/papers/w32719.

Sidhya Balakrishnan, Roberta Costa, Johannes Haushofer, and Fábio Waltenberg. Welfare Effects of a Permanent Unconditional Cash Transfer Program: Evidence from Maricá, Brazil. Technical Report w33089, National Bureau of Economic Research, Cambridge, MA, October 2024. URL http://www.nber.org/papers/w33089.pdf.

Jacob Goldin, Tatiana Homonoff, Neel A Lal, Ithai Lurie, and Katherine Michelmore. Child Tax Benefits and Labor Supply: Evidence from California. Technical Report Working Paper 32343, National Bureau of Economic Research, Cambridge, MA, 2024.

Ufuk Akcigit, Salomé Baslandze, and Stefanie Stantcheva. Taxation and the International Mobility of Inventors. *American Economic Review*, 106(10):2930–2981, October 2016. ISSN 0002-8282. doi: 10.1257/aer.20150237. URL https://pubs.aeaweb.org/doi/10.1257/aer.20150237.

Enrico Moretti and Daniel J. Wilson. The Effect of State Taxes on the Geographical Location of Top Earners: Evidence from Star Scientists. *American Economic Review*, 107(7):1858–1903, July 2017. ISSN 0002-8282. doi: 10.1257/aer.20150508. URL https://pubs.aeaweb.org/doi/10.1257/aer.20150508.

Henrik Jacobsen Kleven, Camille Landais, and Emmanuel Saez. Taxation and International Migration of Superstars: Evidence from the European Football Market. *American Economic Review*, 103(5):1892–1924, August 2013. ISSN 0002-8282. doi: 10.1257/aer.103.5.1892. URL https://pubs.aeaweb.org/doi/10.1257/aer.103.5.1892.

Cristobal Young and Charles Varner. MILLIONAIRE MIGRATION AND STATE TAXATION OF TOP INCOMES: EVIDENCE FROM A NATURAL EXPERIMENT. *National Tax Journal*, 64(2.1):255–283, June 2011. ISSN 0028-0283, 1944-7477. doi: 10.17310/ntj.2011.2.02. URL https://www.journals.uchicago.edu/doi/10.17310/ntj.2011.2.02.

Henrik Jacobsen Kleven, Camille Landais, Emmanuel Saez, and Esben Schultz. Migration and Wage Effects of Taxing Top Earners: Evidence from the Foreigners' Tax Scheme in Denmark*. *The Quarterly Journal of Economics*, 129(1):333–378, February 2014. ISSN 0033-5533, 1531-4650. doi: 10.1093/qje/qjt033. URL https://academic.oup.com/qje/article/129/1/333/1899540.

David R. Agrawal and Dirk Foremny. Relocation of the Rich: Migration in Response to Top Tax Rate Changes from Spanish Reforms. *The Review of Economics and Statistics*, 101(2):214–232, May 2019. ISSN 0034-6535, 1530-9142. doi: 10.1162/rest_a_00764. URL https://direct.mit.edu/rest/article/101/2/214-232/58521.

Enrico Moretti and Daniel J. Wilson. Taxing Billionaires: Estate Taxes and the Geographical Location of the Ultra-Wealthy. *American Economic Journal: Economic Policy*, 15(2):424–466, May 2023. ISSN 1945-7731, 1945-774X. doi: 10.1257/pol.20200685. URL https://pubs.aeaweb.org/doi/10.1257/pol.20200685.

Karen Smith Conway and Jonathan C. Rork. How Has Elderly Migration Changed in the Twenty-First Century? What the Data Can—and Cannot—Tell Us. *Demography*, 53(4):1011–1025, August 2016. ISSN 0070-3370, 1533-7790. doi: 10.1007/s13524-016-0477-7. URL http://link.springer.com/10.1007/s13524-016-0477-7.

Thomas Piketty, Emmanuel Saez, and Gabriel Zucman. Distributional National Accounts: Methods and Estimates for the United State. *The Quarterly Journal of Economics*, 133(2):553–609, May 2018. ISSN 0033-5533, 1531-4650. doi: 10.1093/qje/qjx043. URL https://academic.oup.com/qje/article/133/2/553/4430651.

Gerald Auten and David Splinter. Income Inequality in the United States: Using Tax Data to Measure Long-Term Trends. *Journal of Political Economy*, 132(7):2179–2227, 2024. ISSN 0022-3808, 1537-534X. doi: 10.1086/728741. URL https://www.journals.uchicago.edu/doi/10.1086/728741.

National Academies of Science. *A Roadmap to Reducing Child Poverty*. National Academies Press, Washington, D.C., August 2019. ISBN 978-0-309-48398-8. doi: 10.17226/25246. URL https://www.nap.edu/catalog/25246. Pages: 25246.

Marianne E Page. New Advances on an Old Question: Does Money Matter for Children's Outcomes? *Journal of Economic Literature*, 62(3):891–947, 2024. doi: https://doi.org/10.1257/jel.20231553.

Kim S Rueben, Frank Sammartino, and Kirk J Stark. Upward Mobility and State-Level EITCs: Evaluating California's Earned Income Tax Credit. *TAX LAW REVIEW*, 70, 2017.

Henrik Kleven. The EITC and the Extensive Margin: A Reappraisal. *Journal of Public Economics*, Forthcoming, 2024.

Dmitry Arkhangelsky, Susan Athey, David A. Hirshberg, Guido W. Imbens, and Stefan Wager. Synthetic Difference-in-Difference. *American Economic Review*, 111(12):4088–4118, 2021. doi: 10.1257/aer.20190159.

Pierre-André Chiappori. Collective Labor Supply and Welfare. *Journal of Political Economy*, 100 (3):437–467, June 1992. ISSN 0022-3808, 1537-534X. doi: 10.1086/261825. URL https://www.journals.uchicago.edu/doi/10.1086/261825.

Richard Blundell, Pierre-André Chiappori, and Costas Meghir. Collective Labor Supply with Children. *Journal of Political Economy*, 113(6):1277–1306, December 2005. ISSN 0022-3808, 1537-534X. doi: 10.1086/491589. URL https://www.journals.uchicago.edu/doi/10.1086/491589.

Tuomas Kosonen and Tuomas Matikka. Discrete Labor Supply: Quasi-Experimental Evidence and Implications, 2023.

Claus Thustrup Kreiner, Jakob Roland Munch, and Hans Jørgen Whitta-Jacobsen. Taxation and the long run allocation of labor: Theory and Danish evidence. *Journal of Public Economics*, 127:74–86, July 2015. ISSN 00472727. doi: 10.1016/j.jpubeco.2014.03.012. URL https://linkinghub.elsevier.com/retrieve/pii/S0047272714000619.

Nathaniel Hendren and Ben Sprung-Keyser. A Unified Welfare Analysis of Government Policies*. *The Quarterly Journal of Economics*, 135(3):1209–1318, August 2020. ISSN 0033-5533, 1531-4650. doi: 10.1093/qje/qjaa006. URL https://academic.oup.com/qje/article/135/3/1209/5781614.

Jacob Bastian and Katherine Michelmore. The Long-Term Impact of the Earned Income Tax Credit on Children's Education and Employment Outcomes. *Journal of Labor Economics*, 36(4):841–1163, October 2018. doi: https://doi-org.proxy-um.researchport.umd.edu/10.1086/697477.

David Neumark and Zeyu Li. Too Much of a Good Thing? How Narrow Targeting and Policy Interactions Influence Responses to California's EITC. Technical Report w32883, National Bureau of Economic Research, Cambridge, MA, August 2024. URL http://www.nber.org/papers/w32883.pdf.

John Iselin, Taylor Mackay, and Matthew Unrath. Measuring take-up of the California EITC with state administrative data. *Journal of Public Economics*, 227:105002, November 2023. ISSN 00472727. doi: 10.1016/j.jpubeco.2023.105002. URL https://linkinghub.elsevier.com/retrieve/pii/S0047272723001846.

Matthew Unrath. Filing, Employment, and Earnings Responses to California's Young Child Tax Credit. *Working Paper*, 2024.

Emmanuel Saez. Do Taxpayers Bunch at Kink Points? *American Economic Journal: Economic Policy*, 2(3):180–212, August 2010. ISSN 1945-7731, 1945-774X. doi: 10.1257/pol.2.3.180. URL https://pubs.aeaweb.org/doi/10.1257/pol.2.3.180.

R. Chetty, J. N. Friedman, T. Olsen, and L. Pistaferri. Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records. *The Quarterly Journal of Economics*, 126(2):749–804, May 2011. ISSN 0033-5533, 1531-4650. doi: 10.1093/qje/qjr013. URL https://academic.oup.com/qje/article-lookup/doi/10.1093/qje/qjr013.

Raj Chetty. Bounds on Elasticities With Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply. *Econometrica*, 80(3):969–1018, 2012. ISSN 0012-9682. doi: 10.3982/ECTA9043. URL http://doi.wiley.com/10.3982/ECTA9043.

Henrik Jacobsen Kleven and Esben Anton Schultz. Estimating Taxable Income Responses Using Danish Tax Reforms. *American Economic Journal: Economic Policy*, 6(4):271–301, November 2014. ISSN 1945-7731, 1945-774X. doi: 10.1257/pol.6.4.271. URL https://pubs.aeaweb.org/doi/10.1257/pol.6.4.271.

Jacob Bastian. The Rise of Working Mothers and the 1975 Earned Income Tax Credit. *American Economic Journal: Economic Policy*, 12(3):44–75, August 2020. ISSN 1945-7731, 1945-774X. doi: 10.1257/pol.20180039. URL https://pubs.aeaweb.org/doi/10.1257/pol.20180039.

Andrew Garin, Emilie Jackson, and Dmitri Koustas. New Work or Changes in Reporting? A Framework for Measuring Self-Employment Trends in Administrative Data. Technical report, 2021.

IRS. Statistics for Tax Returns with the Earned Income Tax Credit (EITC). Technical report, Internal Revenue Service, Washington DC, March 2023. URL https://www.eitc.irs.gov/eitc-central/statistics-for-tax-returns-with-eitc/statistics-for-tax-returns-with-the-earned-income#Previous%20Tax%20Years.

Charles Davis and Evan White. Who receives the Earned Income Tax Credit in California? Technical report, California Policy Lab, October 2019. URL https://www.capolicylab.org/who-receives-the-earned-income-tax-credit-in-california/.

Martin Feldstein. Tax Avoidance and the Deadweight Loss of the Income Tax. *The Review of Economics and Statistics*, 81(4):674–680, 1999. URL https://www.jstor.org/stable/2646716.

Raj Chetty, John N Friedman, and Emmanuel Saez. Using Differences in Knowledge Across Neighborhoods to Uncover the Impacts of the EITC on Earnings. *American Economic Review*, 103(7):2683–2721, December 2013. ISSN 0002-8282. doi: 10.1257/aer.103.7.2683. URL https://pubs.aeaweb.org/doi/10.1257/aer.103.7.2683.

Spencer Bastani and Håkan Selin. Bunching and non-bunching at kink points of the Swedish tax schedule. *Journal of Public Economics*, 109:36–49, January 2014. ISSN 00472727. doi: 10.1016/j.jpubeco.2013.09.010. URL https://linkinghub.elsevier.com/retrieve/pii/S0047272713001916.

Austin Nichols and Jesse Rothstein. The Earned Income Tax Credit (EITC). Technical Report w21211, National Bureau of Economic Research, Cambridge, MA, May 2015. URL http://www.nber.org/papers/w21211.pdf.

Andreas Olden and Jarle Møen. The triple difference estimator. *The Econometrics Journal*, 25 (3):531–553, September 2022. ISSN 1368-4221, 1368-423X. doi: 10.1093/ectj/utac010. URL https://academic.oup.com/ectj/article/25/3/531/6545797.

Nada Eissa and Hilary Williamson Hoynes. Taxes and the labor market participation of married couples: the earned income tax credit. *Journal of Public Economics*, 88(9-10):1931–1958, August 2004. ISSN 00472727. doi: 10.1016/j.jpubeco.2003.09.005. URL https://linkinghub.elsevier.com/retrieve/pii/S0047272703001440.

Hilary Hoynes. The Earned Income Tax Credit. *The ANNALS of the American Academy of Political and Social Science*, 686(1):180–203, November 2019. ISSN 0002-7162, 1552-3349. doi: 10.1177/0002716219881621. URL http://journals.sagepub.com/doi/10.1177/0002716219881621.

Maggie R Jones and James P. Ziliak. The Antipoverty Impact of the EITC: New Estimates from Survey and Administrative Tax Records. *National Tax Journal*, 75(3):451–479, 2022. doi: https://doi.org/10.1086/720614. URL https://www.journals.uchicago.edu/doi/abs/10.1086/720614.

Bruno Ferman and Cristine Pinto. Inference in Differences-in-Differences with Few Treated Groups and Heteroskedasticity. *The Review of Economics and Statistics*, 101(3):452–467, July 2019. ISSN 0034-6535, 1530-9142. doi: 10.1162/rest_a_00759. URL https://direct.mit.edu/rest/article/101/3/452-467/58517.

A Colin Cameron, Jonah B Gelbach, and Douglas L Miller. Bootstrap-Based Improvements for Inference with Clustered Errors. *Review of Economics and Statistics*, 90(3):414–427, 2008. doi: https://doi.org/10.1162/rest.90.3.414.

Doruk Cengiz, Arindrajit Dube, Attila Lindner, and David Zentler-Munro. Seeing beyond the Trees: Using Machine Learning to Estimate the Impact of Minimum Wages on Labor Market Outcomes. *Journal of Labor Economics*, 40(S1):S203–S247, April 2022. ISSN 0734-306X, 1537-5307. doi: 10.1086/718497. URL https://www.journals.uchicago.edu/doi/10.1086/718497.

Anna Godøy, Michael Reich, Jesse Wursten, and Sylvia Allegretto. Parental Labor Supply: Evidence from Minimum Wage Changes. *Journal of Human Resources*, 59(2):416–442, March 2024. ISSN 0022-166X, 1548-8004. doi: 10.3368/jhr.1119-10540R2. URL http://jhr.uwpress.org/lookup/doi/10.3368/jhr.1119-10540R2.

Steven Ruggles, Sarah Flood, Matthew Sobek, Danika Brockman, Grace Cooper, Stephanie Richards, and Megan Schouweiler. IPUMS USA, 2023.

Maggie R. Jones and Amy B. O'Hara. Do Doubled-Up Families Minimize Household-Level Tax Burden? *National Tax Journal*, 69(3):613–640, September 2016. ISSN 0028-0283, 1944-7477. doi: 10.17310/ntj.2016.3.05. URL https://www.journals.uchicago.edu/doi/10.17310/ntj.2016.3.05.

David Splinter, Jeff Larrimore, and Jacob Mortenson. Whose Child is this? Shifting of Dependents Among EITC Claimants within the Same Household. *National Tax Journal*, 70(4): 737–758, December 2017. ISSN 0028-0283, 1944-7477. doi: 10.17310/ntj.2017.4.02. URL https://www.journals.uchicago.edu/doi/10.17310/ntj.2017.4.02.

Sarah Flood, Miriam King, Renae Rodgers, Steven Ruggles, J. Robert Warren, and Michael Westberry. Integrated Public Use Microdata Series, Current Population Survey, 2022.

Riley Wilson. The Self-Employment Effects of the EITC in the Gig Economy. *Working Paper*, 2023.

H Joseph Newton, Nicholas J Cox, Christopher F Baum, Nathaniel Beck, Rino Bellocco, Frauke Kreuter, Peter A Lachenbruch, Jens Lauritsen, Stanley Lemeshow, J Scott Long, Roger Newson, Austin Nichols, Marcello Pagano, Sophia Rabe-Hesketh, J Patrick Royston, Lisa Gilmore, David Culwell, and Deirdre Skaggs. Calibrating survey data using iterative proportional fitting (raking). *The Stata Journal*, 14(1):22–59, 2014.

Marianna Kudlyak. Involuntary Part-Time Work a Decade after the Recession. Technical Report 2019-30, Federal Reserve Bank of San Francisco, San Fransisco, November 2019.

Riley Wilson. The EITC and Employment Transitions: Labor Force Attachment and Annual Exit. *National Tax Journal*, 73(1):11–46, 2020. doi: https://doi.org/10.17310/ntj.2020.1.01. URL https://www.journals.uchicago.edu/doi/abs/10.17310/ntj.2020.1.01.

Damian Clarke, Daniel Pailañir, Susan Carleton Athey, and Guido W. Imbens. Synthetic Difference-in-Differences Estimation. *SSRN Electronic Journal*, 2023. ISSN 1556-5068. doi: 10.2139/ssrn.4346540. URL https://www.ssrn.com/abstract=4346540.

Alberto Abadie. Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects. *Journal of Economic Literature*, 59(2):391–425, June 2021. ISSN 0022-0515. doi: 10.1257/jel.20191450. URL https://pubs.aeaweb.org/doi/10.1257/jel.20191450.

Edward Harris and Shannon Mok. How CBO Estimates the Effects of the Affordable Care Act on the Labor Market. Working Paper 2015-09, Congressional Budget Office, Washington, D.C., 2015.

Mark Duggan, Gopi Shah Goda, and Emilie Jackson. The Effect of the Affordable Care Act on Health Insurance Coverage and Labor Market Outcomes. *National Tax Journal*, 72(2):261–322, June 2019. ISSN 0028-0283, 1944-7477. doi: 10.17310/ntj.2019.2.01. URL https://www.journals.uchicago.edu/doi/10.17310/ntj.2019.2.01.

Lizhong Peng, Xiaohui Guo, and Chad D. Meyerhoefer. The effects of Medicaid expansion on labor market outcomes: Evidence from border counties. *Health Economics*, 29(3):245–260, March 2020. ISSN 1057-9230, 1099-1050. doi: 10.1002/hec.3976. URL https://onlinelibrary.wiley.com/doi/10.1002/hec.3976.

Laurens Cherchye, Bram De Rock, and Frederic Vermeulen. Married with Children: A Collective Labor Supply Model with Detailed Time Use and Intrahousehold Expenditure Information. *American Economic Review*, 102(7):3377–3405, December 2012. ISSN 0002-8282. doi: 10.1257/aer.102.7.3377. URL https://pubs.aeaweb.org/doi/10.1257/aer.102.7.3377.

Bradley Hardy, Charles Hokayem, and James P. Ziliak. Income Inequality, Race, and the EITC. *National Tax Journal*, 75(1):149–167, March 2022. ISSN 0028-0283, 1944-7477. doi: 10.1086/717959. URL https://www.journals.uchicago.edu/doi/10.1086/717959.

Hadi Elzayn, Evelyn Smith, Thomas Hertz, Cameron Guage, Arun Ramesh, Robin Fisher, Daniel E Ho, and Jacob Goldin. Measuring and Mitigating Racial Disparities in Tax Audits. *The Quarterly Journal of Economics*, page qjae027, September 2024. ISSN 0033-5533, 1531-4650. doi: 10.1093/qje/qjae027. URL https://academic.oup.com/qje/advance-article/doi/10.1093/qje/qjae027/7778259.

Maggie R Jones. Changes in EITC Eligibility and Participation, 2005—2009. *CARRA Working Paper Series*, Working Paper #2014-04, 2014.

Dana Thomson, Yiyu Chen, Lisa A. Gennetian, and Luis E. Basurto. Earned Income Tax Credit Receipt By Hispanic Families With Children: State Outreach And Demographic Factors: Study examines state outreach efforts and demographic factors in Earned Income Tax Credit receipt by Hispanic families with children. *Health Affairs*, 41(12):1725–1734, December 2022. ISSN 0278-2715, 1544-5208. doi: 10.1377/hlthaff.2022.00725. URL http://www.healthaffairs.org/doi/10.1377/hlthaff.2022.00725.

Nanette Goodman, Michael Morris, and Debra L Brucker. Use of the Earned Income Tax Credit among people with disabilities. *Completed as part of the Employment Policy and Measurement RRTC funded by NIDRR*, 2013.

Jiafeng Chen and Jonathan Roth. Logs with Zeros? Some Problems and Solutions. *The Quarterly Journal of Economics*, 139(2):891–936, March 2024. ISSN 0033-5533, 1531-4650. doi: 10.1093/qje/qjad054. URL https://academic.oup.com/qje/article/139/2/891/7473710.

Bruce D. Meyer and Nikolas Mittag. Using Linked Survey and Administrative Data to Better Measure Income: Implications for Poverty, Program Effectiveness, and Holes in the Safety

Net. *American Economic Journal: Applied Economics*, 11(2):176–204, April 2019. ISSN 1945-7782, 1945-7790. doi: 10.1257/app.20170478. URL https://pubs.aeaweb.org/doi/10.1257/app.20170478.

Bruce D Meyer, Derek Wu, Victoria Mooers, and Carla Medalia. The Use and Misuse of Income Data and Extreme Poverty in the United States. *Journal of Labor Economics*, 39(S1), 2021.

Bruce D Meyer, Derek Wu, Grace Finley, Patrick Langetieg, Carla Medalia, Mark Payne, and Alan Plumley. The Accuracy of Tax Imputations Estimating Tax Liabilities and Credits Using Linked Survey and Administrative Data. In Raj Chetty, John N Friedman, Janet C. Gornick, Barry Johnson, and Arthur Kennickell, editors, *Measuring Distribution and Mobility of Income and Wealth*, volume 80 of *NBER Studies in Income and Wealth*. University of Chicago Press, 2022.

Henrik Kleven. The EITC and the Extensive Margin: A Reappraisal. Technical Report w26405, National Bureau of Economic Research, Cambridge, MA, October 2019. URL http://www.nber.org/papers/w26405.pdf.

Benjamin B. Lockwood. Optimal Income Taxation with Present Bias. *American Economic Journal: Economic Policy*, 12(4):298–327, November 2020. ISSN 1945-7731, 1945-774X. doi: 10.1257/pol.20180762. URL https://pubs.aeaweb.org/doi/10.1257/pol.20180762.

Franchise Tax Board. Tax Year 2019 California Earned Income Tax Credit and Young Child Tax Credit Report. Technical report, California, 2021. URL https://www.ftb.ca.gov/about-ftb/data-reports-plans/index.html.

Cristobal Young, Charles Varner, Ithai Z. Lurie, and Richard Prisinzano. Millionaire Migration and Taxation of the Elite: Evidence from Administrative Data. *American Sociological Review*, 81(3):421–446, June 2016. ISSN 0003-1224, 1939-8271. doi: 10.1177/0003122416639625. URL http://journals.sagepub.com/doi/10.1177/0003122416639625.

Joshua Rauh and Ryan Shyu. Behavioral Responses to State Income Taxation of High Earners: Evidence from California. Technical Report w26349, National Bureau of Economic Research, Cambridge, MA, October 2019. URL http://www.nber.org/papers/w26349.pdf.

Ben Brewer, Karen Smith Conway, and Jonathan C. Rork. Protecting the Vulnerable or Ripe for Reform? State Income Tax Breaks for the Elderly—Then and Now. *Public Finance Review*, 45(4):564–594, July 2017. ISSN 1091-1421, 1552-7530. doi: 10.1177/1091142116665903. URL http://journals.sagepub.com/doi/10.1177/1091142116665903.

Salla Kalin, Antoine B Levy, and Mathilde Muñoz. Pensioners Without Borders: Agglomeration and the Migration Response to Taxation. Technical Report 32890, National Bureau of Economic Research, Cambridge, MA, August 2024. URL http://www.nber.org/papers/w32890.

Karen Smith Conway and Jonathan C. Rork. State "Death" Taxes and Elderly Migration—The Chicken or the Egg? *National Tax Journal*, 59(1):97–128, March 2006. ISSN 0028-0283,

1944-7477. doi: 10.17310/ntj.2006.1.05. URL https://www.journals.uchicago.edu/doi/10.17310/ntj.2006.1.05.

Karen Smith Conway and Jonathan C. Rork. THE GENESIS OF SENIOR INCOME TAX BREAKS. *National Tax Journal*, 65(4):1043–1068, December 2012a. ISSN 0028-0283, 1944-7477. doi: 10.17310/ntj.2012.4.16. URL https://www.journals.uchicago.edu/doi/10.17310/ntj.2012.4.16.

Charles Adam Bee and Joshua Mitchell. Do Older Americans Have More Income Than We Think? *SSRN Electronic Journal*, 2017. ISSN 1556-5068. doi: 10.2139/ssrn.3015870. URL https://www.ssrn.com/abstract=3015870.

Peter J. Brady and Steven Bass. Comparing the Current Population Survey to Income Tax Data. *SSRN Electronic Journal*, 2021. ISSN 1556-5068. doi: 10.2139/ssrn.4025470. URL https://www.ssrn.com/abstract=4025470.

Henrik Kleven, Camille Landais, Mathilde Muñoz, and Stefanie Stantcheva. Taxation and Migration: Evidence and Policy Implications. *Journal of Economic Perspectives*, 34(2):119–142, May 2020. ISSN 0895-3309. doi: 10.1257/jep.34.2.119. URL https://pubs.aeaweb.org/doi/10.1257/jep.34.2.119.

Ning Jia, Raven Molloy, Christopher Smith, and Abigail Wozniak. The Economics of Internal Migration: Advances and Policy Questions. *Journal of Economic Literature*, 61 (1):144–180, March 2023. ISSN 0022-0515. doi: 10.1257/jel.20211623. URL https://pubs.aeaweb.org/doi/10.1257/jel.20211623.

Raven Molloy, Christopher L. Smith, and Abigail Wozniak. Job Changing and the Decline in Long-Distance Migration in the United States. *Demography*, 54(2):631–653, April 2017. ISSN 0070-3370, 1533-7790. doi: 10.1007/s13524-017-0551-9. URL https://read.dukeupress.edu/demography/article/54/2/631/167691/Job-Changing-and-the-Decline-in-Long-Distance.

Tamim Bayoumi and Jelle Barkema. Stranded! How Rising Inequality Suppressed US Migration and Hurt Those "Left Behind". Technical Report 122, International Monetary Fund, Washington DC, June 2019.

William W. Olney and Owen Thompson. The Determinants of Declining Internal Migration. Technical Report 23123, National Bureau of Economic Research, Cambridge, MA, February 2024. URL https://www.nber.org/papers/w32123.

Kristina Komissarova. Location Choices over the Life Cycle: The Role of Relocation for Retirement. *Working Paper*, 2022.

Jon Bakija and Joel Slemrod. Do the Rich Flee from High State Taxes? Evidence from Federal Estate Tax Returns. Technical Report w10645, National Bureau of Economic Research, Cambridge, MA, July 2004. URL http://www.nber.org/papers/w10645.pdf.

Charles Varner, Cristobal Young, Allen Prohofsky, and Allen Prohofsky. Millionaire Migration in California:. *Working Paper*, 2018.

Joshua Rauh and Ryan Shyu. Behavioral Responses to State Income Taxation of High Earners: Evidence from California. *American Economic Journal: Economic Policy*, 16(1):34–86, February 2024. ISSN 1945-7731, 1945-774X. doi: 10.1257/pol.20200500. URL https://pubs.aeaweb.org/doi/10.1257/pol.20200500.

Katrine Jakobsen, Henrik Kleven, Jonas Kolsrud, Camille Landais, and Mathilde Muñoz. Taxing Top Wealth: Migration Response and Their Aggregate Economic Implications. Technical Report 32153, National Bureau of Economic Research, Cambridge, MA, February 2024. URL http://www.nber.org/papers/w32153.

Arun Advani, Helen Hughson, and Andy Summers. How much tax do the rich really pay? Evidence from the UK. *Oxford Review of Economic Policy*, 39(3):406–437, August 2023. ISSN 0266-903X, 1460-2121. doi: 10.1093/oxrep/grad032. URL https://academic.oup.com/oxrep/article/39/3/406/7245704.

Marius Brülhart and Raphaël Parchet. Alleged tax competition: The mysterious death of bequest taxes in Switzerland. *Journal of Public Economics*, 111:63–78, March 2014. ISSN 00472727. doi: 10.1016/j.jpubeco.2013.12.009. URL https://linkinghub.elsevier.com/retrieve/pii/S004727271300251X.

Marius Brülhart, Jonathan Gruber, Matthias Krapf, and Kurt Schmidheiny. Behavioral Responses to Wealth Taxes: Evidence from Switzerland. *American Economic Journal: Economic Policy*, 14(4):111–150, November 2022. ISSN 1945-7731, 1945-774X. doi: 10.1257/pol.20200258. URL https://pubs.aeaweb.org/doi/10.1257/pol.20200258.

Marko Köthenbürger, Costanza Naguib, Christian Stettler, and Michael Stimmelmayr. Income Taxes and the Mobility of the Rich: Evidence from Us and UK Households in Switzerland. *SSRN Electronic Journal*, 2023. ISSN 1556-5068. doi: 10.2139/ssrn.4422851. URL https://www.ssrn.com/abstract=4422851.

Arun Advani and Hannah Tarrant. Behavioural responses to a wealth tax. *Fiscal Studies*, 42(3-4):509–537, September 2021. ISSN 0143-5671, 1475-5890. doi: 10.1111/1475-5890.12283. URL https://onlinelibrary.wiley.com/doi/10.1111/1475-5890.12283.

Raven Molloy, Christopher L Smith, and Abigail Wozniak. Internal Migration in the United States. *Journal of Economic Perspectives*, 25(3):173–196, August 2011. ISSN 0895-3309. doi: 10.1257/jep.25.3.173. URL https://pubs.aeaweb.org/doi/10.1257/jep.25.3.173.

Karen Smith Conway and Jonathan C. Rork. The Changing Roles of Disability, Veteran, and Socioeconomic Status in Elderly Interstate Migration. *Research on Aging*, 33(3):256–285, May 2011. ISSN 0164-0275, 1552-7573. doi: 10.1177/0164027511399107. URL https://journals.sagepub.com/doi/10.1177/0164027511399107.

Karen Smith Conway and Jonathan C. Rork. No Country for Old Men (or Women): Do State Tax Policies Drive Away the Elderly? *National Tax Journal*, 65(2):313–356, June 2012b. ISSN

0028-0283, 1944-7477. doi: 10.17310/ntj.2012.2.03. URL https://www.journals.uchicago.edu/doi/10.17310/ntj.2012.2.03.

Marco Badilla Maroto, Benjamin Faber, Antoine Levy, and Mathilde Munoz. Senior Migration, Local Economic Development and Spatial Inequality. *Working Paper*, 2024.

Amy Finkelstein, Matthew Gentzkow, and Heidi Williams. Place-Based Drivers of Mortality: Evidence from Migration. *American Economic Review*, 111(8):2697–2735, August 2021. ISSN 0002-8282. doi: 10.1257/aer.20190825. URL https://pubs.aeaweb.org/doi/10.1257/aer.20190825.

Ben Brewer, Karen Smith Conway, and Jonathan C. Rork. Do income tax breaks for the elderly affect economic growth? *Contemporary Economic Policy*, 40(1):7–27, January 2022. ISSN 1074-3529, 1465-7287. doi: 10.1111/coep.12549. URL https://onlinelibrary.wiley.com/doi/10.1111/coep.12549.

Jon Bakija. Documentation for a Comprehensive Historical U.S Federal and State Income Tax Calculator Program. Version 2019.01, Williams College, Williamstown, MA, March 2019.

Raj Chetty, John N. Friedman, Emmanuel Saez, and Danny Yagan. The SOI Databank: A case study in leveraging administrative data in support of evidence-based policymaking. *Statistical Journal of the IAOS*, 34(1):99–103, February 2018. ISSN 18747655, 18759254. doi: 10.3233/SJI-170418. URL https://www.medra.org/servlet/aliasResolver?alias=iospress&doi=10.3233/SJI-170418.

Kevin Pierce. SOI Migration Data: A New Approach. Technical report, Internal Revenue Service, 2015.

Danny Yagan. Employment Hysteresis from the Great Recession. *Journal of Political Economy*, 127(5):2505–2558, 2019. doi: 0022-3808/2019/12705-0012$10.00.

Brad Foster. Agree to Disagree? Comparing IRS, NCOA, and Census Bureau Survey Migration Measures, 2023. URL https://www.fcsm.gov/events/2023-fcsm-conference/.

Sergio Correia, Paulo Guimarães, and Thomas Zylkin. ppmlhdfe: Fast Poisson Estimation with High-Dimensional Fixed Effects. *The Stata Journal: Promoting communications on statistics and Stata*, 20(1):95–115, March 2020. ISSN 1536-867X, 1536-8734. doi: 10.1177/1536867X20909691. URL http://arxiv.org/abs/1903.01690. arXiv:1903.01690 [econ].

The Urban Institute and The Urban-Brookings Tax Policy Center. State and Local Government Finance Data | Tax Policy Center, 2024. URL https://state-local-finance-data.taxpolicycenter.org.

Thomas Piketty and Emmanuel Saez. Income Inequality in the United States, 1913–1998. *Quarterly Journal of Economics*, 118(1):1–41, 2003.

Gerald Auten and David Splinter. Comment: Tax Evasion at the Top of the Income Distribution: Theory and Evidence. *Comment*, 2021.

Daniel Reck, Max Risch, and Gabriel Zucman. Response to a Comment by Auten and Splinter on "Tax Evasion at the Top of the Income Distribution: Theory and Evidence". *Comment*, October 2021.

William Gale, John Sabelhous, and Samuel I. Thorpe. Measuring income inequality: A primer on the debate. Commentary, The Brookings Instution, Washington D.C., December 2023. URL https://www.brookings.edu/articles/measuring-income-inequality-a-primer-on-the-debate/.

Thomas Piketty, Emmanuel Saez, and Gabriel Zucman. Income Inequality in the United States: A Comment, 2024.

Gerald Auten and David Splinter. Reply to Piketty, Saez, and Zucman (2024): Income Inequality in the United States. 2025a.

Gerald Auten and David Splinter. Reply to Iselin and Reck (2024): Income Inequality in the United States. 2025b.

John Guyton, Patrick Langetieg, Daniel Reck, Max Risch, and Gabriel Zucman. The Distribution of Undetected Under-Reported Income Identified by Detection-Controlled Estimation Methods. Technical report, (Availible in the appendix of Guyton et al. (2023b)), November 2023a.

Jonathan S. Feinstein. Detection Controlled Estimation. *The Journal of Law and Economics*, 33 (1):233–276, April 1990. ISSN 0022-2186, 1537-5285. doi: 10.1086/467205. URL https://www.journals.uchicago.edu/doi/10.1086/467205.

Jonathan S. Feinstein. An Econometric Analysis of Income Tax Evasion and Its Detection. *The RAND Journal of Economics*, 22(1):14, 1991. ISSN 07416261. doi: 10.2307/2601005. URL http://doi.wiley.com/10.2307/2601005.

Brian Erard and Jonathan S. Feinstein. Econometric Models for Multi-Stage Audit Processes: An Application to the IRS National Research Program. In James Alm, Jorge Martinez-Vazquez, and Benno Torgler, editors, *Developing Alternative Frameworks for Explaining Tax Compliance*. Taylor & Francis Group, London, 1st edition, 2010.

Brian Erard and Jonathan Feinstein. The Individual Income Reporting Gap: What We See and What We Don't. Technical Report Publication 1500 (rev 4-2012), Internal Revenue Service, Washington D.C., 2011.

Gerald Auten and Patrick Langetieg. The Distribution of Unreported Income: What We Can Learn from the NRP. February 2023.

Matthew Smith, Danny Yagan, Owen Zidar, and Eric Zwick. Capitalists in the Twenty-First Century. *Quarterly Journal of Economics*, 134(4):1675–1745, 2019. doi: https://doi.org/10.1093/qje/qjz020. URL https://academic.oup.com/qje/article-abstract/134/4/1675/5542244.

Wojciech Kopczuk and Eric Zwick. Business Incomes at the Top. *Journal of Economic Perspectives*, 34(4):27–51, November 2020. ISSN 0895-3309. doi: 10.1257/jep.34.4.27. URL https://pubs.aeaweb.org/doi/10.1257/jep.34.4.27.

John Guyton, Patrick Langetieg, Daniel Reck, Max Risch, and Gabriel Zucman. Tax Evasion at the Top of the Income Distribution: Theory and Evidence. Working Paper 28542, National Bureau of Economic Research, Cambridge, MA, December 2023b.

Joel Slemrod. Cheating Ourselves: The Economics of Tax Evasion. *Journal of Economic Perspectives*, 21(1):25–48, January 2007. ISSN 0895-3309. doi: 10.1257/jep.21.1.25. URL https://pubs.aeaweb.org/doi/10.1257/jep.21.1.25.

David Leite. Micro Evidence and Aggregate Implications of Consumption Through the Firm. Working Paper 2024-16, Saîd Business School, 2024. URL https://oxfordtax.sbs.ox.ac.uk/sitefiles/wp2416-leite-david.pdf&ved=2ahUKEwiYsqHy1YOMAxXQF1kFHZzBNYcQFnoECA8QAQ&usg=AOvVaw2X95xalgf4EU5tzTk9gz6r.

Emily Black, Ryan Hess, Rebecca Lester, Jacob Goldin, Daniel E Ho, Mansheej Paul, and Annette Portz. The Spiderweb of Partnership Tax Structures. 2023.

Daniel J. Hemel, Janet Holtzblatt, and Steve Rosenthal. The Tax Gap's Many Shades of Gray. *SSRN Electronic Journal*, 2021. ISSN 1556-5068. doi: 10.2139/ssrn.3934044. URL https://www.ssrn.com/abstract=3934044.

Theodore Black, Kim Bloomquist, Edward Emblom, Andrew Johns, Alan Plumley, and Esmeralda Stuk. Federal Tax Compliance Research: Tax Year 2006 Tax Gap Estimation. *IRS Research, Analysis & Statistics Working Paper*, 2012.

Barry W. Johnson and Peter J. Rose. Federal Tax Compliance Research: Tax Gap Estimates for Tax Years 2011–2013. Technical Report 1415, Internal Revenue Service, Research, Applied Analytics & Statistics, Washington D.C., 2019.

Sébastien Laffitte. The Market for Tax Havens, 2024.

Kavya Vaghul and Ben Zipperer. Historical State and Sub-state Minimum Wages., 2022. URL https://github.com/benzipperer/historicalminwage/releases/tag/v1.4.0.

Marianne P. Bitler, Hilary W. Hoynes, and John Iselin. Cyclicality of the U.S. Safetey Net: Evidence from the 200s and implications for the COVID-19 Crisis. *National Tax Journal*, 73(3):759–780, September 2020. ISSN 0028-0283, 1944-7477. doi: 10.17310/ntj.2020.3.06. URL https://www.journals.uchicago.edu/doi/10.17310/ntj.2020.3.06.

Nolan G. Pope. The Effects of DACAmentation: The Impact of Deferred Action for Childhood Arrivals on Unauthorized Immigrants. *Journal of Public Economics*, 143:98–114, November 2016. ISSN 00472727. doi: 10.1016/j.jpubeco.2016.08.014. URL https://linkinghub.elsevier.com/retrieve/pii/S0047272716301268.

Andrew Garin, Emilie Jackson, and Dmitri Koustas. New Gig Work or Changes in Reporting? Understanding Self-Employment Trends in Tax Data. *American Economic Journal: Applied Economics*, Forthcoming, 2024. URL https://www.aeaweb.org/articles?id=10.1257/app.20220483&from=f.

Katharine G. Abraham, John C. Haltiwanger, Claire Hou, Kristin Sandusky, and James R. Spletzer. Reconciling Survey and Administrative Measures of Self-Employment. *Journal of Labor Economics*, 39(4):825–860, October 2021. ISSN 0734-306X, 1537-5307. doi: 10.1086/712187. URL https://www.journals.uchicago.edu/doi/10.1086/712187.

Katharine G. Abraham, Brad Hershbein, Susan N. Houseman, and Beth C. Truesdale. The Independent Contractor Workforce: New Evidence on Its Size and Composition and Ways to Improve Its Measurement in Household Surveys. *ILR Review*, 77(3):336–365, May 2024a. ISSN 0019-7939, 2162-271X. doi: 10.1177/00197939241226945. URL https://journals.sagepub.com/doi/10.1177/00197939241226945.

Katharine G Abraham, John Haltiwanger, Kristin Sandusky, and James Spletzer. The Rise of the Gig Economy. 2024b.

Marianne Bertrand, Esther Duflo, and Sendhil Mullainathan. How Much Should We Trust Differences-in-Differences Estimates? *Quarterly Journal of Economics*, 119(1):249–275, 2004. URL https://www.jstor.org/stable/25098683.

Alberto Abadie, Susan Athey, Guido W Imbens, and Jeffrey M Wooldridge. When Should You Adjust Standard Errors for Clustering? *The Quarterly Journal of Economics*, 138(1):1–35, December 2022. ISSN 0033-5533, 1531-4650. doi: 10.1093/qje/qjac038. URL https://academic.oup.com/qje/article/138/1/1/6750017.

James G. MacKinnon, Morten Ørregaard Nielsen, and Matthew D. Webb. Cluster-robust inference: A guide to empirical practice. *Journal of Econometrics*, 232(2):272–299, February 2023. ISSN 03044076. doi: 10.1016/j.jeconom.2022.04.001. URL https://linkinghub.elsevier.com/retrieve/pii/S0304407622000781.

Stephen G Donald and Kevin Lang. Inference with Difference-in-Differences and Other Panel Data. *Review of Economics and Statistics*, 89(2):221–233, May 2007. ISSN 0034-6535, 1530-9142. doi: 10.1162/rest.89.2.221. URL https://direct.mit.edu/rest/article/89/2/221-233/57678.

James G. MacKinnon and Matthew D. Webb. Wild Bootstrap Inference for Wildly Different Cluster Sizes. *Journal of Applied Econometrics*, 32(2):233–254, March 2017. ISSN 0883-7252, 1099-1255. doi: 10.1002/jae.2508. URL https://onlinelibrary.wiley.com/doi/10.1002/jae.2508.

Luis Alvarez and Bruno Ferman. Extensions for Inference in Difference-in-Differences with Few Treated Clusters, February 2023. URL http://arxiv.org/abs/2302.03131. arXiv:2302.03131 [econ].

Timothy G Conley and Christopher R Taber. Inference with "Difference in Difference With a Small Number of Policy Changes. *The Review of Economics and Statistics*, 93(1):113–125, 2011. URL https://www.jstor.org/stable/23015923.

James G. MacKinnon and Matthew D. Webb. Randomization inference for difference-in-differences with few treated clusters. *Journal of Econometrics*, 218(2):435–450, October 2020. ISSN 03044076. doi: 10.1016/j.jeconom.2020.04.024. URL https://linkinghub.elsevier.com/retrieve/pii/S0304407620301445.

Alberto Abadie, Alexis Diamond, and Jens Hainmueller. Comparative Politics and the Synthetic Control Method. *American Journal of Political Science*, 59 (2):495–510, 2015. ISSN 1540-5907. doi: 10.1111/ajps.12116. URL http://onlinelibrary.wiley.com/doi/abs/10.1111/ajps.12116. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ajps.12116.

James G. MacKinnon and Matthew D. Webb. Wild Bootstrap Randomization Inference for Few Treated Clusters. In *The Econometrics of Complex Survey Data*, volume 39 of *Advances in Econometrics*, pages 61–85. Emerald Publishing Limited, January 2019. ISBN 978-1-78756-726-9 978-1-78756-725-2. doi: 10.1108/S0731-905320190000039003. URL https://doi.org/10.1108/S0731-905320190000039003.

Andreas Hagemann. Inference with a Single Treated Unit. 2024.

Andrew Johns and Joel Slemrod. The Distribution of Income Tax Noncompliance. *National Tax Journal*, 63(3):397–418, September 2010. ISSN 0028-0283, 1944-7477. doi: 10.17310/ntj.2010.3.01. URL https://www.journals.uchicago.edu/doi/10.17310/ntj.2010.3.01.

Mark J. Mazur and Alan H. Plumley. Understanding the Tax Gap, 2007.

Roger H. Gordon and Joel Slemrod. Are "Real" Responses to Taxes Simply Income Shifting Between Corporate and Personal Tax Bases? In Joel Slemrod, editor, *Does Atlas Shrug? The Economic Consequences of Taxing the Rich.* Harvard University Press, Cambridge, MA, 2002.

Andrew Samwick. Tax Shelters and Passive Losses After the Tax Reform Act of 1986. Technical Report w5171, National Bureau of Economic Research, Cambridge, MA, July 1995. URL http://www.nber.org/papers/w5171.pdf.

Internal Revenue Service. Federal Tax Compliance Research: Individual Income Tax Gap Estimates for 1985, 1988, and 1992. Technical Report Publicaiton 1415, U.S. Department of the Treasury, Washington D.C., 1996.