

To the Graduate Council:

I am submitting herewith a dissertation written by Diego A. Guerrero entitled "Essays on the Economics of Federal Programs." I have examined the final paper copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Economics.

Donald J. Bruce, Major Professor

We have read this dissertation
and recommend its acceptance:

Matthew C. Harris

Celeste K. Carruthers

Luke P. Rodgers

Accepted for the Council:

Dixie L. Thompson

Vice Provost and Dean of the Graduate School

To the Graduate Council:

I am submitting herewith a dissertation written by Diego A. Guerrero entitled “Essays on the Economics of Federal Programs.” I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Economics.

Donald J. Bruce, Major Professor

We have read this dissertation
and recommend its acceptance:

Matthew C. Harris

Celeste K. Carruthers

Luke P. Rodgers

Accepted for the Council:

Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Essays on the Economics of Federal Programs

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Diego A. Guerrero

May 2024

Copyright © 2024 by Diego A. Guerrero
All rights reserved.

Acknowledgements

I am deeply grateful for the support and extensive comments provided by my dissertation committee: Don Bruce, Celeste Carruthers, Matt Harris, and Luke Rodgers. My thanks extend to the generous financial support and the dedicated family at the the University of Tennessee's Boyd Center for Business and Economic Research: Bill Fox, Erin Hatfield, Enda Hagarden, Kira Rasmussen, Larry Kessler, Michael Kofoed, LeAnn Luna, Seth Neller, Tim Kuhn, Alex Norwood, Emily Pratt, and Ed Taylor.

The Department of Economics at the University of Tennessee has been my academic home throughout this half-decade. I extend my gratitude to Marianne Wanamaker, who guided the early stages of this project, and my Second Year Paper Committee: Andrew Hanson, Christian Vossler, Charles Sims, and Scott Holladay. I want to give a special mention to Richard Beem, Lucas Garcez, Scott Gilpatric, Mohammed Mohsin, Maria Padilla-Romo, Sherri Pinkston, Georg Schaur, Thomas Verville, Ge Wu, and attendees of the UT Brownbag Series. Beyond the Economics Department, I have enjoyed the unwavering friendship of Nicolas Arguello. The High-Performance & Scientific Computing group and the Division of Research Administration provided invaluable assistance.

A supportive network of economists and scholars continuously enriches my thinking. This dissertation benefited from comments by Ana Pena (Colorado State University), Caroline Weber (University of Kentucky), participants at the Southern Economic Conference and the National Tax Association, and several anonymous

referees. Katherine Baicker, Mireille Jacobson, Wendy Koslicki, and Federico Masera kindly shared their data, code, and answered inquiries. I thank the staff at the Bureau of Labor Statistics and NORC at the University of Chicago. I extend my thanks to Mana Komai, Lynn MacDonald, and Monica Garcia-Perez for their lasting support, and the faculty at the University of Colorado Denver who welcomed me in their seminars. I am also grateful for being awarded the Adam Smith Fellowship by the Mercatus Center, and extend my appreciation to the participants who engaged in stimulating debates on policy-making and governance, specially Peter Boettke, Rosolino Candela, Christopher Coyne, Erwin Dekker, Paul Dragos, Stefanie Haeffele, Logan Hansen, Arielle John, and Jordan Lofthouse. Moreover, my gratitude goes to the Inter-American Development Bank and, specifically, to Paloma Acevedo, Leandro Adrian, Javier Beverinotti, Emily Diaz, and Osmel Manzano. I also thank my collaborators and seniors, Juan Pantano, Jose Saboin, and Wladimir Zaroni.

Lastly, I am indebted to my friends and family, whose patience sustained my ambitions. My friends, too many to enumerate, supported me when they were needed most and when they were not. To my mother, who received fewer phone calls; to my father, whose contagious work ethics was an inspiration; and to my wife, I am immensely fortunate for your understanding and love. And to Chubby, whose impatience for meals and walking breaks taught me a lesson on priorities.

Abstract

This dissertation consists of three essays in public economics and the political economy of the public sector. The document contributes empirical evidence on the impact of public policies on households, firms, and government.

This first essay revisits the evidence on the intergenerational effects of the earned income tax credit (EITC) on the adult income of individuals who received the tax credit during childhood. The empirical strategy exploits formulaic changes localized across socioeconomic status in the tax schedule, and finds that tax credit transfers during early childhood increase adult income between 1.7 and 2.7 percent per \$1,000. Furthermore, the paper links the intergenerational impact of the tax credit with maternal labor supply, arguing that the effect is driven by children of mothers working the fewest hours.

The second essay addresses the association between the EITC and employment, wages, and establishments using establishment surveys. The empirical strategy leverages differences in state roll-out and differences in industry and occupational exposure. The results point that the EITC is related with a transitory decrease in wages between 2 and 8 percent, but there is no discernible impact on employment and the number of establishments.

The third chapter is a collaboration with Eunsik Chang and Matthew Harris. The paper studies the relationship between police departments acquiring military hardware and the failure to disclose the homicide of civilians to the federal government. We examine the timing of non-disclosed deaths and the acquisition

of military hardware, and find evidence supporting moral hazard. There is also suggestive evidence of adverse selection among agencies. These results call into question the interpretation of previous results, and allow me to conclude that the federal government has failed to monitor the local law enforcement agencies to whom it gave lethal capabilities.

Table of Contents

1	Introduction	1
2	The Intergenerational Effects of the EITC on the Adult Income of Children	8
2.1	The Earned Income Tax Credit	15
2.2	Data	17
2.3	Empirical Strategy	23
2.4	Results	34
2.4.1	Dynamic complementarities and aggregated EITC	39
2.4.2	Heterogeneities	41
2.4.3	Mechanism analysis: maternal labor supply	47
2.5	Discussion	49
3	Business as usual? Firm Responses to the EITC	65
3.1	The EITC supplement programs	70
3.2	Empirical Strategy	75
3.3	Data	77
3.4	Results	83
3.4.1	Industry Heterogeneities	86
3.4.2	Occupational Heterogeneities	93
3.5	Robustness	95
3.5.1	Event Studies	95

3.5.2	Timing in policy enactment	96
3.5.3	Variations in the exposure threshold	96
3.6	Discussion	97
4	Not Just About Optics: Adverse Selection and Moral Hazard in Police Militarization and Citizens Deaths	110
4.1	Institutional Background	116
4.1.1	Police Killings of Citizens	116
4.1.2	Institutional Details on the 1033 Program	118
4.1.3	Mixed Evidence on Police Militarization and Civilian Deaths .	120
4.2	Data	122
4.2.1	Characterizing reports of homicides by police	126
4.2.2	Characterizing 1033 participation	133
4.3	Empirical Strategy	134
4.4	Results	142
4.4.1	First stage	142
4.4.2	The impact of military hardware on homicides by police . . .	143
4.4.3	The impact of military hardware on the disclosure of killings by police	145
4.4.4	Leading indicators of 1033 take-up	152
4.5	Discussion	157
A	Appendix A	164
A.1	Income and tax credits	165
A.2	Policy Variation	170
A.3	EITC Exposure in the CNLSY	174
A.4	First Stage	175
A.5	Sensitivity to covariates	179
A.6	Placebo Test	180
A.7	Maternal Labor Supply	183

B	Appendix B	188
B.1	Industry Estimates	189
B.2	Occupational Estimates	191
B.3	Additional robustness tests	198
C	Appendix C	204
C.0.1	Descriptive Tables and Figures	205
C.0.2	Robustness tests	211
C.0.3	Transitional Matrices	219
C.0.4	Assets Forfeiture	222
C.0.5	Lagged under-reports	226
C.0.6	Event studies	229
	Vita	242

List of Tables

2.1	Descriptive statistics.	26
2.2	Calculated EITC Exposure and Imputed EITC benefits in childhood.	27
2.3	OLS and 2SLS coefficients of imputed earned income tax credit on adult income between age 22 and 40.	38
2.4	OLS and 2SLS effects of tax credits on adult income using single endogenous variables and total EITC in childhood.	43
3.1	Roll-out year in state supplementary EITC programs.	73
3.2	Descriptive Statistics in 2018.	80
3.3	Average impact of the EITC on industry outcomes.	87
3.4	Average impact of the EITC on occupational outcomes.	88
4.1	Total quantity of 1033 items shipped by category.	128
4.2	Balance test of agency characteristics and disclosure of civilian homicides by police.	132
4.3	Balance test of agency characteristics and 1033 take-up.	138
4.4	Balance test of 1033 take-up per sworn officer and disclosure of civilian homicides by police.	139
4.5	The effect of military items on the probability of under-reporting.	150
4.6	Military items and under-reporting conditional on a civilian being killed.	155
A.1	Moments of the distribution of tax credit variables.	170

A.2	OLS and 2SLS coefficients of earned income tax credit exposure on adult income between age 22 and 40.	174
A.3	Correlation between policy variation and covariates across childhood stages.	177
A.4	Correlation between policy variation and covariates.	178
B.1	Average impact of the EITC on industry outcomes in logs.	189
B.2	Estimates of the EITC on the log of occupational outcomes.	191
B.3	Estimates of the EITC on occupational outcomes without occupational exposure.	192
B.4	Impact of the EITC on occupational outcomes conditional on refundability.	193
B.5	Population weighted coefficients of the EITC on industry outcomes. .	198
C.1	Total Original Acquisition Value of 1033 items shipped by category. .	206
C.2	The effect of military items on killings by police with alternative instruments.	212
C.3	The effect of one additional standard deviation in military items on the probability of under-reporting.	213
C.4	The effect of military items on total under-reports.	214
C.5	The effect of logarithmic military items on total under-reports.	215
C.6	The effect of military items (OAV) on the probability of under-reporting.	216
C.7	The effect of military items on the probability of under-reporting using categories in Masera (2021a).	217
C.8	The effect of total military items on the probability of under-reporting with alternative instruments.	218
C.9	Previous under-reports and the stock of 1033 items conditional on former UCR respondents.	227
C.10	Previous under-reports and the stock of 1033 items.	228

List of Figures

2.1	Maximum EITC benefit by number of dependents.	18
2.2	The individual's budget constraint and the EITC schedule during the Welfare Reform.	31
2.3	Standardized transfers and policy variation.	32
2.4	OLS Coefficients of imputed earned income tax credit on the adult income by childhood stage.	37
2.5	Effects of tax credits on adult income across different intervals of childhood stages.	42
2.6	Event study of years of childhood exposure to the 1993 EITC reform on adult income.	45
2.7	Subgroup analyses of tax credits on adult income between ages 22 and 40.	46
2.8	Maternal labor supply after birth by EITC status.	53
2.9	Intergenerational effects of the EITC on adult income conditional on the predicted maternal labor intensity.	54
3.1	EITC exposure across economic sectors.	81
3.2	EITC exposure across major occupations.	82
3.3	Heterogeneous effects of the EITC across economic sectors: Establishments.	89
3.4	Heterogeneous effects of the EITC across economic sectors: Employment.	90
3.5	Heterogeneous effects of the EITC across economic sectors: Wages.	91

3.6	Heterogeneous effects of the EITC and industry exposure.	92
3.7	Heterogeneous effects of the EITC and occupation exposure.	94
3.8	Event study of EITC roll-out and employment across industries.	98
3.9	Event study of EITC roll-out and wages across industries.	99
4.1	Civilians killed by police recorded in UCR and Fatal Encounters datasets.	117
4.2	Total non-disclosed homicides by police and the stock of military items.	129
4.3	Percent of agencies under-reporting per state (2000-2020).	130
4.4	Gap between Fatal Encounters and UCR among agencies that killed a civilian.	131
4.5	Under-reported homicides by police and years since 1033 take up, by shipment type.	140
4.6	Boots on the ground and the national stock of items.	147
4.7	Distance to the closest disposition center and predicted quantity of items.	148
4.8	The effect of military items on homicides using different instruments and datasets.	149
4.9	Racial and sex heterogeneities in the effect of under-reporting.	154
4.10	Probability of under-reports as police departments transition in the stock of items.	156
A.1	EITC Schedule for Single and Married Households in 2018.	167
A.2	Distribution of EITC exposure and imputed EITC, by cohort.	168
A.3	Distribution of income by percentile EITC exposure and by percentile of imputed EITC.	169
A.4	Standardized policy variation across alternative constructions of the instrument.	171
A.5	2SLS Coefficients of imputed earned income tax credit on the adult income by childhood stage.	172
A.6	2SLS subgroup analyses of tax credits on adult income between ages 22 and 40.	173

A.7	Sensitivity of first stage estimates to covariates.	176
A.8	Sensitivity to covariates on adult income.	179
A.9	Distribution of 2SLS coefficients with placebo instruments.	182
A.10	Maternal Labor Supply after birth by EITC status.	183
A.11	Maternal income after birth by EITC status.	184
A.12	Family income after birth by EITC status.	185
A.13	Intergenerational effects of the EITC on adult income conditional on maternal labor intensity prior to childbearing.	186
A.14	LATE of the EITC on adult income conditional on maternal labor intensity during early childhood.	187
B.1	Event study of EITC roll-out and establishments.	190
B.2	Heterogeneous effects of the EITC across major occupations: Employ- ment.	194
B.3	Heterogeneous effects of the EITC across major occupations: Hourly wages.	195
B.4	Event study of EITC roll-out and employment across occupations. . .	196
B.5	Event study of EITC roll-out and wages across occupations.	197
B.6	Sensitivity of coefficients to the exclusion of zeroes in the sample. . .	199
B.7	Correlation between state EITC roll-out and policy state covariates. .	200
B.8	Correlation between state EITC roll-out and state sociodemographic characteristics.	201
B.9	Sensitivity of treatment effects to various thresholds of exposure. . . .	202
B.10	Sensitivity of treatment effects to year of exposure.	203
C.1	Total civilians killed by police by race.	205
C.2	Distribution of the gap in civilian killings between Fatal Encounters and UCR.	207
C.3	Percent of UCR respondents that do not disclose a homicide by police.	208

C.4	Average gap between Fatal Encounters and UCR conditional on recent reporting behavior.	209
C.5	The distribution of 1033 take-up year.	209
C.6	The distribution of 1033 take-up year by shipment type.	210
C.7	Boots on the ground	211
C.8	Probability of under-reports as police departments transition in the stock of weapons.	219
C.9	Probability of under-reports as police departments transition in the stock of optics.	220
C.10	Probability of under-reports as police departments transition in the stock of vehicles.	221
C.11	Average real dollars per capita in assets seized as police departments transition in the stock of 1033 items.	222
C.12	Average real dollars per capita in assets seized as police departments transition in the stock of weapons.	223
C.13	Average real dollars per capita in assets seized as police departments transition in the stock of optics.	224
C.14	Average real dollars per capita in assets seized as police departments transition in the stock of vehicles.	225
C.15	The probability of under-reporting after acquiring 1033 items (All items).	230
C.16	The probability of under-reporting after acquiring 1033 weapons. . . .	231
C.17	The probability of under-reporting after acquiring 1033 vehicles. . . .	232
C.18	The probability of under-reporting after acquiring 1033 optics. . . .	233
C.19	The probability of under-reporting after acquiring 1033 clothes. . . .	234
C.20	The probability of under-reporting after acquiring 1033 other items. .	235
C.21	The probability of under-reporting after acquiring 1033 items (All items).	236
C.22	The probability of under-reporting after acquiring 1033 weapons. . . .	237
C.23	The probability of under-reporting after acquiring 1033 vehicles. . . .	238
C.24	The probability of under-reporting after acquiring 1033 optics. . . .	239

C.25 The probability of under-reporting after acquiring 1033 clothes. . . .	240
C.26 The probability of under-reporting after acquiring 1033 other items. .	241

CHAPTER 1

Introduction

This dissertation consists of three essays in public economics and the political economy of the public sector. The three studies are empirical and exploit state-of-the-art econometrics to estimate the impact of two federal programs: the earned income tax credit (EITC), and the Law Enforcement Support Office's (LESO) 1033 Program. Each essay contributes an original analysis on the impact of public policies on households, firms, and government.

The Earned Income Tax Credit (EITC) is one of the largest social welfare programs in the United States. The EITC is a refundable tax credit aimed at working taxpayers and can be considered as a conditional cash transfer to low-income households. A refundable tax credit is one that may result in negative tax liabilities, which in turn entitles taxpayers to a refund. The amount of the transfer is determined by family earnings, the number of children, and marital status. The credit gradually increases (phase-in) until it reaches a maximum (plateau), after which it decreases (phase-out) with increasing pre-tax earnings, until the household no longer qualifies for welfare. The first two chapters of this dissertation analyze two dimensions of the program. Chapter 2 evaluates the long-term impact of the EITC on the adult income of individuals who were exposed to the program during childhood. Chapter 3 studies the impact of the program on firms and the labor market using establishment surveys.

Research on the tax credit nurtures public policies because cash transfers are a popular alternative to compensate for adverse shocks to households, either due to policy reforms or technological changes that upset a vulnerable share of the population. This translates to developing economies aiming to eradicate distortionary policies by compensating target populations, and developed countries seeking to offset technological shocks.¹ With this in mind, policy-makers and scholars have a growing interest in the costs and benefits of conditional and unconditional cash transfers

¹For an example of compensatory policies, see the case of the removal of fossil fuel subsidies in Iran (Zarepour and Wagner, 2022; Salehi-Isfahani et al., 2015; Guillaume et al., 2011) and proposals in Venezuela (Rodríguez et al., 2012). The case of technological shocks is perhaps best represented by the negative labor market effects following artificial intelligence and robotics (De Stefano, 2019; Bruun and Duka, 2018).

([Hoynes and Rothstein, 2019](#); [Ghatak and Maniquet, 2019](#)). The EITC is part of the former, a means-tested program that conditions the magnitude of the transfer on the earnings and size of the household. The importance of this line of research is enhanced by the fact that the EITC is one of the main policy tools to combat poverty in America. In 2021 the Internal Revenue Service (IRS) transferred tax credits totaling \$60 billion or \$2,411 per average taxpayer, making the EITC one of the largest means-tested programs ([IRS, 2022](#)).

My empirical investigation is motivated by the additional effects derived from labor market responses to the policy. The key takeaway is that the EITC effectively targets its intended population by drawing non-working individuals into the labor market, but the specific eligibility conditions result in a range of heterogeneous effects ([Eissa and Hoynes, 2006](#); [Meyer, 2002](#); [Meyer and Rosenbaum, 2001](#)). To begin with, as the EITC targets families with children, I build on ([Bastian and Micheltore, 2018](#)) and analyze the evidence of intergenerational effects on the adult earnings of recipient children. Additionally, due to the expected impact of the policy on the labor market, I also estimate the policy impact on businesses. These analyses offer a comprehensive understanding of the various effects of the EITC on both individuals and businesses, thus contributing to the ongoing economic literature on the behavioral responses to welfare programs.

Chapter 4 studies information asymmetries in state and local law enforcement departments that participate in the 1033 Program, a federal grant that supplies excess military equipment to police agencies. The Law Enforcement Support Office (LESO) manages the program since its inception in 1997, distributing over \$7.6 billion worth of equipment to participating agencies. Scholars have examined how the 1033 Program affects various outcomes, including civilian complaints, use of force, police fatalities, and crime rates ([Masera, 2021a](#); [Gunderson et al., 2021](#); [Harris et al., 2017](#); [Bove and Gavrilova, 2017](#)). Military hardware may be an input to violent behavior and homicides of civilians, but it is unclear how military gear affects reporting behavior by police departments.

The research question is motivated by the discrepancy between police reports and third-party sources of civilian fatalities at the hands of law enforcement officers. While the Federal Bureau of Investigation (FBI) reports approximately 400 civilian deaths per year, third-party sources indicate the number could be as high as one thousand or more homicides.² Furthermore, studies on the impact of section 1033 on police killings of civilian are highly sensitive to these issues in data quality. With the same methodology, acquisition of military equipment has a causal effect on killings by police when using third party sources (Masera, 2021a) but a negligible impact with self-report data (Harris et al., 2017). This paper documents police reporting behavior and estimates the empirical effect of police militarization on the gap between self-reported and third-party statistics.

²The Department of Justice (DoJ) estimates that only half of the killings by police are reported, with much variation across states (Banks et al., 2015). Recent tracking by The Washington Post records only a third of fatal police encounters are reported by police agencies. See <https://www.washingtonpost.com/investigations/interactive/2022/fatal-police-shootings-unreported>

References

- Banks, D., Couzens, L., Blanton, C., and Cribb, D. (2015). Arrest-related deaths program assessment: Technical report. Technical report, Department of Justice, Office of Justice Programs, Bureau of Justice Statistics.
- Bastian, J. and Micheltore, K. (2018). The long-term impact of the earned income tax credit on children’s education and employment outcomes. *Journal of Labor Economics*, 36(4):1127–1163.
- Bove, V. and Gavrilova, E. (2017). Police officer on the frontline or a soldier? the effect of police militarization on crime. *American Economic Journal: Economic Policy*, 9(3):1–18.
- Bruun, E. P. and Duka, A. (2018). Artificial Intelligence, jobs and the future of work: Racing with the machines. *Basic Income Studies*, 13(2).
- De Stefano, V. (2019). Negotiating the algorithm: Automation, Artificial Intelligence and labour protection. *Artificial Intelligence and Labour Protection (May 16, 2018). Comparative Labor Law & Policy Journal*, 41(1).
- Eissa, N. and Hoynes, H. W. (2006). Behavioral responses to taxes: Lessons from the EITC and Labor Supply. *Tax Policy and the Economy*, 20:73–110.
- Ghatak, M. and Maniquet, F. (2019). Universal Basic Income: some theoretical aspects. *Annual Review of Economics*, 11:895–928.
- Guillaume, D. M., Zyttek, R., and Farzin, M. R. (2011). Iran: The chronicles of the subsidy reform. *IMF Working Papers*, 2011(167).

- Gunderson, A., Cohen, E., Schiff, K. J., Clark, T. S., Glynn, A. N., and Owens, M. L. (2021). Counterevidence of crime-reduction effects from federal grants of military equipment to local police. *Nature Human Behaviour*, 5(2):194–204.
- Harris, M. C., Park, J., Bruce, D. J., and Murray, M. N. (2017). Peacekeeping force: Effects of providing tactical equipment to local law enforcement. *American Economic Journal: Economic Policy*, 9(3):291–313.
- Hoynes, H. and Rothstein, J. (2019). Universal Basic Income in the United States and advanced countries. *Annual Review of Economics*, 11:929–958.
- IRS (2022). EITC Reports and Statistics. <https://www.irs.gov/credits-deductions/individuals/earned-income-tax-credit/eitc-reports-and-statistics>. Internal Revenue Service.
- Masera, F. (2021). Police safety, killings by the police, and the militarization of us law enforcement. *Journal of Urban Economics*, 124:103365.
- Meyer, B. D. (2002). Labor supply at the extensive and intensive margins: The EITC, welfare, and hours worked. *American Economic Review*, 92(2):373–379.
- Meyer, B. D. and Rosenbaum, D. T. (2001). Welfare, the earned income tax credit, and the labor supply of single mothers. *The Quarterly Journal of Economics*, 116(3):1063–1114.
- Rodríguez, P. L., Morales, J., and Monaldi Marturet, F. (2012). Direct distribution of oil revenues in Venezuela: A viable alternative? *Center for Global Development Working Paper*, (306).
- Salehi-Isfahani, D., Wilson Stucki, B., and Deutschmann, J. (2015). The reform of energy subsidies in Iran: The role of cash transfers. *Emerging Markets Finance and Trade*, 51(6):1144–1162.

Zarepour, Z. and Wagner, N. (2022). Cash instead of subsidy: Assessing the impact of the iranian energy subsidy reform on households. *Energy Policy*, 168:113145.

CHAPTER 2

The Intergenerational Effects of the EITC on the Adult Income of Children

The Earned Income Tax Credit (EITC) is a major component of the social safety net in the United States, improving the incomes of low-income families via targeted cash transfers contingent on employment. With an annual expenditure of \$60 billion, it ranks among the largest means-tested program in the country.¹ Most studies document the immediate or short-term effects of EITC receipts, leaving an important gap in our understanding of the program’s long-term and intergenerational implications. However, the long-term effects are critical for assessing the policy effectiveness in advancing mobility and mitigating structural poverty. [Bastian and Micheltore \(2018\)](#) presents the first causal evidence of improvements in long-term earnings and human capital among children exposed to the tax credit. In this paper, I build upon and expand this work, leveraging new survey data and a novel source of variation to estimate the effects of the tax credit on the adult income of individuals from families who were eligible for the program during their childhood. My empirical investigation focuses on identifying the optimal timing for investing during childhood and the interaction between maternal labor supply and the intergenerational effects of the program.

The intergenerational effects of EITC transfers are ambiguous *a priori* due to heterogeneous responses of parents to eligibility conditions. On the one hand, income effects may alleviate budget constraints within families, improving the health and home environment of children and their mothers ([Dench and Joyce, 2020](#); [Averett and Wang, 2018](#); [Hoynes et al., 2015](#)), and raising academic achievement and the likelihood of college enrollment for children ([Bastian and Micheltore, 2018](#); [Manoli and Turner, 2018](#); [Chetty et al., 2011](#)). On the other hand, the program affects

¹In 2021 the Internal Revenue Service (IRS) transferred \$60 billion in income tax credits or \$2,411 per recipient ([IRS, 2022](#)). Medicaid, the largest means-tested program in the United States, had costs of \$730 billion in medical assistance programs and \$30 billion more in administrative costs during 2021 ([Centers for Medicare & Medicaid Services, 2022](#)). Total transfers and administrative costs from the Supplemental Nutrition Assistance Program (SNAP) reached \$111 billion in 2021 due to temporary pandemic relief, although in previous years the program budget was \$63 billion ([CBPP, 2022](#)). The Supplemental Security Income (SSI) or Social Security amounts to \$56 billion in spending ([Social Security Administration, 2021](#)). Meanwhile, Housing Assistance adds to \$40 billion in Federal grants to local and state expenditures, excluding tax benefits like the mortgage interest deduction ([Urban Institute, 2022](#)).

time allocation in the household because eligibility conditions incentivize maternal employment (Eissa and Hoynes, 2006, 2004; Meyer, 2002). This substitution effect results in mothers shifting from childcare to the workforce, increasing childcare spending and allowing providers to capture part of the subsidy (Pilkauskas and Micheltore, 2023; Micheltore and Pilkauskas, 2021; Rodgers, 2018). Furthermore, the question of whether increased labor force attachment among mothers adversely impacts children remains contentious (Caetano et al., 2021; Blau and Currie, 2006).²

Trade-offs arising from income and substitution effects may lead to seemingly contradictory outcomes in the observed impacts of the program. Prior studies indicate that EITC exposure received during adolescence has its greatest impact on long-run income and educational achievement, with negligible effects from exposure in earlier childhood periods (Bastian and Micheltore, 2018). However, localized estimates show that children have higher adult earnings if they are born shortly before the eligibility cutoff of January 1st for tax credits and deductions targeted at families with children (Barr et al., 2022). The later finding closely aligns with the existing literature on the importance of early childhood investments (Carneiro et al., 2021; Doyle et al., 2017; Cunha and Heckman, 2007) and raises the question about the optimal timing of investments. The EITC’s impact during early childhood is valuable for policymakers designing complementary policies and advances the economic understanding of parental investment in children (Carneiro et al., 2021). In addition, it may help clarify whether the intergenerational benefits of early childhood cash transfers are offset by eligibility requirements that encourage mothers to join the workforce soon after birth (Kuka and Shenhav, 2020).

In this paper, I revisit the prevailing evidence on how the EITC affects the adult income of individuals whose families were eligible for the program during their childhood. The empirical investigation leverages restricted access data from the

²Inspired by these issues, studies have attempted to define the effects of the program on parental investment in children. Bastian and Lochner (2020) characterize maternal time use into parental investment or quality time with children and other activities. There is no evidence that exposure to tax credits reduces the quality time mothers spend with their children .

National Longitudinal Survey of Youth of 1979 Children & Young Adult supplement (CNLSY).³ This longitudinal survey data has rich labor market records up to age 40, and family characteristics during childhood originating from maternal characteristics in the NLSY. I calculate tax credit transfers across the life cycle of children by imputing tax credit values calculated with TAXSIM using family income, number of children, maternal marital status, state, and year.⁴ Endogenous family responses to taxes are mitigated with plausibly exogenous variation in the tax credit schedule keeping household responses constant, exploiting cross-state and cross-year formulaic changes localized at the levels of income. Finally, the paper tests the hypothesis that children whose maternal labor supply increases due to the policy experience less impact than those whose mothers' work fewer hours. To analyze this mechanism, I estimate the intergenerational effects of the program across the distribution of work hours of mothers.

Causal identification is challenged by the endogenous nature of earnings and work hours, as households adapt their labor choices to maximize tax benefits (Saez, 2010). Numerous studies circumvent the issue by relying on policy exposure, which varies with fertility, year, and state of residency but is independent of income and marital status (McInnis et al., 2023; Bastian and Lochner, 2020; Micheltore and Lopoo, 2021a; Braga et al., 2020; Bastian and Micheltore, 2018). EITC exposure, while useful for guiding policymakers, offers a lower-bound intent-to-treat estimate by assuming that untreated units are affected by the maximum tax credit transfer available according to fertility across states and years.⁵ Furthermore, within this

³Sometimes referred as the Children of the NLSY. The survey longitudinally follows the children born from mothers that were surveyed in the original 1979 NLSY cohort.

⁴NBER TAXSIM is a popular tax calculator in the public finance literature, accounting for federal and state regulations (Feenberg and Coutts, 1993).

⁵Also, there remain concerns that exposure may be affected by endogenous migration and fertility. Dynamic complementarities during childhood could amplify the influence of these unobservables. For instance, exposure to the tax credit may impact domestic violence (Cesur et al., 2022; Edmonds et al., 2022), which has long-term impacts on child development. Migration patterns, fertility, and unobserved changes in parenting are also debated in the literature (Bastian and Lochner, 2020; Micheltore and Lopoo, 2021a; Micheltore and Pilkauskas, 2021; Pilkauskas and Micheltore, 2019; Baughman and Dickert-Conlin, 2009, 2003).

dataset, policy exposure shows no statistical association with adult income while imputed tax credits consistently identify early childhood effects.⁶

To address endogeneity issues, I introduce mother fixed effects and construct a simulated instrument that captures policy variation independent of any household responses, allowing the EITC parameters to vary across families, states, and over time while holding household characteristics constant.⁷ There are advantages to this approach, as it exploits local variation according to the family’s tax credit schedule. I graphically show that the policy variation instrument behaves differently to imputed tax credits and exposure, with spikes in years of federal and state reforms that support its exogeneity. A battery of robustness tests supports that the instrument is strong, and the results are stable across different specifications and alternative instrument formulations.

The results show that children receiving the EITC have an annual adult income between 1.7 and 2.7 percent higher than the average for every \$1,000 in total tax credits received during early childhood. The effects are consistent in both OLS and 2SLS analyses and are robust across different specifications. At face value, the intergenerational effects of the program fall within the bounds of previous estimates (Barr et al., 2022; Bastian and Micheltore, 2018). These effects are larger than the long-run benefits of SNAP and comparable to those of Medicaid. The intergenerational effects of the EITC on long-run income are statistically significant during early childhood, as opposed to effects in late childhood, which remain small and insignificant across specifications. The early childhood coefficients are in line with Barr et al. (2022), where adult income is 1 to 3 percent higher for babies born

⁶OLS estimates of EITC exposure on adult income are shown in the Appendix A.3. I complement the analysis with 2SLS estimates of EITC exposure instrumented with policy variation, as explained in the text.

⁷This instrument gained popularity following Gruber and Saez (2002) and draws inspiration from Cummins et al. (1994). It is particularly useful for studying federal policies implemented uniformly across geographies and lacking a balanced control group. The methodology has been applied in studies of Medicaid (Currie and Gruber, 1996), marginal tax rates (Gruber and Saez, 2002; Bruce, 2002), and more recently, SNAP (Leung and Seo, 2023) and the child care tax credit (Rodgers, 2023). The EITC has also been analyzed using this approach, especially regarding its impact on marriage (Dickert-Conlin and Houser, 2002).

around the cutoff date that makes them eligible for tax benefits during their first year of life. Moreover, the analysis suggests dynamic complementarities, indicating that intergenerational gains increase with additional years of transfers.

The mechanism analysis provides suggestive evidence that the intergenerational effects of the tax credit are offset by increases in maternal labor supply. Given that the US safety net primarily targets working mothers, the program encourages workforce attachment among mothers shortly after birth among those with toddlers (Micheltmore and Pilkauskas, 2021; Kuka and Shenhav, 2020), leading to larger spending in childcare arrangements (Pilkauskas and Micheltmore, 2023; Micheltmore and Pilkauskas, 2021). I document that the intergenerational effects of the program are null for children of mothers above the first quintile and above the median of hours worked, while positive and significant effects prevail among children whose mothers worked fewer hours during their early childhood. In the same spirit, the heterogeneity analysis reveals that the children that benefit the most from the program are white boys with married mothers. The heterogeneities by race and marital status of mothers are consistent with the hypothesis that increased maternal labor supply drives weaker long-run effects for children: research finds that the labor supply of black mothers is highly responsive to tax credits (Jiao et al., 2022), and married mothers are able to reduce their labor force participation and rely on their spouses as the primary earners (Eissa and Hoynes, 2004). Lastly, there is a broad consensus among researchers that boys are more sensitive to household resources than girls (Autor et al., 2019; Chetty et al., 2016; Bertrand and Pan, 2013).

This paper has three contributions to the economics literature. First, this paper contributes a novel mechanism analysis that sheds light on the interaction between maternal labor supply and the long-term intergenerational effects of the EITC. This adds to research on maternal labor supply and skill development in children (Caetano et al., 2021; Blau and Currie, 2006), and studies on the relationship between childcare arrangements and tax credits (Pilkauskas and Micheltmore, 2023; Micheltmore and Pilkauskas, 2021; Rodgers, 2018). To the best of my knowledge, I provide the first

evidence that imposing work conditions to social programs mitigates or offsets long-term anti-poverty effects.

Second, I contribute to the growing body of research studying the relationship between childhood resources and the long-term outcomes of low-income individuals. This paper adds new evidence leveraging novel variation localized in the segment of the tax schedule specific to families, rather than the program’s plateau. Instead, previous estimates find null effects in early childhood using policy exposure at the plateau ([Bastian and Micheltore, 2018](#)), and positive effects based on regression discontinuity designs around the cutoff birth date determining eligibility for tax benefits and deductions for children ([Barr et al., 2022](#)). Additionally, my data is well-suited to studying targeted families, which are over-represented in the sample. My empirical approach is idoneous as it mitigates endogenous responses in parental behavior and is relevant to the tax schedule of low-income families.

Lastly, this paper contributes to public economics literature evaluating the EITC. Existing studies explore the program’s long-term impacts on maternal labor markets ([Neumark and Shirley, 2020](#)), and children’s education ([Bastian and Micheltore, 2018](#); [Manning, 1990](#)), health ([Braga et al., 2020](#)), and fertility and early marriage ([Micheltore and Lopoo, 2021a](#)). In these instances, the predominant source of variation is the exposure to the maximum tax credit transfers available across years and states. Difference-in-difference frameworks are also frequent in this field following major program reforms ([Hoynes et al., 2015](#); [Eissa and Hoynes, 2006, 2004](#)). In contrast, my research exploits variation in policy rules throughout the entire tax credit schedule, which depends on the socioeconomic characteristics of families, as well as cross-state and annual policy variations. In the EITC literature, similar variation was used to study academic achievement ([Dahl et al., 2009](#)) and marriage ([Dickert-Conlin and Houser, 2002](#)).

The importance of these results is accentuated by the rise of cash transfers as a policy alternative and more stringent work requirements to access the social safety net. It remains contentious whether work incentives in means-tested programs actually

promote employment (Gray et al., 2023; Han, 2022; Fetter and Lockwood, 2018; Hoynes and Schanzenbach, 2012; Autor and Duggan, 2007; Besley and Coate, 1992) but there is agreement that the EITC has positive employment effects among single women –not without recent debate (Schanzenbach and Strain, 2021; Kleven, 2023). Mothers fulfill a key role as caretakers (Caetano et al., 2021; Blau and Currie, 2006) and yet workfare programs drive them to the workplace. Changes in time allocation affect parenting, leading to increased spending patterns in childcare (Pilkauskas and Micheltore, 2023; Micheltore and Pilkauskas, 2021; Kuka and Shenhav, 2020; Rodgers, 2018), and implying that workfare programs have unknown effects on children and their long-term development.

There are substantial benefits to incentivize work among women, particularly in reducing gender inequality and improving the labor outcomes of women (Kuka and Shenhav, 2020). Nevertheless, comprehensive policy-making must address complementary programs that support maternal labor supply while providing high-quality alternatives to childcare that boost intergenerational outcomes (Karademir et al., 2024). Further research should examine complementarities between workfare programs and affordable access to quality daycare and parental leave during early childhood.

2.1 The Earned Income Tax Credit

The EITC is a workfare program supporting low-income families while incentivizing work through a wage subsidy.⁸ The EITC has been a refundable cash transfer policy targeting low-income families since 1975 and has undergone nine reforms in the last four decades. Taxpayers apply by filing regular taxes (line 27a in Form 1040) and a form disclosing all qualifying children. The transfer amount is determined by the

⁸An excellent literature review is available in Hotz and Scholz (2001) but much has been written since then. The EITC emerged amidst a broader debate on the eligibility conditions to receive welfare, proposed against a Negative Income Tax and unconditional cash transfers. The ongoing momentum towards a Universal Basic Income (UBI) and other cash transfer programs signals this discussion remains current (Hoynes and Rothstein, 2019; Banerjee et al., 2019).

taxpayer’s adjusted gross income, the number of children, marital status, and the state of residency. Every January, the IRS publishes a tax schedule featuring three regions: a phase-in that increases linearly subsidizing pre-tax earnings for any income above zero to incentivize non-working taxpayers to seek employment; a plateau where the benefit is at its maximum; and a phase-out that gradually reduces the subsidy. Additionally, households can claim larger benefits for each of up to three children. To qualify, children must be younger than 19 years old or full-time students younger than 24 years old, reside in the household for at least half the year, and the claimant must be their parent, grandparent, or foster parent. The phase-out range is determined by the taxpayers’ marital status. Lastly, residency affects the benefits due to complementary state policies that match a percentage of the federal benefit.⁹

[Crandall-Hollick \(2018\)](#) categorizes the history of the program in three major periods each featuring their own policy goals. During the 70s and 80s, Senator Russell Long, in opposition to unconditional cash transfers, introduced the policy as a “work bonus” aiming to reduce welfare dependence. In the second stage, policy-makers focused on targeting transfers on the basis of income, and emphasized poverty alleviation by increasing the magnitude of the credit and introducing benefits to childless workers. This was the advent of the 1990s Welfare Reform marked by The Omnibus Reconciliation Act of 1990 (OBRA), OBRA 1993, and the Personal Responsibility and Work Opportunity Act of 1996 (PRWORA). The expansion into a poverty alleviation program came along with concerns over fraud and noncompliance, leading to a new phase-out range to move high earners out of entitlements. Although the reform limited eligibility for high earners, [Hotz and Scholz \(2001\)](#) argue the policy increased progressivity, resulting in effective marginal tax rates of negative 30 percent for low-income earners by the end of the 1990s. In its third period, Congress attempted to address concerns about marital status and family size. The Economic Growth and Tax Relief Reconciliation Act of 2001 (EGTRRA) progressively reduced marriage penalties from two single parents entitled to more tax credit benefits than

⁹An illustrative tax credit schedule is shown in appendix [A.1](#).

a married couple. In the same decade, the American Recovery and Reinvestment Act (ARRA) increased the phaseout threshold for married couples (2008), and the program’s schedule grew in generosity for families with three or more children.

In 1984, Wisconsin instituted a program to complement EITC transfers with a percent of the Federal tax credits for EITC eligible families. Today, twenty seven states offer state tax credits matching up to 75 percent of the Federal EITC, depending on family characteristics and state laws. For example, South Carolina adopted the policy as late as 2018 matching up to 20 percent of the federal transfers but is set to progressively expand the program to 125 percent by 2023. Figure 2.1 pictures the time and state-level variation in the program. The solid line is the maximum Federal tax credit while the maximum state transfers are in grey. For states that provide the program, benefits range between \$200 and \$2,572 with changes in the percent matched over time. For the purpose of this paper, the combined Federal and state transfers in Figure 2.1 visualize cross-state and annual variation core to the identification strategy described next.

2.2 Data

This study uses data from the the Bureau of Labor Statistics’ National Longitudinal Survey of Youth (NLSY) and the Children and Young Adult (CNLSY) supplement from 1979 to 2018. In 1979, the BLS began interviewing one cohort of 12,686 youths between the ages of 14 and 22. This longitudinal panel has continued for over 40 years and these individuals are now adults with their own families. Beginning in 1986, the BLS began interviewing 11,545 children born to 4,943 mothers in the original NLSY. Households were interviewed annually until 1994, and biannually thereafter. This dataset is known as the CNLSY. The NLSY provides rich socioeconomic and psychological characteristics of respondents, and the CNLSY records labor market outcomes of their children. The restricted-access supplement provides the state of residency for all interviewees during each year.

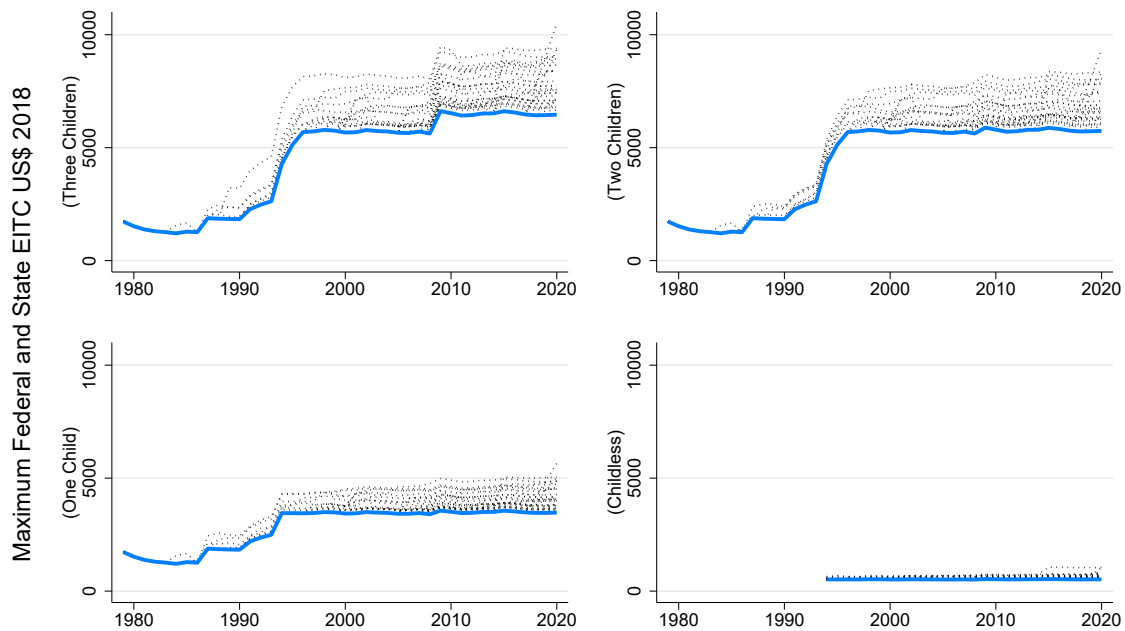


Figure 2.1: Maximum EITC benefit by number of dependents.

Notes: The figure plots the maximum real tax credit transfers available adding the Federal and state benefits per year in the plateau of the tax credit schedule. The solid line is the Federal tax credit benefit while grey dotted lines are the combined federal and state tax credits in each state. The upper panel shows the benefits for families with three and two children. The bottom panel shows the benefits for families with one and no children.

I match the adult labor market characteristics in the CNLSY with the NLSY to obtain family background during childhood. The final sample consists of 5,253 children from 2,740 unique mothers.¹⁰ Descriptive statistics are presented in Table 2.1.¹¹ The first panel illustrates the average income from salaries and wages for adults who were children in the CNLSY. Income is measured after the child is 21 years old, for comparability with previous research. We observe that income is increasing in age from 14 thousand dollars in their twenties to 30 thousand dollars when individuals reach their forties, in line with stylized facts in the age-profile of earnings (Mincer, 1974). The NLSY is characterized by oversampling of Hispanics, economically disadvantaged non-blacks and non-Hispanics. Consequently, the adult income of individuals is 70 percent of median personal income in 2018 (38 thousand dollars).

Family characteristics during childhood are measured at age 18 in the second panel of Table 2.1. The average child in the sample lived in a family making a combined income of 66 thousand by the time they were 18.¹² Oversampling is further observed by the racial composition of the sample: a third of individuals are black children and a fourth are Hispanic. Families have 3 children on average, while 90 percent of the

¹⁰Sample restriction criteria follow from Bastian and Micheltore (2018). First, maternal socioeconomic characteristics are observed at least once when the child is between the ages of 0 to 5 (early childhood), 6 to 12 (mid childhood), and 13 to 18 (late childhood). This is unavoidable in order to calculate tax credit benefits during childhood. Second, the sample is restricted to children who respond to their adult earnings at least once between the ages of 22 and 27. The analytical sample includes individuals without reported earnings, but those with persistent attrition during their adulthood are excluded. Non-interview years are imputed via linear interpolation.

¹¹All currency amounts in this paper are adjusted for inflation in 2018 dollars using the national Consumer Price Index.

¹²Family income, a key variable in calculating tax credits, is estimated by adding the income from salary and wages for the mother and the spouse. The NLSY also has a net family income variable available but its construction has some recorded inconsistencies. First, total net family income includes compensation from salaries and wages, military income, net business income, farm income, unemployment compensation, and other forms of welfare benefits and scholarships. Second, net family income does not include partners' income and earnings until 1994 and consisted of separate questions before. As such, and to avoid the inclusion of government transfers into the estimation of tax credits, I construct family income by adding a couple's income from salary, wages, and tips.

sample graduated high school and 47 percent received some college education. In the sample, 88 percent of mothers have ever been married by the time the child is 18.¹³

In the table we further show the composition of cohorts in the CNLSY using the birth year of children. While the NLSY-79 consists of one cohort, the CNLSY encompasses three birth cohorts beginning in the 1970s. Two thirds of the sample were born in the 1980s, allowing for a large degree of exposure to the EITC. A quarter of individuals are born in the 1990s while only 8 percent are from the earliest 1970s cohort. The cohort-structure implies that, because the CNLSY interviews all children born from women in the original NLSY survey, the sample has 23 percent of children without siblings, 40 percent of sibling pairs, 22 percent of children with two other siblings in the sample, and families with up to 11 siblings. In consequence, later cohorts tend to be siblings of earlier cohorts, which may entitle them to larger tax credits or affect endogenous labor market outcomes.¹⁴

The composition of this data has several distinguishing features from previous research that relied on the Panel Survey of Income Dynamics. First, the CNLSY has a larger analytical sample compared with research using the PSID ([Bastian and Micheltore, 2018](#)). Second, unlike the PSID, the CNLSY is not a nationally representative sample but may be more adequate to study low-income families due to the oversampling of economically disadvantaged households.¹⁵ Average family income in the analytical sample is 18 thousand dollars lower than in the PSID.¹⁶ The PSID includes 18 percent black families and 1 percent Hispanics, in comparison with the 32 and 23 percent over-sample in the CNLSY. The sample of mothers in [Bastian and](#)

¹³This may be a high marriage rate but it may find its explanation in the birth year of mothers belonging to the 1979 NSLY cohort. These mothers are born between 1957 and 1964. By the time women in the NLSY-79 were of age, over 70 percent of women in America were married perhaps explaining a large proportion of ever-married mothers.

¹⁴See the distribution of tax credits by cohort in the appendix [A.1](#).

¹⁵The CNLSY is susceptible to nonrandom migration and fertility patterns among NLSY respondents ([Cooksey, 2018](#)). Therefore, the sample predominantly includes low-income households and siblings. The literature and NLSY manuals emphasize the potential benefits for studying low-income households and the safety net, but these features are likely to drive different findings with respect to the established literature.

¹⁶Average family income in the PSID is 77.76 thousand 2013 dollars or 88.8 thousand 2018 dollars ([Bastian and Micheltore, 2018](#)).

[Micheltmore \(2018\)](#) are also more educated, with 5 percent of high school dropouts and 70 percent that have some college education, and has the same rate of ever-married mothers by the time children are 18 (88 percent). Lastly, differences in the cohort structure imply that children in later cohorts of the CNLSY are not comparable with later cohorts in the PSID. However, the presence of siblings allow for mother fixed effects that control for unobservable family characteristics that cannot be accounted for in the cross-sectional structure in [Bastian and Micheltmore \(2018\)](#).

I calculate two measurements of EITC benefits summarized in Table 2.2.¹⁷ First, cumulative EITC exposure at each childhood stage represents the maximum available federal and state benefit given the year, state of residency, and number of children. The cumulative measure adds the value across a given age range, capturing annual variation in the tax credit schedule, and variation based on state of residency and number of siblings, independent of family income and marital status following [Bastian and Micheltmore \(2018\)](#).¹⁸ In the sample, the mean child was exposed to \$14 thousand dollars between ages 0 and 5, 27 thousand during mid childhood, and 30 thousand during teenage years. Total EITC exposure for the average child was of 71 thousand dollars.

Second, I use TAXSIM to calculate imputed EITC as the combined federal and state benefits the family is eligible for given the year, state of residency, number

¹⁷The NLSY does not accurately or consistently record the tax credits received by a household.

¹⁸EITC exposure has been the preferred measurement of tax credit benefits in much of the literature. For example, see [McInnis et al. \(2023\)](#); [Bastian and Black \(2022\)](#); [Micheltmore and Pilkauskas \(2021\)](#); [Micheltmore and Lopoo \(2021a,b\)](#); [Braga et al. \(2020\)](#); [Bastian and Lochner \(2020\)](#). In addition, a branch of the literature evaluates the program using difference-in-difference estimators that compare women with and without children ([Bastian, 2020](#)) or women with different educational levels ([Eissa and Hoynes, 2006, 2004](#)). There is much debate on these methods. In a working paper, [Kleven \(2023\)](#) suggests that EITC expansions are usually correlated with other social programs, and these estimates confound other components of the welfare system. Following this discussion, [Schanzenbach and Strain \(2021\)](#) disagrees with the inclusion of business cycle controls as these fluctuations have heterogeneous effects on single mother with children that are targeted by the EITC. [Bastian \(2017\)](#) criticism is of a different vein, the working paper suggests that traditional treatment and control groups in the EITC are unstable, due to the endogeneity of marital status and fertility. EITC exposure, which is estimated based on the number of children in the family, shares this weakness with difference-in-difference estimates.

of children, marital status, and their earnings.¹⁹ Imputed tax credits are a reliable approximation of actual EITC benefits. Recent analyses compare IRS records with imputed tax credit values, suggesting that childless taxpayers tend to over-report their earnings to the IRS (Jones and Ziliak, 2022). Therefore, imputed values are likely representative for my sample restricted to mothers. To date, IRS data finds that 80 percent of eligible households claim the EITC.²⁰ The average child in the analytical sample was eligible to receive 2 thousand dollars during early childhood, five thousand dollars in mid childhood, and 14 thousand as a teenager. Total average tax credits during childhood add up to one thousand dollars. Imputed tax credits are naturally smaller than EITC exposure, as the former captures family income and the latter represents the maximum benefits available at the plateau (families in the phase-in and phase-out are eligible to fewer transfers than those at the plateau).

For illustrative purposes, consider the simplest family structure consisting of an only child born in 1990 in a state without the supplemental EITC program. Regardless of family income, this child will be exposed to the federal EITC program from childhood until their teenage years, amounting to \$60 thousand in tax credit transfers. Specifically, exposure would be \$16 thousand during early childhood, \$24 thousand in the subsequent seven years, and \$20 thousand in late childhood.

In contrast, determining imputed EITC values for this individual requires further assumptions about family income. In 1990, AGI had to be less than \$20,200 to be eligible for tax credits. Assume that this hypothetical child had family income of \$6 thousand dollars (the average for a child born in 1990 to parents earning less than the maximum AGI). As the only eligible dependent in the household, the child would be entitled to \$1,614 in his first year (nominal value of \$840). Assuming constant family characteristics throughout childhood, the family received \$13 thousand in

¹⁹Earnings are estimated with the combined income of the mother and her spouse. If the mother does not report being married, income from their partner is ignored for calculating program eligibility.

²⁰Take-up rates range from 71 percent to 87 percent across states. The most recent figures by each state are available on the IRS website <https://www.etc.irs.gov/etc-central/participation-rate/etc-participation-rate-by-states>.

early childhood, \$21 thousand between ages 6 and 12, and \$15 thousand during late childhood, resulting in \$50 thousand in EITC transfers across childhood.

2.3 Empirical Strategy

I seek to estimate the impact of EITC transfers received during childhood on adult earnings. Former studies exploit a cross-sectional design where variation in adult income is explained by variation in tax credit benefits across different childhood stages (Bastian and Micheltore, 2018). Early childhood is captured by the cumulative tax credit transfers received from age 0 to 5, mid childhood is represented by total transfers between age 6 and 12, and late childhood are is combined EITC transfers during adolescence. The estimating equation is:

$$Y_i = \beta_1 \text{EITC}_{i,(0-5)} + \beta_2 \text{EITC}_{i,(6-12)} + \beta_3 \text{EITC}_{i,(13-18)} + \eta X_i + \psi V_{s,t} + \lambda_s + \lambda_t + \epsilon_i \quad (2.1)$$

Equation 2.1 is the estimator for the i th individual's adult income Y , living in state s at age 18, and born in year t . The parameters of interest β_1, β_2 , and β_3 are interpreted as the marginal dollar-to-dollar association between total tax credits received in childhood stage a in $\text{EITC}_{i,(a)}$ and adult income. $\text{EITC}_{i,(a)}$ are the total tax credit transfers during the age interval a with each interval characterizing early childhood, mid childhood, and late childhood, respectively. The marginal effect of a \$1,000 change in stage a may be interpreted as a \$200 increase over a five year interval or a \$1,000 expansion in a single year.²¹

My main specification fully leverages the NLSY-79 by including mother fixed effects that account for unobservable family background. X_i is a vector that controls for family characteristics and fixed effects by race and gender interacted with the

²¹The results section discusses dynamic complementarities of cash transfers and nuances associated with the aggregation of childhood stages. I test the sensitivity of estimates to the length of interval a and show these are volatile when age is disaggregated to one year. The aggregate tax transfers during all childhood are also examined and show positive results, contrasting with previous research where total tax credits in childhood are insignificant (Bastian and Micheltore, 2018).

state and year. The term includes cohort fixed effects, quadratic and cubic age in 2018, indicators for black, Hispanic, female, whether the mother has been married by age 18, an indicator for whether the mother graduated high school, an indicator for whether the mother has some college education, and fixed effects for the number of siblings at age 18. The specification includes the three-way interactions between race, ethnicity, and sex by number of siblings, birth year and state guaranteeing that treatment variation is independent of family composition.²²

The term $V_{s,t}$ accounts for state-policy indicators measured at age 18 recommended in the literature, such as GDP per capita, unemployment rate, the top marginal tax rate, the minimum wage, tax revenues, higher education spending, and the maximum welfare benefits.²³ The specification further controls for state fixed effects measured at age 18 λ_s , and year of birth fixed effects λ_t controlling for idiosyncratic time variation across children. Standard errors are clustered at the state level.

There are several concerns related to endogenous transfers and long-term labor market outcomes in Equation 2.1. The main challenge in the literature relates to endogenous income required to calculate tax credits in $EITC_{i,(a)}$. The concern is driven by well-documented bunching in the EITC schedule (Mortenson and Whitten, 2020; Kleven, 2016; Chetty et al., 2013) which concentrates among the self-employed (Kuka and Shenhav, 2020; Saez, 2010). Additionally, maternal labor supply is, in theory and not without debate, impacted by the tax credit (Kleven, 2023; Eissa and

²²This specification follows Bastian and Micheltore (2018), however, I exclude indicators for fathers' education, which are not readily available in the CNLSY. The PSID surveys families while the CNLSY is a supplement that surveys children of women in the NLSY. Some sociodemographics, such as mother's characteristics, race, and ethnicity are redundant in the presence of mother fixed effects. Additionally, my results are robust to state-specific time trends that account for additional time varying state shocks suggested in Bastian and Micheltore (2018). I exclude time trends from the main results to avoid confounding the dynamic effects of the program (Wolfers, 2006).

²³GDP per capita is sourced from the Bureau of Economic Analysis. Total Tax Revenue data is retrieved from the Annual Survey of State Government Tax Collections (STC) by the Census Bureau. The unemployment rate, state minimum wage, and consumer price index are provided by the U.S. Bureau of Labor Statistics, with data available at the Federal Reserve Bank of St. Louis. Per capita spending on higher education is published by the State Higher Education Executive Officers (SHEEO). Lastly, maximum welfare benefits for a family with three children per state is found in the Urban Institute's Welfare Rules Database.

Hoynes, 2006, 2004) driving substitution effects in time allocation that decreases time spent with children (Bastian and Lochner, 2020) and may distort care-giving (Caetano et al., 2021; Rodgers, 2018; Micheltmore and Pilkauskas, 2021).

Research favors an empirical design that estimates the impact of EITC exposure, which is independent of income and absorbs cross-year and cross-state variation in the plateau and changes in the number of children in the family (Bastian and Micheltmore, 2018). Exposure, while useful to guide policy-making, assumes that untreated units are affected by the maximum available transfer and likely yields a lower bound intent-to-treat. Also, both exposure and imputed tax credits might be influenced by endogenous fertility and migration.²⁴ Discussions on the policy’s effect on fertility continue, with most recent evidence suggesting that state generosity in welfare increases the probability of having a child (Bastian, 2017; Baughman and Dickert-Conlin, 2009, 2003). Endogenous migration is understudied in the literature but there is a stark contrast in tax credits across states, especially in recent years with states becoming more generous. For instance, as of 2023, a family of three in Tennessee qualifies for a maximum of \$7,430 but moving a few miles entitles them to \$13,032 in South Carolina. Even if interstate migration is unaffected for low-income households, rural-urban migration spurred due to the influx of the tax credit program (Bastian and Micheltmore, 2018). Another issue is that the cross-sectional design prevents the inclusion of child fixed effects that could alleviate issues regarding unobserved patterns in childcare spending and childcare quality influenced by the program (Micheltmore and Pilkauskas, 2021; Pilkauskas and Micheltmore, 2019), as well as time allocation and home quality factors (Edmonds et al., 2022; Cesur et al., 2022; Bastian and Lochner, 2020; Averett and Wang, 2018).²⁵

²⁴Households may also respond to the program by changing their marital status. The literature remains inconclusive, suggesting small changes in the probability of marriage and divorce (Isaac, 2020; Bastian, 2017; Herbst, 2011; Dickert-Conlin and Houser, 2002). Marital status is omitted in the construction of EITC exposure but adds an additional layer of concern on imputed tax credits as marital status may influence the long-term development of children.

²⁵My specification includes mother-fixed effects, previously omitted in the literature, reducing these factors.

Table 2.1: Descriptive statistics.

	Mean	SD
Mean Adult Income (000s \$2018)		
Adult income ages 22-27	14.44	12.63
Adult income ages 22-30	20.20	15.96
Adult income ages 22-40	27.42	20.61
Child Characteristics at age 18		
Family income (000s \$2018)	65.97	71.90
Female	0.49	0.50
Black	0.32	0.47
Hispanic	0.23	0.42
Dependents	3.11	1.11
Mother graduated HS	0.90	0.30
Mother has some college	0.47	0.50
Mother ever married	0.88	0.33
Cohort 1970s	0.08	0.27
Cohort 1980s	0.66	0.48
Cohort 1990s	0.26	0.44
Observations	5253	

Notes: Summary statistics are computed with the CNLSY custom sample weights. All monetary values are in thousand real 2018 dollars. Minimum and maximum values are omitted in compliance with the BLS agreement. Mean adult income is the real value of the average adult income of children in the CNLSY a given age range, conditional on reporting income at least once during the period. Child characteristics are measured at age 18.

Table 2.2: Calculated EITC Exposure and Imputed EITC benefits in childhood.

	Mean	SD
EITC exposure between ages 0 and 5	13.51	7.62
EITC exposure between ages 6 and 12	27.48	9.99
EITC exposure between ages 13 and 18	30.79	4.67
EITC exposure between ages 0 and 18	71.79	19.30
Imputed EITC between ages 0 and 5	2.14	3.96
Imputed EITC between ages 6 and 12	5.46	8.17
Imputed EITC between ages 13 and 18	6.66	9.30
Imputed EITC between ages 0 and 18	14.25	17.23
Imputed EITC at age 18	1.13	1.90
EITC Eligible between ages 0 and 18	0.74	0.44
Observations	5253	

Notes: Summary statistics are computed with the CNLSY custom sample weights. All monetary values are in thousand real 2018 dollars. Minimum and maximum values are omitted in compliance with the BLS agreement. The EITC variables are cumulative values in each age range. EITC exposure is the total maximum federal and state benefit available to a household given the specific year, state of residency, and number of children. Imputed EITC is the total federal and state benefits the family is eligible for given the year, state of residency, number of children, marital status, and their earnings, calculated with NBER TAXSIM.

There are advantages to using imputed tax credits and a simulated IV strategy, mitigating endogenous household responses.²⁶ The primary advantage is the alignment between $EITC_{i,(a)}$ and the most likely transfers received by households given their income bracket. Calculating tax credits at localized income levels enables the identification of effects for low-income households targeted by the policy. Another benefit is that imputed tax credits can be instrumented using a methodology from the public finance literature designed to estimate localized treatment effects using income levels and keeping household responses constant. Factors such as marital status, fertility, state of residency, and family environment are all likely to be influenced by the program but are held constant using this methodology. By construction, the simulated instrument is only relevant to imputed tax credits and does not apply to policy exposure, which is not localized across the tax credit schedule.

The simulated IV builds on [Gruber and Saez \(2002\)](#) and assumes independence between policy changes in the EITC schedule and family responses. My goal is to estimate the variation in the EITC a family is eligible to receive due to policy changes, independent of potential behavioral responses from the household. The first step requires an approximation of family income captured by predicted family income $\hat{E}_{i,t}$ for individual i during childhood year t . I estimate predicted income using lagged income and sociodemographic characteristics of the mother and test whether results are robust to different functional forms of $\hat{E}_{i,t}$.²⁷

²⁶Imputed tax credits, as they are imputed values and not actual cash transfers received by families, are a measure of exposure as well. Variation, however, differs as imputed values are determined using income.

²⁷The literature on marginal tax rates recognizes two challenges in estimating policy variation. First, reversion to the mean implies that high income in a given year is likely followed by lower income in the upcoming years, generating a negative correlation. Second, area widening income distribution implies a positive correlation between tax policy and unobservables in the income. To alleviate these issues, researchers estimate predicted income using a spline of lagged income ([Dahl and Lochner, 2012](#); [Gruber and Saez, 2002](#)). I find that there are negligible differences in the distribution of family income between a linear approximation and a spline and test my results across four specifications of $\hat{E}_{i,t}$. First, my main estimates regress family income on lagged wages, year fixed effects, and mother fixed effects. Second, I predict family income using an autoregressive function. Third, income is regressed on a fifth-order spline of lagged income as in [Dahl and Lochner \(2012\)](#). And fourth, I use the mother's net family income from her adolescence in 1979, corresponding to the grand-father of i 'th's generation, and keep income fixed adjusting for inflation. The latter specification exploits

The following step involves calculating two simulated earned tax credit values. First, I simulate $\chi \left[\hat{E}_{it}, \varsigma_{it} | \tau_t \right]$: the tax credit χ the family would be eligible for with tax rules τ at time t given current predicted earnings $\hat{E}_{it}$ and current family characteristics ς_{it} (number of children, marital status, and state of residency). Second, the counterfactual tax credit the family would have been eligible for under tax rules τ of the previous period $t - 1$ while holding fixed current predicted income and family characteristics, or $\chi \left[\hat{E}_{it}, \varsigma_{it} | \tau_{t-1} \right]$. The difference between both values $Z_{it} \equiv \Delta\chi$ represents policy-induced variation in the EITC schedule by varying tax credit rules, *ceteris paribus*.

I illustrate the theoretical differences between exposure and the simulated instrument in Figure 2.2. The figure plots a microeconomic budget constraint framework (black solid line) experienced by a family making \$7 an hour and endowed with 16 hours per day and a 5-days work week. The head of household is single, has 3 eligible dependents, and lives in a state without a matching EITC supplement. Under 1993 rules, the blue line is the EITC schedule the family is eligible to: with a phase-in at the right of the budget constraint (when leisure hours are at the maximum), a plateau that begins around 3000 leisure hours, and the phase-out at the left of the budget constraint. The red line is the EITC schedule in 1994 after the reform and keeping these parameters.

The top panel in Figure 2.2 captures EITC exposure in 1994 following the tax reforms in OBRA 1993. EITC exposure is the dashed gold line atop the 1994 kinked budget constraint, where every individual living in this state and with 3 dependents is assigned variation from the maximum available tax credit in 1994. In the bottom panel in Figure 2.2, I plot again a budget constraint and the tax credit schedules for a household with equal characteristics. The simulated instrument is policy variation in 1994, calculated with fixed household characteristics and represented by the vertical difference between the 1994 tax credit schedule (red kinked constraint) and in 1993

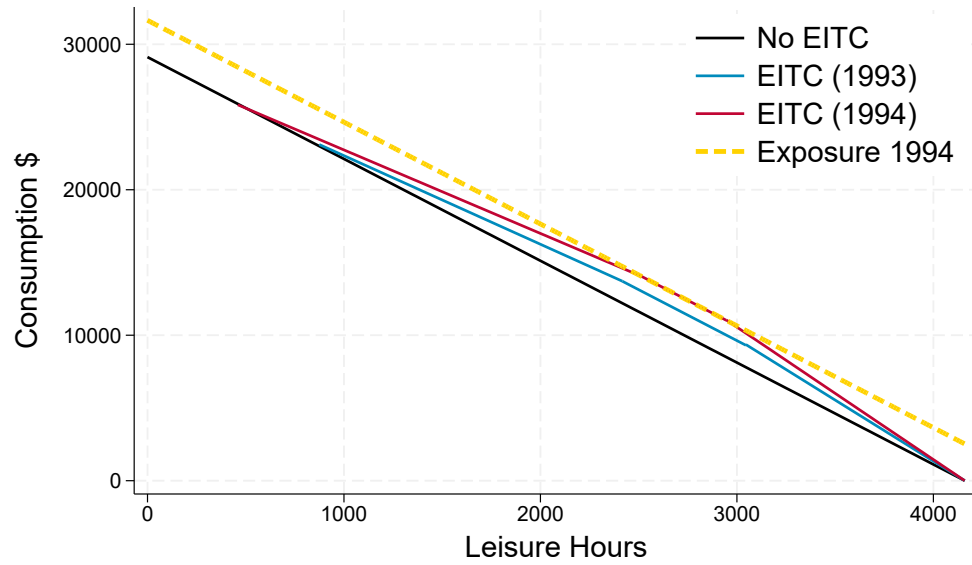
exogenous income from the mother's family but disconnects estimated policy variation from the EITC schedule of mothers. Appendix A.4 shows that the average of these instruments behave similarly.

(blue line). This vertical distance at any given point is drawn as black arrows in the graph, and shows the change in the tax credit schedule Z_{it} keeping work hours (predicted income) fixed across the horizontal axis as well as family characteristics, which remain as a single mother with 3 children in a fixed state. Because the schedule is changing but family characteristics are held constant, the vertical distance is the policy variation holding constant household responses to tax policy. The vertical distance between the 1994 and 1993 tax schedules is used to predict the imputed tax credits in 1994, calculated with observed instead of predicted income.

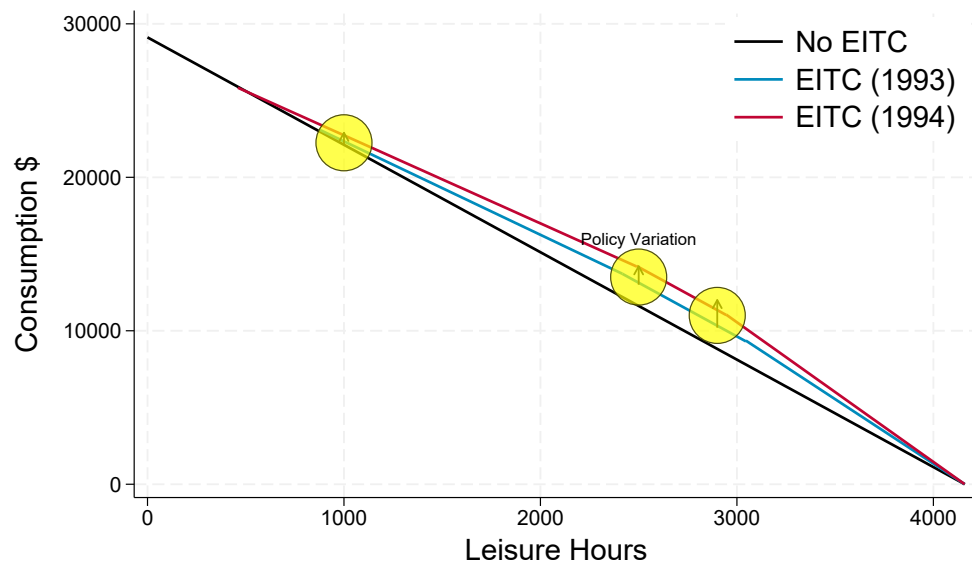
Notice that policy variation differs for individuals across segments of the tax credit schedule while exposure does not. Ineligible individuals may receive no variation from the policy unless their income is close enough to an expansion in the maximum AGI. The phase-in rate, phase-out rate, and plateau have also experienced changes across the program's history. It follows that the instrument assigns localized variation across the tax credit schedule instead of uniform exposure to all individuals in the same state and family composition. The instrument captures cross-time and cross-state variation and policy variation at the segment of the tax credit schedule for each individual.

Empirically, this instrument behaves differently to other tax credit measurements. Figure 2.3 normalizes and compares the average imputed EITC, EITC exposure, and the instrument in the sample.²⁸ The graph argues for the exogeneity of the instrument, showing that variation in the simulated IV is representative of reforms in the tax schedule instead of family characteristics. During the 1990s, a period marked by the Welfare Reform, policy variation exhibits spikes that are three times the magnitude of imputed tax credit transfers and exposure.

²⁸Due to differences in scale between tax credit transfers and policy variation, the graph standardizes the values relative to a baseline. Each line is an average in a year divided by mean transfers in all periods. The transformation allows for comparing these variables despite the fact that exposure averages in the thousands of dollars, average imputed values are hundreds of dollars, and policy variation tends to be smaller as it is the difference between two imputed tax credits. The sign of these variables also varies, as imputed tax credits and EITC exposure are always positive values but policy variation is potentially negative in the years when states reduce or eliminate their matching state tax credit programs. The moments of the distributions are shown in the appendix Table A.2.



EITC exposure in 1994.



The policy variation instrument in 1994.

Figure 2.2: The individual's budget constraint and the EITC schedule during the Welfare Reform.

Notes: The figure plots the budget constraint (black line) for an individual making 7 \$/hour. The blue and red lines are the kinks from an expanded budget constraint resulting from the EITC schedules in 1993 and 1994, respectively. The top panel plots the maximum available tax credit in 1994 with a dashed gold line (EITC Exposure). The bottom panel plots three arrows capturing the vertical distance between the tax credit benefits in 1994 and 1993, keeping pre-tax income fixed in three different levels.

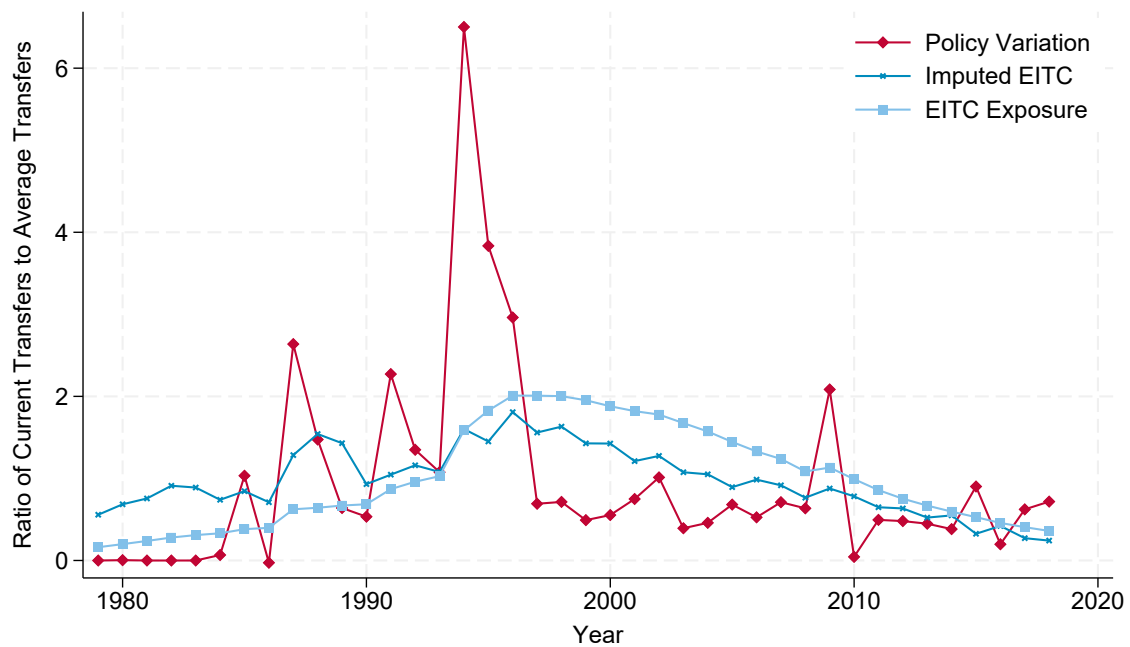


Figure 2.3: Standardized transfers and policy variation.

Notes: The figure plots the average imputed EITC and EITC exposure values (blue) and the average policy variation instrument (red). Imputed transfers and policy variation are scaled using the ratio between the value in each year and a baseline (the mean of the series).

The figure shows distinctions between endogenous tax credits and policy variation. First, the instrument captures institutional changes in the tax credit schedules. The most prominent spikes coincide with federal reforms in the 1980s, the Welfare Reform of the 1990s, and subsequent reform periods in 2002, 2008, and 2009. Policy variation in the remaining years reflects the rollout, expansion, and contraction of state programs. Furthermore, EITC exposure and imputed tax credits have trends, potentially driven by the age-profile of earnings and labor force attachment, policy variation does not show such trends.²⁹

Finally, the instrument is used to predict the three endogenous variables in Equation 2.1. The multiple endogenous variable approach requires estimating three first-stage equations where the right-hand side instrument are the aggregated policy changes across each stage of childhood $Z_{i,(a)}$.³⁰ Each first-stage equation is estimated as follows:

$$\text{EITC}_{i,(a)} = \phi_1 Z_{i,(0-5)} + \phi_2 Z_{i,(6-12)} + \phi_3 Z_{i,(13-18)} + \eta_1 X_i + \psi_1 V_{s,t} + \lambda_{1,s} + \lambda_{1,t} + \varepsilon_i \quad (2.2)$$

²⁹The exogeneity of the instrument may be hampered if there is correlation between policy variation captured by Z_{it} and family characteristics included when estimating $\hat{E}_{it}$ and ς_{it} . I test these correlations within a simple regression framework. Further robustness tests show that results are consistent across alternative specifications in the construction of the instrument using different functional forms in predicting $\hat{E}_{i,t}$, and calculating χ with lagged predicted income $\hat{E}_{i,t-1}$ and independent family characteristics. In the latter case, all families are assumed to be composed of single parents with three children living in the mother's state of residency when the child was born.

³⁰The strength of the instruments is tested with two methods. First, standard inference testing in models with multiple endogenous variables calculate the Sanderson-Windmeijer (SW) multivariate F test (Sanderson and Windmeijer, 2016). I present the F-test and its corresponding p-value in each regression. This F-statistic is a corrected version of the Cragg-Donald rank test, developed for models with two or more endogenous variables. Second, I run 100 Monte Carlo simulations where policy variation is calculated using placebo income levels and family characteristics (number of dependents and marital status) are randomly assigned to families. All simulations use a log-normal distribution of income, matching the moments of the actual distribution in the analytical sample. The number of permutations is restricted by computing power, as they cannot be performed in the BLS' restricted-access server and instead are constructed in a sample with 10 million observations representing every observation in the sample as if they lived in each state. Simulated policy variation is assigned in the state of residency of mother's in 1979. In the test, I regress the irrelevant instrument on imputed tax credits and test whether estimates approach the actual estimates. The results are presented in Appendix A.9 and suggests that placebo policy variation does not replicate the main estimates.

This empirical strategy is well-suited for cross-sectional studies of long-term outcomes. There are a multitude of unobserved heterogeneities influencing the life cycle and associated with program participation, socioeconomic status, and family characteristics. For example, the EITC is believed to influence skill development in childhood (Chetty et al., 2011; Averett and Wang, 2018), which may stimulate dynamic complementarities related to brain plasticity (Noble, 2021; Noble et al., 2015; Cunha and Heckman, 2007). There is also discussion on whether the program has influenced unobserved factors at home, such as time spent with children as well as time allocation (Bastian and Lochner, 2020), domestic violence (Edmonds et al., 2022; Cesur et al., 2022), and children’s health, skills and home quality environment (Averett and Wang, 2018). These mechanisms, while not entirely understood, contribute to a complex system of endogenous intertwined relationships. Under the assumption that policy variation is exogenous to idiosyncratic heterogeneities, the instrumental variable regression provides LATE estimates driven by policy variation that influences individuals solely through the EITC schedule.³¹

2.4 Results

Figure 2.4 comprises the main results in this paper.³² The plot presents the marginal effects of \$1,000 in tax credits across different childhood stages on adult income. The figure includes results from three different regressions. First, regressing adult income

³¹Ultimately, this approach assumes that policy variation affects households only through the current year tax refunds that include their tax credit transfers. That is, I assume that total family income E_{it} is composed of permanent income $\bar{E}_i$, income shocks μ_{it} , and the program such that $E_{it} = \bar{E}_i + \text{EITC}_{it}(E_{it-1}, Z_{it}) + \mu_{it}$. The simulated approach uses Z_{it} as a predictor of EITC and requires two assumptions. First, federal and state reforms in the tax credit schedule are independent of households responses; and, second, that households do not adjust to policy variation Z_{it} . The economic intuition is that a household responses aim at maximizing tax refunds but not as a dynamic adjustment to year-to-year changes in the tax schedule. It seems reasonable to think that households do not revise the historic changes in Publication 596 when filing their taxes. Appendix A.4 presents evidence that the instrument is uncorrelated with family characteristics, including family income, which may ease concerns that households dynamically respond to policy variation.

³²The robustness of results to covariates is presented in Appendix A.5. The results are also robust to the instrumental variable approach. See Appendix Figure A.5.

between ages 22 and 27 on the EITC across all three childhood stages (blue diamond line) following the adult age range in [Bastian and Micheltmore \(2018\)](#). Second, average adult income between ages 22 and 30 regressed on tax credits (blue triangle). Lastly, a regression on adult income up to the year the child is 40 years old. The evidence suggests that tax credit transfers during early childhood increase adult income by \$487.70 when the child turns forty years, or 1.7 percent of average income in the sample. Transfers during other stages of childhood have negligible effects.

My findings highlight strong early childhood effects that grow with age. Similarly, [Barr et al. \(2022\)](#) finds that tax benefits and deductions during the year immediately after birth increase adult income during the mid twenties by approximately 1 percent and 3 percent for adult income in the mid thirties. Figure 2.4 highlights this increasing pattern over adulthood. Transfers in early childhood are insignificant for adult income up to age 27, but larger and significant coefficients when accounting for individuals up to thirty and forty years old.

The effects of tax credits dampen and are statistically insignificant in mid and late childhood in all three regressions. This contrast with evidence in [Bastian and Micheltmore \(2018\)](#) where exposure to the program during late childhood increases earnings by 2 percent when the child is between 22 and 27 years old, arguing that cash-at-hand for teenagers had strong effects boosting skill acquisition or college enrollment ([Bastian and Micheltmore, 2018](#); [Manoli and Turner, 2018](#)). Estimators based on the instrumental variable suggests that tax credit transfers after early childhood may have a negative association with adult income, although this insignificant coefficient is negligible and close to zero.

The empirical findings are consistent in the instrumental variable approach. Table 2.3 presents the estimate of \$1,000 in EITC during childhood on the mean adult income up to age 40.³³ The first columns show OLS results equivalent to Figure 2.4, while the next six columns exploit policy variation with alternative constructions as

³³The results cannot be replicated using EITC exposure and the instrument is a weak predictor of exposure. In fact, the CNLSY sample does not show [Bastian and Micheltmore \(2018\)](#)'s any evidence that exposure during late childhood affects adult earnings. See Appendix A.3.

robustness test. My preferred estimates, in column 2, indicate that early childhood effects increase annual income by \$587.96 (2.1 percent) for every \$1,000 in cash transfers. Transfers during mid and late childhood are insignificant. These estimates are calculated using the simplest construction of policy variation, with a linear prediction of family income during childhood using lagged income, year fixed effects, and mother fixed effects.³⁴ Estimates of early childhood effects are robust to excluding family characteristics from the inputs of policy variation (column 3) and calculating policy variation with income from the previous year (column 4), which may alleviate the role of endogenous characteristics.

Columns 5 and 6 in Table 2.3 present the smallest early childhood effects. Column 5 uses an autorregressive function to calculate predicted income as an input of policy variation. In this case, an increase of \$1,000 in tax credits during early childhood is associated with nearly \$400 larger income during adulthood. Column 6 uses a spline of lagged income to predict earnings and shows a correlation of nearly \$500 per \$1,000 in cash transfers. In both cases, coefficients are weakly significant at 10 percent levels.

The bottom rows in Table 2.3 provide conventional statistics, including the number of observations, adjusted R-squared values, and the F-test. All results in columns 2 to 6 have a strong instrument according to the tests for weak identification with multiple endogenous variables (Sanderson and Windmeijer, 2016). In columns 2 to 6, policy variation uses predicted family earnings. However, column 7 uses the mother's net family income in 1979, which is a weak instrument and produces coefficients with large and insignificant coefficients. In that specification, net family income captures maternal background in her upbringing, which appears disconnected from her offspring's.³⁵

³⁴The instrument is constructed by inputting predicted family income, number of dependents, marital status, and state of residency during the current year and calculating the difference with eligible transfers during the previous period holding all inputs constant.

³⁵The first stage coefficient is stable to the choice of covariates. Appendix A.4 evidences that the first stage is independent of observables. Appendix A.9 further tests the strength of the instrument, highlighting that policy variation based on random family characteristics does not replicate these main results. The results assert that policy variation is a strong instrument to cumulative tax credits.

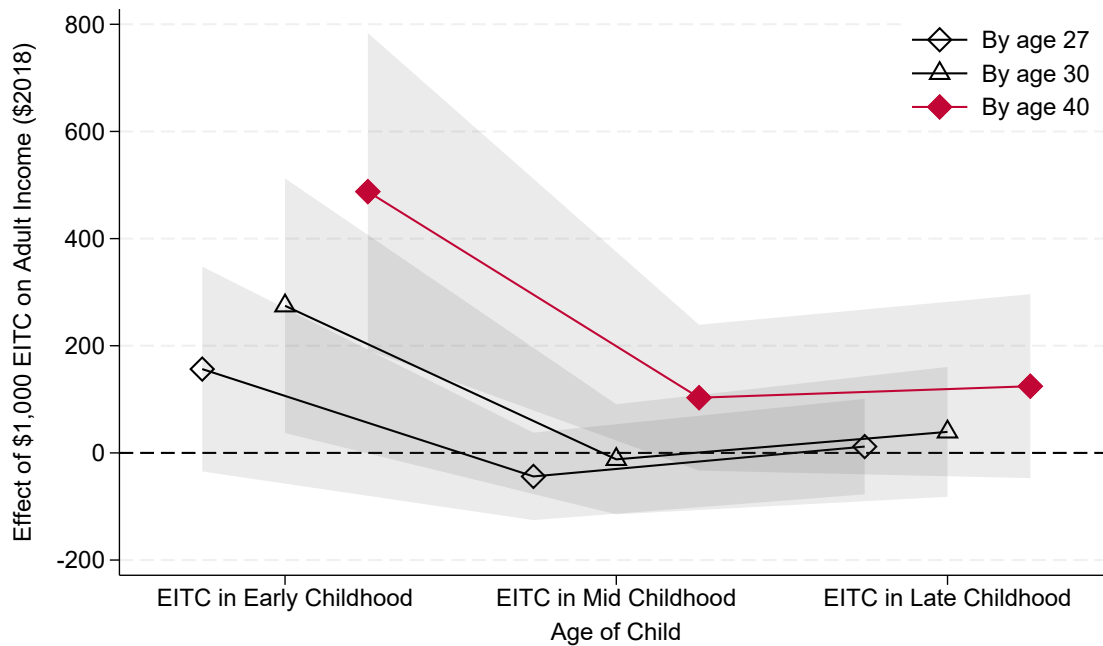


Figure 2.4: OLS Coefficients of imputed earned income tax credit on the adult income by childhood stage.

Notes: The figure plots the coefficients of \$1,000 in imputed tax credit transfers on adult income. The corresponding childhood stage of each coefficient is shown in the horizontal axis. Black diamonds are coefficients on the adult income between age 22 and 27, the triangle represents adult income between age 22 and 30, and the red diamond is adult income up to age 40. 95% confidence interval clustered at the state level in grey. The regressions control for mother fixed effects, cohort fixed effects, quadratic and cubic age in 2018, indicators for black, Hispanic, female, whether the mother has been married by age 18, an indicator for whether the mother graduated high school, an indicator for whether the mother has some college education, and fixed effects for the number of siblings at age 18, the three-way interactions between race, ethnicity, and sex by number of siblings, birth year and state. The specifications account for state policy variables measured at age 18 (GDP per capita, unemployment rate, the top marginal tax rate, the minimum wage, tax revenues, higher education spending, and the maximum welfare benefits), state fixed effects, and year of birth fixed effects. The specifications also account for state policy variables measured at age 18.

Table 2.3: OLS and 2SLS coefficients of imputed earned income tax credit on adult income between age 22 and 40.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
EITC in Early Childhood	487.70*** (147.01)	586.96** (227.16)	585.58** (243.50)	652.88** (315.36)	394.86* (227.77)	495.36* (289.23)	5111.77 (9203.58)
EITC in Mid Childhood	103.09 (67.60)	-19.46 (130.15)	-64.92 (147.07)	-117.92 (171.03)	11.24 (167.41)	-46.42 (166.41)	1421.19 (9407.73)
EITC in Late Childhood	124.53 (85.22)	41.36 (169.70)	212.83 (217.95)	431.60 (296.22)	158.05 (197.54)	124.18 (264.90)	3719.42 (20692.98)
Obs.	5253	5253	5253	5253	5253	5253	5253
R2	0.59	0.07	0.07	0.06	0.07	0.07	-0.92
Estimator	OLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
KP Wald F		37.37	68.50	34.04	78.86	42.42	0.02
SW p-value (1)		0.00	0.00	0.00	0.00	0.00	0.75
SW p-value (2)		0.00	0.00	0.00	0.00	0.00	0.74
SW p-value (3)		0.00	0.00	0.00	0.00	0.00	0.75
Pred. Income Period		Current	Current	Lagged	Current	Current	1979
Pred. Income Function		Linear	Linear	Linear	Autoregressive	Spline	Fixed
IV Demographics		Yes	No	No	No	No	No

Notes: The Table presents coefficients of \$1,000 in EITC transfers on mean adult income between ages 22 and 40. 2SLS coefficients are instrumented with cumulative policy variation in each childhood stage. The rows show the number of observations, adjusted R-squared, the Kleibergen-Paap F statistic (KP Wald F), and the Sanderson-Windmeijer (SW) multivariate F test corresponding to each first stage equation (Sanderson and Windmeijer, 2016). Column 1 is estimated with OLS, and Columns 2-7 are estimated with 2SLS. Policy variation is calculated holding predicted family income constant for the current period of tax credits, the previous year (lagged), or with net family income of mothers in 1979. The linear predicted income function is regressed with lagged wages, year fixed effects, and mother fixed effects. The autoregressive function uses one income lag. The spline uses a fifth-order spline of lagged income. Fixed predicted income is the grandfather's income in 1979 adjusted for inflation. The instrument is computed with demographics (number of children, marital status, and state of residency), or assuming that all mothers are single with three children residing in the state where the child was born. Regressions control for mother fixed effects, cohort fixed effects, quadratic and cubic age in 2018, indicators for black, Hispanic, female, whether the mother has been married by age 18, an indicator for whether the mother graduated high school, an indicator for whether the mother has some college education, and fixed effects for the number of siblings at age 18, the three-way interactions between race, ethnicity, and sex by number of siblings, birth year and state. The specifications also accounts for state policy variables measured at age 18 (GDP per capita, unemployment rate, the top marginal tax rate, the minimum wage, tax revenues, higher education spending, and the maximum welfare benefits), state fixed effects, and year of birth fixed effects. Robust standard errors clustered at the state-level in parentheses. *** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$.

2.4.1 Dynamic complementarities and aggregated EITC

My results highlight that the timing of investments in children influences the effectiveness of social programs. In this section, I test the evidence on the timing of investments by re-estimating Equation 2.1 using different aggregation criteria across age intervals. The analysis begins by including one imputed EITC variable for each childhood stage in an OLS setting. Then, I group childhood stages by pairs and triples. The results highlight the sensitivity of coefficients to the aggregation decisions in Bastian and Micheltore (2018). In an attempt to shed light on these divergences, I propose a difference-in-difference estimator based on variation in years of exposure to the 1993 reform.

Figure 2.5 shows the analysis of age intervals. The left panel includes individual coefficients for EITC transfers estimated at each age. The results are volatile with large standard errors, and some coefficients have negative values. This behavior is attenuated by expanding the interval size. The middle panel in Figure 2.5 presents stages of childhood aggregated in pairs of years. The results remain statistically insignificant at conventional levels, but the coefficients smooth the observed negative values around age 3. Therefore, all early childhood coefficients are positive while mid childhood and late transfers continue showing negative correlations. In the third panel, where age is grouped in triplets, the smooth dampening shape is clearer and both of the early childhood intervals show positive and significant coefficients.

Further analyses hint that the aggregate tax credits during childhood may have positive effects. Table 2.4 summarizes these results. First, columns 1 to 3 include estimates of early, mid, and late childhood with only one endogenous variable in the specification, instead of multiple variables. Omitting transfers during different childhood stages does not change the general result: early childhood transfers have a robust positive relationship with adult income. Column 4 shows the association between cumulative tax credit transfers between age 0 and 18 on adult income. The estimate is smaller than those of early childhood effects but remains positive,

suggesting that the null tax credits in mid and late childhood decrease the average effects of the policy. Estimating early childhood transfers with an instrumental variable continues to yield significant estimates (column 7) but all other childhood intervals are insignificant and have a negative association in columns 6 and 7. These negative coefficients during mid and late childhood may offset the impact of cumulative transfers when considering all of childhood in column 8.

In summary, the evidence is inconclusive, but transfers may have dynamic complementarities. The term was advanced in the early childhood literature to describe structural models where investments during later stages of childhood are augmented by early period investments (Cunha and Heckman, 2007). An intuitive example is that a teenager receiving free calculus textbook may benefit from the investment far less than a peer who also received algebra textbooks during elementary grades, *ceteris paribus*. In that sense, investments during early childhood complement later-life investments. These complementarities in treatment effects may be related to income growth in later adulthood, perhaps due to investments in skills and improved health that materialize in the thirties.

I take an additional step to explore dynamic complementarities. The strategy hinges on differences in program eligibility and differences in birth cohort exposure to the 1993 EITC reform, an approach inspired in Goodman-Bacon (2021).³⁶ There are at least two disadvantages in this setting. First, defining treatment as program eligibility confounds additional safety net expansions during 1993 that share eligibility criteria with the tax credit. Second, eligibility criteria changes across the years, violating the assumption of stable treatment and control groups. In addition, the

³⁶The method relies on an event study that follows equation $Y_i = \alpha + \sum_{\tau=-3}^{-2} \delta_{\tau} 1(\text{EITC}_i > 0) \tau + \sum_{\tau=0}^T \delta_{\tau} 1(\text{EITC}_i > 0) \tau + \pi X_i + \omega_t + \omega_s + \mu_i$. In this setting, exposure is the difference between the year the individual turns 18 and the policy onset in 1993. A person born in 1976 turns 18 in 1994 and is exposed to OBRA 1993 for only one year, becoming ineligible after age 19; while an individual born in 1990 is exposed to the program for 15 years. OBRA 1993 is selected instead of other reforms because it presents the largest policy variation in Figure 2.3. Data limitations are an additional concern, as two-thirds of the sample are born in the 1980s. Note that we do not observe fewer years of adult income for individuals fully exposed to the program during childhood (someone born in 1993 is 25 years old by the end of the sample).

program has undergone multiple reforms in each decade. Altogether, these features suggest that the framework may be suggestive of the underlying correlations between safety net expansion and intergenerational mobility, but causal inference is flawed.

The event study in Figure 2.6 shows that individuals exposed more years to the 1993 expansion have up to 15 thousand dollars more income. The vertical axis graphs the magnitude of coefficients on the real adult income at age 40, while the horizontal axis captures the years of exposure to the tax credit expansion. There is an increasing trend in adult income such that individuals exposed more years to the policy have larger returns to the program than individuals exposed fewer years to the 1993 reform. The significant coefficients begin for individuals exposed through their first four and five years of childhood, consistent with the patterns previously documented. This is suggestive evidence of dynamic complementarities as one additional year of exposure to the program appears to increase the program's effectiveness. The event study further indicates that pre-trends are balanced in the sample.

2.4.2 Heterogeneities

I explore heterogeneities in four dimensions: sex, race and ethnicity, mothers' education, and mothers' marital status. I estimate the association between tax credit transfers and adult income in a restricted sample defined by the dimension of interest. Figure 2.7 shows the results of the analysis, with the magnitude of the coefficients plotted in the horizontal axis and the subgroup of interest indicated in the horizontal axis. The coefficients for early childhood transfers are at the top of the figure, and mid childhood tax credits at the middle followed by late childhood transfers at the bottom. Early childhood transfers show significant heterogeneities for white boys and children of married mothers standing out as the subgroups that benefit the most from the intergenerational effects of the program.³⁷

³⁷One major issue in the analysis of heterogeneities by mother's education and marital status is that these are potentially endogenous. However, heterogeneities persist when using the instrument as seen in Appendix Figure A.6.

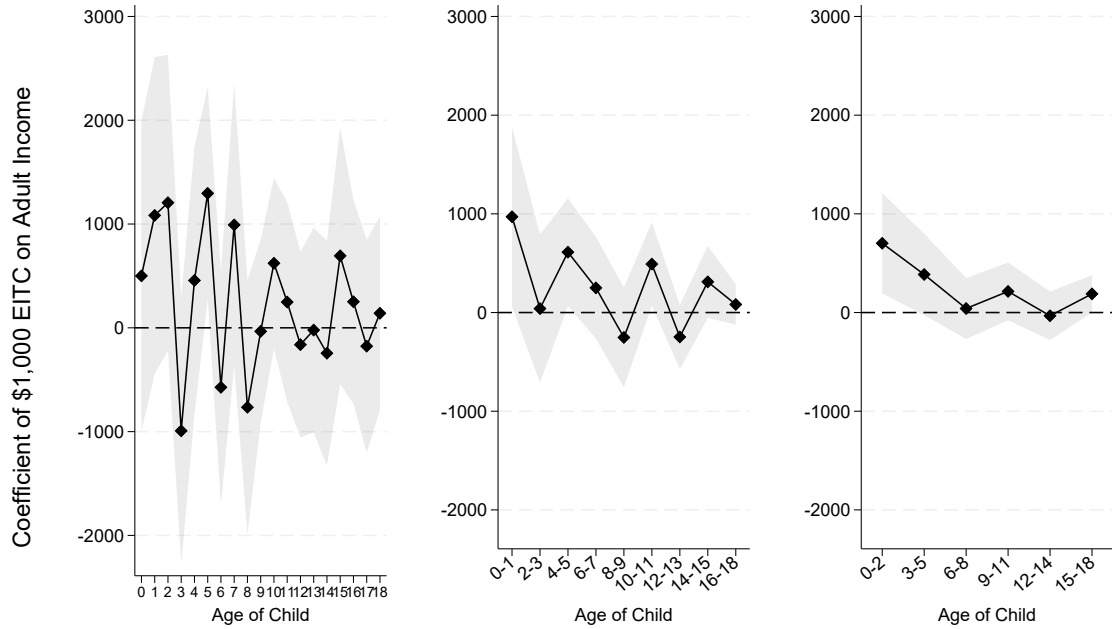


Figure 2.5: Effects of tax credits on adult income across different intervals of childhood stages.

Notes: The Figure plots OLS coefficients of \$1,000 in EITC transfers on mean adult income between ages 22 and 40. All regressions control for mother fixed effects, cohort fixed effects, quadratic and cubic age in 2018, indicators for black, Hispanic, female, whether the mother has been married by age 18, an indicator for whether the mother graduated high school, an indicator for whether the mother has some college education, and fixed effects for the number of siblings at age 18, the three-way interactions between race, ethnicity, and sex by number of siblings, birth year and state. The specifications also accounts for state policy variables measured at age 18 (GDP per capita, unemployment rate, the top marginal tax rate, the minimum wage, tax revenues, higher education spending, and the maximum welfare benefits), state fixed effects, and year of birth fixed effects. The 95% confidence intervals clustered at the state-level are plotted in grey.

Table 2.4: OLS and 2SLS effects of tax credits on adult income using single endogenous variables and total EITC in childhood.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
EITC in Early Childhood	429.99*** (135.96)				564.66*** (189.83)			
EITC in Mid Childhood		68.99 (58.28)				-166.54 (172.27)		
EITC in Late Childhood			27.91 (80.24)				-123.45 (160.28)	
EITC between ages 0-18				154.69** (63.66)				59.50 (94.11)
Obs.	5253	5253	5253	5253	5253	5253	5253	5253
R2	0.59	0.58	0.58	0.59	0.07	0.06	0.06	0.07
Estimator	OLS	OLS	OLS	OLS	2SLS	2SLS	2SLS	2SLS
KP Wald F					359.41	370.70	213.58	200.39

Notes: The Table presents OLS and 2SLS coefficients of \$1,000 in total EITC transfers on mean adult income. The dependent variable is the mean income between ages 22 and 40. The rows show the number of observations, adjusted R-squared, and the Kleibergen-Paap rk Wald F statistic (KP Wald F). Columns 1-4 are estimated with OLS, and Columns 5-8 are estimated with 2SLS instrumented with total policy variation. Policy variation is calculated with a linear prediction of income and family characteristics demographics in the current year. All regressions control for mother fixed effects, cohort fixed effects, quadratic and cubic age in 2018, indicators for black, Hispanic, female, whether the mother has been married by age 18, an indicator for whether the mother graduated high school, an indicator for whether the mother has some college education, and fixed effects for the number of siblings at age 18, the three-way interactions between race, ethnicity, and sex by number of siblings, birth year and state. The specifications also accounts for state policy variables measured at age 18 (GDP per capita, unemployment rate, the top marginal tax rate, the minimum wage, tax revenues, higher education spending, and the maximum welfare benefits), state fixed effects, and year of birth fixed effects. Robust standard errors clustered at the state-level in parentheses. *** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$.

The results suggest that the adult income of men is \$857.26 (2.7 percent) larger for every \$1,000 in tax credit transfers received during early childhood but girls have smaller and statistically insignificant benefits. These heterogeneities are consistent with studies on the long-term impact of tax benefits aimed at children where boys show significant gains but girls do not (Barr et al., 2022). The general consensus is that boys are more sensitive to household resources than females during childhood (Autor et al., 2019; Chetty et al., 2016; Bertrand and Pan, 2013). For example, Chetty et al. (2016) suggests that traditional gender disparities reverse at the bottom of the income distribution, such that boys tend to do poorly in their academic and labor market outcomes in comparison with girls and these disparities persists into adulthood.

The results reveal an intriguing pattern in terms of race: white children appear to benefit more than their non-white counterparts. The difference is weak as the impact on black children is significant at 10 percent levels, but the finding diverges from the extant literature. Empirical research documents that tax credit expansions are associated with reductions in the racial income gap and increased income for black households, relative to their white counterparts (Hardy et al., 2022). The literature also argues that married black mothers exhibit larger responses to the program than white married mothers (Jiao et al., 2022). One possibility for the heterogeneity is that income effects are offset by the intensive margin responses in the labor supply of non-white mothers. Furthermore, there are no heterogeneities manifested for Hispanics and non-Hispanics. Taken together, the evidence calls for further research on racial disparities in the intergenerational income elasticity among recipients of the EITC.

Finally, the last two rows in Figure 2.7 present heterogeneities by the marital status of mothers. I find that the intergenerational effect of the policy is only significant for children of married mothers. Research suggests that married mothers respond to the EITC by reducing their labor supply and financially relying on their spouse (Eissa and Hoynes, 2004), which may increase investment in their kids during early childhood. In a similar vein, the impact of tax credits appears stronger in

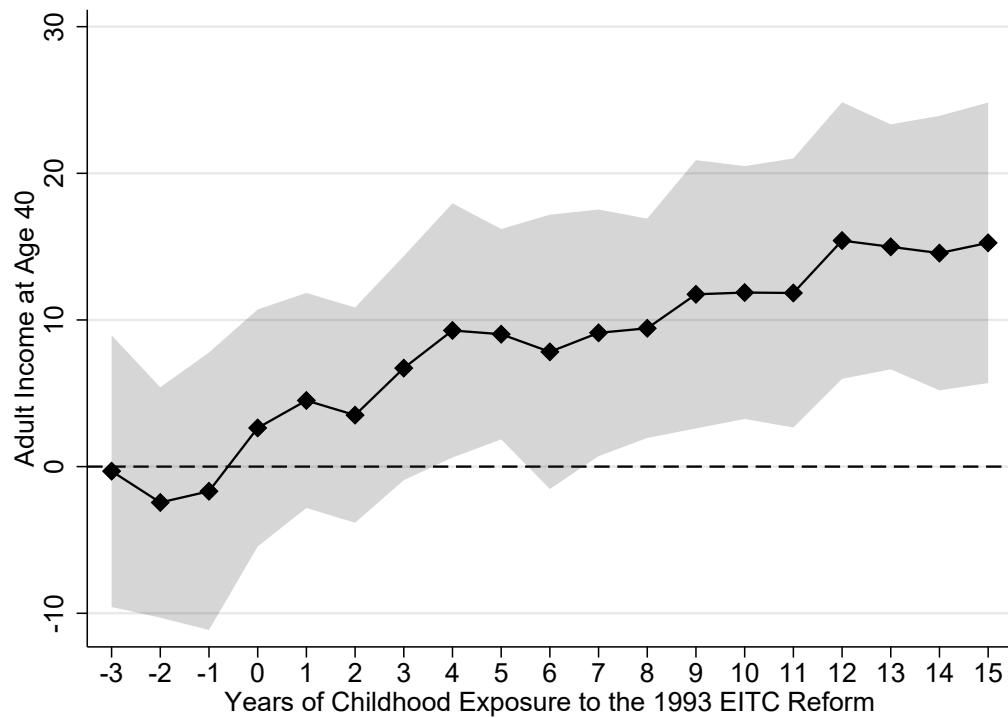


Figure 2.6: Event study of years of childhood exposure to the 1993 EITC reform on adult income.

Notes: The Figure presents difference-in-difference estimates of the impact of the 1993 welfare reform and the adult income of exposed children. The dependent variable is the mean income between ages 22 and 40. The event study exploits differences in birth cohort and differences in EITC eligibility. All controls are included. The 95 percent confidence intervals are plotted in grey.

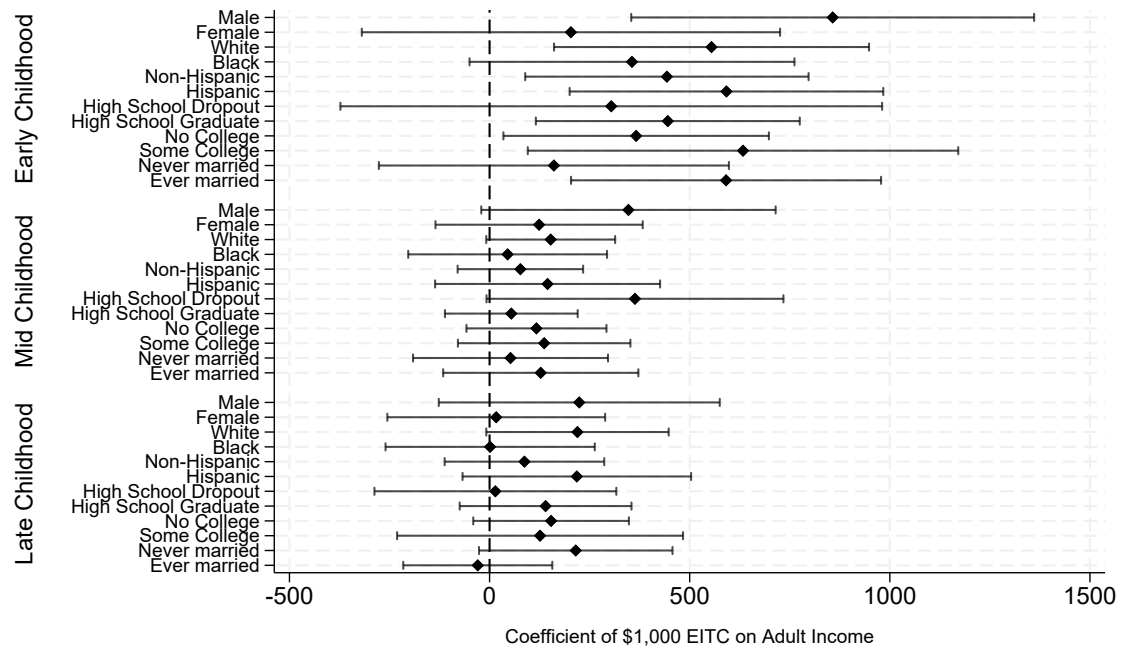


Figure 2.7: Subgroup analyses of tax credits on adult income between ages 22 and 40.

Notes: The figure plots the coefficients of \$1,000 in imputed tax credit transfers on adult income across subgroups. The vertical axis is the size of coefficients. 95 percent confidence intervals are clustered at the state level. The horizontal axis shows the subgroup of relevance. Coefficients for early childhood tax credits are at the top, mid-childhood tax credits in the middle, and late childhood at the bottom. The regressions control for mother fixed effects, cohort fixed effects, quadratic and cubic age in 2018, indicators for black, Hispanic, female, whether the mother has been married by age 18, an indicator for whether the mother graduated high school, an indicator for whether the mother has some college education, and fixed effects for the number of siblings at age 18, the three-way interactions between race, ethnicity, and sex by number of siblings, birth year and state. The specifications account for state policy variables measured at age 18 (GDP per capita, unemployment rate, the top marginal tax rate, the minimum wage, tax revenues, higher education spending, and the maximum welfare benefits), state fixed effects, and year of birth fixed effects.

the maternal labor supply of unmarried mothers with toddlers (Micheltmore and Pilkauskas, 2021), and unmarried mothers with children under three spend more in childcare services after program expansions (Pilkauskas and Micheltmore, 2023; Micheltmore and Pilkauskas, 2021). Kuka and Shenhav (2020) also evidences that maternal labor supply is most responsive to tax credits recently after birth. In this sense, these heterogeneities are explained if increased childcare spending offsets the program benefits for single mothers, for whom the policy is most effective in increasing labor supply. Married mothers, on the other hand, boost the program intergenerational effectiveness when staying at home.³⁸ Therefore, the relationship between maternal labor supply and the intergenerational impact of the EITC is an important mechanism to examine.³⁹

2.4.3 Mechanism analysis: maternal labor supply

Maternal labor supply is the proposed mechanism that may offset intergenerational effects. Increased labor force attachment in mothers, driven by the EITC policy, may lead to additional expenses in childcare, thereby mitigating the positive income effect during early childhood (Pilkauskas and Micheltmore, 2023; Micheltmore and Pilkauskas, 2021; Rodgers, 2018). Mothers staying at home contribute to additional dimensions of parenting, such as behavior monitoring, rules enforcement, and early stimuli.⁴⁰ Lastly, studies argue that maternal labor supply reduces skill development among low-income children (Caetano et al., 2021; Blau and Currie, 2006).⁴¹ To test the

³⁸This would also explain racial heterogeneities, given the high responsiveness of black mothers' labor supply to the program (Jiao et al., 2022).

³⁹Figure 2.7 also presents heterogeneities by educational achievement of mothers. There are no effects among the children of mothers without a high school degree. However, sample size restrictions may cause these disparities as only 639 observations are children born to high school dropout mothers. There are no statistical difference between children with mothers without a college degree and those with a college degree, although the coefficient is twice as large for the latter group.

⁴⁰Child development is most sensitive in early childhood. For example, vocabulary exposure in quantity and quality during early ages accounts for a 30 million word gap heard among children with low and high socioeconomic status (Romeo et al., 2018; Suskind et al., 2015).

⁴¹In recent years, the literature has developed a growing concern about maternal labor supply and childcare, particularly due to social programs imposing work conditions that distort time allocation and change the supply of household goods, including childcare. The question, then, is about the

mechanism, I examine whether the policy has a weaker impact on children whose mothers spend more hours working.

This investigation begins by examining the intensive margin of maternal labor supply. The sample is divided into EITC eligible mothers and a control group made of those who are ineligible. Figure 2.8 plots the change in women’s weekly work hours after having a child when they are EITC eligible and when they are not, relative to two years before birth. This is a striking picture: while both groups are not statistically different before birth, eligible mothers receive less of a pregnancy penalty upon birth or, alternatively, ineligible mothers disengage more from the labor market after pregnancy. That is, the EITC increases the likelihood that mothers remain attached to the job market. Previous literature supports Figure 2.8 as EITC expansions have larger effects on maternal labor supply immediately after first birth (Kuka and Shenhav, 2020).⁴²

To link maternal work intensity to the intergenerational effects of the EITC, I condition Equation 2.1 on whether maternal work hours are above or below the median. However, maternal labor supply is potentially endogenous. I estimate maternal labor supply intensity using predicted hours of work based on endowed characteristics: race, ethnicity, education, state of residency, and year fixed effects. The intergenerational effects of the program are then estimated in the sample of children whose mothers had low labor intensity (worked below the median on average

counterfactual childcare arrangement in the absence of maternal labor supply. Bastian and Lochner (2020), for instance, studies the type of time spent with children at home by mothers and finds that social programs do not affect the allocation of “childcare investments” but, instead, reduce leisure. Recent studies find an increase in formal childcare spending induced by the program (Pilkauskas and Michelmore, 2023; Michelmore and Pilkauskas, 2021).

⁴²Appendix A.7 shows the unconditional work hours of mothers by the age of their children. Mothers who are EITC eligible work fewer hours before having a child, compared with mothers who are not eligible. After the child is born, work hours converge for both groups. Appendix Figure A.11 expands the analysis to maternal income, which recovers faster for EITC eligible mothers. Both groups had similar income trends prior to having a child, but mothers eligible for the program reattach to labor markets and receive larger incomes. Appendix Figure A.12 shows that, unlike maternal income, family income is larger for the control group, which is to be expected as only low income families are eligible for the program.

during the individual’s childhood) and in the sample of children whose mothers had high labor intensity (worked above the median).⁴³

Suggestive evidence in Figure 2.9 indicates that the positive long-run effects of early childhood tax credits are manifested only for children whose mothers dedicated fewer hours to work during their childhood. The implications are that intergenerational policy effects on the adult income of children are driven mostly by the group of mothers that benefits from the income effects of the tax credits while working fewer hours. This result is consistent with maternal labor supply intensity prior to childbearing (Appendix Figure A.13) and the instrumental variable analysis (Appendix Figure A.14). The evidence hints that mothers with fewer working hours receive the positive income shock without allocating these resources to daycare centers, thereby mitigating any substitution effects that offset the intergenerational effectiveness of the program. This result aligns with subgroup analyses and a body of literature exploring program interactions with childcare.⁴⁴

2.5 Discussion

In this paper, I revisit the estimates of the long-term effects of the EITC on the adult income of individuals who received the policy during their childhood. The paper studies the optimal timing for investing in children and analyzes the mechanisms of

⁴³I test two additional strategies with similar results. First, estimating the number of hours worked prior to childbearing. The sample is then divided into children whose mothers had low labor intensity (below the median) and high labor intensity (above the median). The results are presented in Appendix Figure A.13. In a second robustness test, the sample is divided by endogenous maternal labor supply intensity during early childhood. Therefore, I take the average labor supply intensity of women when the child is between 0 and 5 years old and categorize women according to the quintile in the distribution of work hours: first quintile (low labor intensity), second (mid-low intensity), third (mid-high intensity), and the top quintile (high intensity). The intergenerational effects of the program are then estimated in each of the four groups. In this case, endogenous maternal responses justify the instrumental variable, as using policy variation allows for keeping maternal labor supply constant. This analysis is shown in Appendix Figure A.14.

⁴⁴For example, married mothers show positive intergenerational effects, which is consistent with previous research on how married mothers leverage the income of their spouses to reduce their labor supply (Eissa and Hoynes, 2004). The result may also explain racial disparities, as the labor supply of black mothers is more sensitive to the EITC (Jiao et al., 2022). The analysis is consistent with reductions in labor market penalties for motherhood (Kuka and Shenhav, 2020).

the program. This project contributes to the existing literature by introducing a novel research design, a plausibly exogenous instrument capturing localized policy variation across the tax credit schedule, and variation in stage and annual generosity. I also draw from rich multi-generational labor data to link these intergenerational effects with maternal labor supply intensity. The paper sheds light on how work conditions interact with the intergenerational effects of the program and mitigate the effectiveness of the policy in reducing long-term poverty.

The findings point that an additional \$1,000 in EITC transfers during early childhood boosts adult income by \$487.70 each year (between 1.7 and 2.7 percent). The evidence is stable across specifications using OLS and an instrumental variable approach. I document that mothers eligible to receive the tax credit tend to work more after birth, but it is among mothers that work fewer hours that we observe the positive intergenerational effects of the program. The analysis of heterogeneities is consistent with maternal labor supply as the key mechanism offsetting the program's impact. There are significant heterogeneities for boys, white families, and children of married women. The latter two groups have mothers whose labor supply is less sensitive to the program. Contrary to previous research emphasizing a 2 percent increase in income arising from EITC exposure during late childhood [Bastian and Micheltore \(2018\)](#), I find little evidence of effects during mid and late childhood and robust positive effects for transfers received during early childhood. My findings are consistent with [Barr et al. \(2022\)](#), who estimates that children born around the cutoff to receive the EITC during their first year have between 1 and 3 percent larger income as adults.

To contextualize these numbers, meta-analyses regarding TANF/AFDC suggest that an additional \$1,000 cause a 0.05-0.06 standard deviation improvement in student achievement that roughly translate into 1 percent gains in earnings ([Hoynes and Schanzenbach, 2018](#)). [Brown et al. \(2015\)](#) find that a one-standard deviation increase in Medicaid eligibility increases female earnings by 1.9 percent by age 28. Meanwhile, the SNAP program roll-out has statistically insignificant effects on adult

income [Hoynes et al. \(2016\)](#). In comparison, investing \$1,000 in children through the EITC generates a long-run income equivalent of four months in food supplements every year.

There are several limitations in this study. Most importantly, we do not observe actual tax records in the data and instead impute tax credit values using a calculator. Therefore, these estimates are a similar measure of exposure as in the literature, but they leverage variation in family income to localize effects for low income families. Another key limitation is that the CNLSY is not a representative sample of the overall US population. On one hand, low income households are over-represented in the sample, which may be advantageous to study means-tested programs. On the other hand, there may be systematic differences across cohorts, for the CNLSY tracks children of a single cohort of mothers. Later cohorts are likely to be born from older mothers or be the younger sibling in the family. Also, my sample size remains limited despite being larger than the analytical sample used by previous research. As a consequence, these estimates may be noisy. This is particularly important when the sample is restricted, such as in the mechanism analysis and studying heterogeneities.

From an economic standpoint, time allocation and work conditions yield income and substitution effects in the family that drive variation in the labor market outcomes of children. From a policy perspective, the nuanced effects of the program offer guidance on complementary policies that may boost the effectiveness of a \$60 billion program that already shows promise in poverty alleviation. Future research could benefit from further use of administrative data, similar to [Barr et al. \(2022\)](#), and randomized research designs, in analyzing the optimal timing of EITC interventions and exploring whether work conditions mitigate the program effectiveness. Additionally, variations in state rollout and birth cohorts, would offer valuable insights and contribute a deeper understanding of the program.

This paper has three takeaways. First, the timing and duration of policy exposure is an important factor in the intergenerational effects of social programs. The analysis highlights the heterogeneous impacts of tax credits received during early childhood

compared to those received later in childhood. This underlines the significance of early investments and its impact in shaping future outcomes. This finding aligns with the theoretical and empirical consensus on the crucial role of early childhood investments ([Barr et al., 2022](#); [Carneiro et al., 2021](#); [Cunha and Heckman, 2007](#)). Second, the link between imposing work conditions and changes in maternal time allocation may impact the policy effects on future generations. The intergenerational effects of the program on children are higher for those mothers who worked fewer hours during childhood, suggesting that reductions in the supply or quality of parental care when mothers work more hours mitigates the positive effects of the program. Supporting working mothers may have benefits in reducing gender inequality and improving the labor outcomes of women ([Kuka and Shenhav, 2020](#)). However, we must consider that distortions in time allocation imply changes in childcare arrangements. Emerging literature continues to study this topic and further analyses of childcare and parental leave are necessary ([Pilkauskas and Micheltore, 2023](#); [Micheltore and Pilkauskas, 2021](#); [Bastian and Lochner, 2020](#); [Rodgers, 2018](#)). Finally, the contrast with previous research using different data and sources of variation highlights the need for continued research on conditional cash transfers. Overall, the findings underscore the broader context in shaping anti-poverty programs while accounting for the timing of interventions and family responses to workfare.

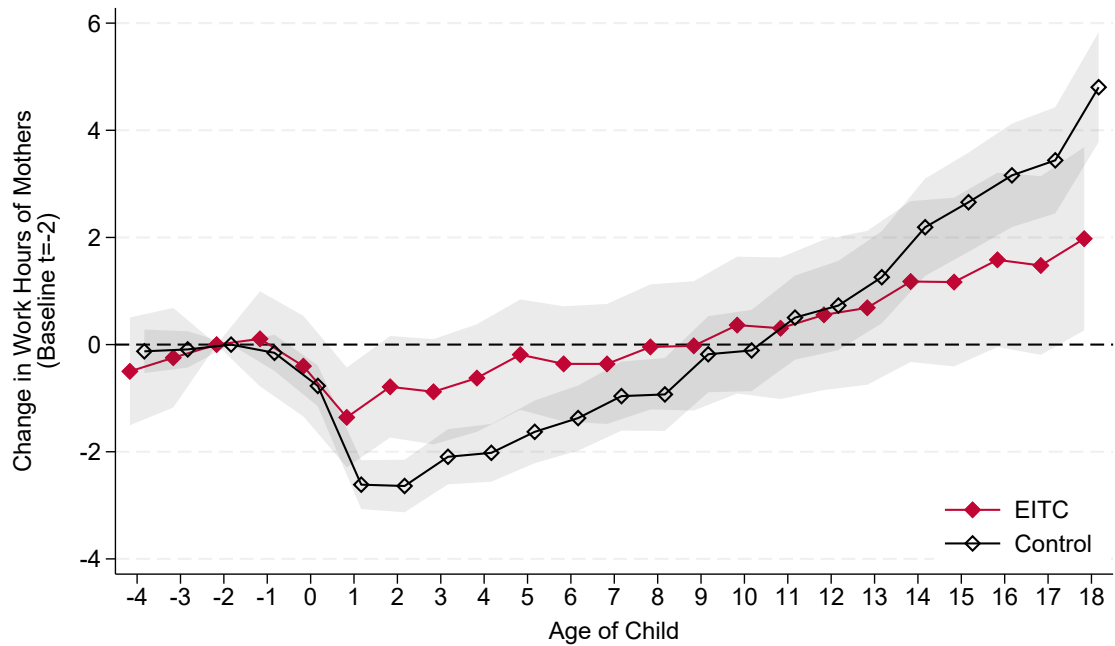


Figure 2.8: Maternal labor supply after birth by EITC status.

Notes: The figure plots the coefficients resulting from two two-way fixed effects regression. The coefficients capture the correlation between weekly work hours of mothers and the age of children. The first regression includes a sample of households that are EITC eligible (red) while the second is the sample of non eligible households (blue). The 95% confidence interval is constructed with robust standard errors clustered at the state-level.

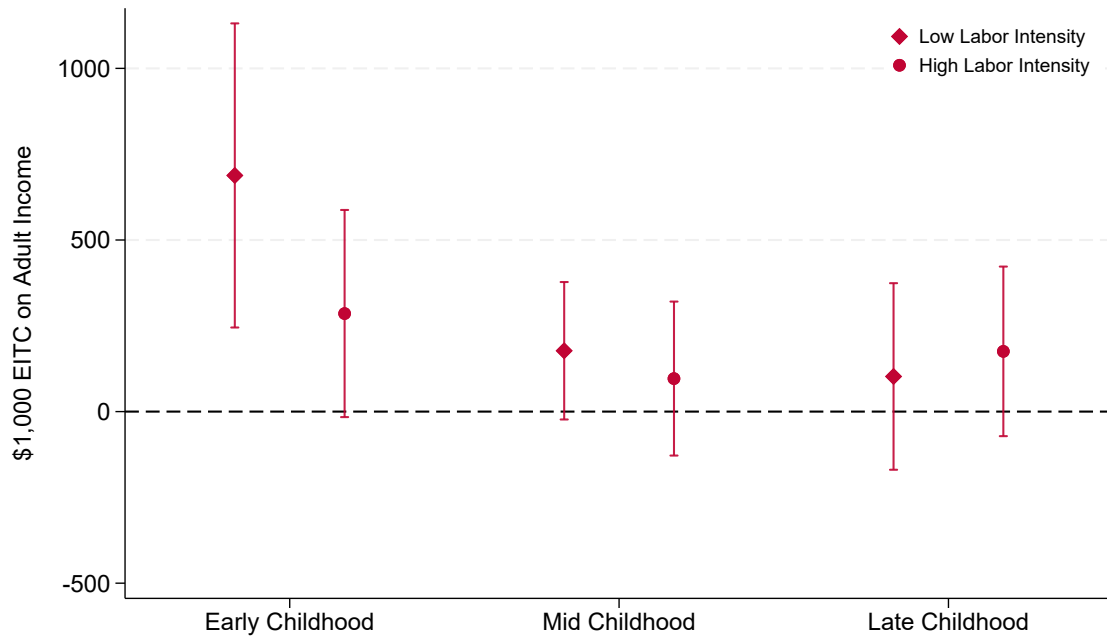


Figure 2.9: Intergenerational effects of the EITC on adult income conditional on the predicted maternal labor intensity.

Notes: The figure plots the coefficients of \$1,000 in imputed tax credit transfers on adult income conditional on maternal work intensity. The vertical axis is the size of coefficients. The 95 percent confidence intervals are clustered at the state level. The horizontal axis is the childhood stage in which tax credits are received. Low labor intensity (red diamond) are children whose mothers are below the median in the distribution of predicted work hours. High labor intensity (red circle) are children whose mothers are above the median in the distribution of predicted work hours. Maternal work hours are predicted in a regression with endowed characteristics: race, ethnicity, education, state and year fixed effects. The regressions control for mother fixed effects, cohort fixed effects, quadratic and cubic age in 2018, indicators for black, Hispanic, female, whether the mother has been married by age 18, an indicator for whether the mother graduated high school, an indicator for whether the mother has some college education, and fixed effects for the number of siblings at age 18, the three-way interactions between race, ethnicity, and sex by number of siblings, birth year and state. The specifications account for state policy variables measured at age 18 (GDP per capita, unemployment rate, the top marginal tax rate, the minimum wage, tax revenues, higher education spending, and the maximum welfare benefits), state fixed effects, and year of birth fixed effects.

References

- Autor, D., Figlio, D., Karbownik, K., Roth, J., and Wasserman, M. (2019). Family disadvantage and the gender gap in behavioral and educational outcomes. *American Economic Journal: Applied Economics*, 11(3):338–381.
- Autor, D. H. and Duggan, M. G. (2007). Distinguishing income from substitution effects in disability insurance. *American Economic Review*, 97(2):119–124.
- Averett, S. and Wang, Y. (2018). Effects of higher eitc payments on children’s health, quality of home environment, and noncognitive skills. *Public Finance Review*, 46(4):519–557.
- Banerjee, A., Niehaus, P., and Suri, T. (2019). Universal basic income in the developing world. *Annual Review of Economics*, 11:959–983.
- Barr, A., Eggleston, J., and Smith, A. A. (2022). Investing in infants: The lasting effects of cash transfers to new families. *The Quarterly Journal of Economics*, 137(4):2539–2583.
- Bastian, J. (2017). Unintended Consequences? More Marriage, More Children, and the EITC. In *Proceedings. Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association*, volume 110, pages 1–56.
- Bastian, J. (2020). The rise of working mothers and the 1975 earned income tax credit. *American Economic Journal: Economic Policy*, 12(3):44–75.

- Bastian, J. and Black, D. (2022). Overcoming Financial Constraints and Migrating Out of Rural and Distressed America. Unpublished Manuscript. Available at SSRN 4166950.
- Bastian, J. and Lochner, L. (2022). 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.
- Bastian, J. and Micheltore, K. (2018). The long-term impact of the earned income tax credit on children’s education and employment outcomes. *Journal of Labor Economics*, 36(4):1127–1163.
- Baughman, R. and Dickert-Conlin, S. (2003). Did expanding the eitc promote motherhood? *American Economic Review*, 93(2):247–251.
- Baughman, R. and Dickert-Conlin, S. (2009). The earned income tax credit and fertility. *Journal of Population Economics*, 22(3):537–563.
- Bertrand, M. and Pan, J. (2013). The trouble with boys: Social influences and the gender gap in disruptive behavior. *American Economic Journal: Applied Economics*, 5(1):32–64.
- Besley, T. and Coate, S. (1992). Workfare versus welfare: Incentive arguments for work requirements in poverty-alleviation programs. *The American Economic Review*, 82(1):249–261.
- Blau, D. and Currie, J. (2006). Pre-school, day care, and after-school care: who’s minding the kids? *Handbook of the Economics of Education*, 2:1163–1278.
- Braga, B., Blavin, F., and Gangopadhyaya, A. (2020). The long-term effects of childhood exposure to the earned income tax credit on health outcomes. *Journal of Public Economics*, 190:104249.

- Brown, D. W., Kowalski, A. E., and Lurie, I. Z. (2015). Medicaid as an investment in children: what is the long-term impact on tax receipts? Technical report, National Bureau of Economic Research.
- Bruce, D. (2002). Taxes and entrepreneurial endurance: Evidence from the self-employed. *National Tax Journal*, pages 5–24.
- Caetano, C., Caetano, G., Nielsen, E., and Sanfelice, V. (2021). The Effect of Maternal Labor Supply on Children: Evidence from Bunching. Unpublished Manuscript.
- Carneiro, P., García, I. L., Salvanes, K. G., and Tominey, E. (2021). Intergenerational mobility and the timing of parental income. *Journal of Political Economy*, 129(3):757–788.
- CBPP (2022). Policy Basis: The Supplemental Nutrition Assistance Program (SNAP). <https://www.cbpp.org/research/food-assistance/the-supplemental-nutrition-assistance-program-snap>. Center on Budget and Policy Priorities.
- Centers for Medicare & Medicaid Services (2022). Financial Management Report for FY 2021. <https://www.medicaid.gov/medicaid/financial-management/state-expenditure-reporting-for-medicaid-chip/expenditure-reports-mbescbes/index.html>.
- Cesur, R., Rodríguez-Planas, N., Roff, J., and Simon, D. (2022). Domestic violence and income: Quasi-experimental evidence from the earned income tax credit. Technical report, National Bureau of Economic Research.
- Chetty, R., Friedman, J. N., and Rockoff, J. (2011). New evidence on the long-term impacts of tax credits. *IRS Statistics of Income White Paper*.
- Chetty, R., Friedman, J. N., and Saez, E. (2013). Using differences in knowledge across neighborhoods to uncover the impacts of the eitc on earnings. *American Economic Review*, 103(7):2683–2721.

- Chetty, R., Hendren, N., Lin, F., Majerovitz, J., and Scuderi, B. (2016). Childhood environment and gender gaps in adulthood. *American Economic Review*, 106(5):282–288.
- Cooksey, E. C. (2018). Using the national longitudinal surveys of youth (nlsy) to conduct life course analyses. *Handbook of Life Course Health Development*, pages 561–577.
- Crandall-Hollick, M. L. (2018). The earned income tax credit (eitc): A brief legislative history. *Congressional Research Service*.
- Cummins, J. G., Hassett, K. A., Hubbard, R. G., Hall, R. E., and Caballero, R. J. (1994). A reconsideration of investment behavior using tax reforms as natural experiments. *Brookings Papers on Economic Activity*, 1994(2):1–74.
- Cunha, F. and Heckman, J. (2007). The technology of skill formation. *American Economic Review*, 97(2):31–47.
- Currie, J. and Gruber, J. (1996). Saving babies: The efficacy and cost of recent changes in the medicaid eligibility of pregnant women. *Journal of Political Economy*, 104(6):1263–1296.
- Dahl, G. B. and Lochner, L. (2012). The impact of family income on child achievement: Evidence from the earned income tax credit. *American Economic Review*, 102(5):1927–56.
- Dahl, M., DeLeire, T., and Schwabish, J. (2009). Stepping stone or dead end? the effect of the eitc on earnings growth. *National Tax Journal*, 62(2):329–346.
- Dench, D. and Joyce, T. (2020). The earned income tax credit and infant health revisited. *Health Economics*, 29(1):72–84.
- Dickert-Conlin, S. and Houser, S. (2002). Eitc and marriage. *National Tax Journal*, pages 25–40.

- Doyle, O., Harmon, C., Heckman, J. J., Logue, C., and Moon, S. H. (2017). Early skill formation and the efficiency of parental investment: a randomized controlled trial of home visiting. *Labour Economics*, 45:40–58.
- Edmonds, A. T., Moe, C. A., Adhia, A., Mooney, S. J., Rivara, F. P., Hill, H. D., and Rowhani-Rahbar, A. (2022). The earned income tax credit and intimate partner violence. *Journal of Interpersonal Violence*, 37(13-14):NP12519–NP12541.
- Eissa, N. and Hoynes, H. W. (2004). Taxes and the labor market participation of married couples: the earned income tax credit. *Journal of Public Economics*, 88(9-10):1931–1958.
- Eissa, N. and Hoynes, H. W. (2006). Behavioral responses to taxes: Lessons from the EITC and Labor Supply. *Tax Policy and the Economy*, 20:73–110.
- Feenberg, D. and Coutts, E. (1993). An introduction to the TAXSIM model. *Journal of Policy Analysis and Management*, 12(1):189–194.
- Fetter, D. K. and Lockwood, L. M. (2018). Government old-age support and labor supply: Evidence from the old age assistance program. *American Economic Review*, 108(8):2174–2211.
- Goodman-Bacon, A. (2021). The long-run effects of childhood insurance coverage: Medicaid implementation, adult health, and labor market outcomes. *American Economic Review*, 111(8):2550–2593.
- Gray, C., Leive, A., Prager, E., Pukelis, K., and Zaki, M. (2023). Employed in a snap? the impact of work requirements on program participation and labor supply. *American Economic Journal: Economic Policy*, 15(1):306–341.
- Gruber, J. and Saez, E. (2002). The elasticity of taxable income: evidence and implications. *Journal of Public Economics*, 84(1):1–32.

- Han, J. (2022). The impact of snap work requirements on labor supply. *Labour Economics*, 74:102089.
- Hardy, B., Hokayem, C., and Ziliak, J. P. (2022). Income inequality, race, and the etc. *National Tax Journal*, 75(1):149–167.
- Herbst, C. M. (2011). The impact of the Earned Income Tax Credit on marriage and divorce: Evidence from flow data. *Population Research and Policy Review*, 30(1):101–128.
- Hotz, V. J. and Scholz, J. K. (2001). The earned income tax credit. Technical report, National Bureau of Economic Research.
- Hoynes, H., Miller, D., and Simon, D. (2015). Income, the earned income tax credit, and infant health. *American Economic Journal: Economic Policy*, 7(1):172–211.
- Hoynes, H. and Rothstein, J. (2019). Universal basic income in the united states and advanced countries. *Annual Review of Economics*, 11:929–958.
- Hoynes, H., Schanzenbach, D. W., and Almond, D. (2016). Long-run impacts of childhood access to the safety net. *American Economic Review*, 106(4):903–934.
- Hoynes, H. W. and Schanzenbach, D. W. (2012). Work incentives and the food stamp program. *Journal of Public Economics*, 96(1-2):151–162.
- Hoynes, H. W. and Schanzenbach, D. W. (2018). Safety net investments in children. Technical report, National Bureau of Economic Research.
- IRS (2022). EITC Reports and Statistics.
- Isaac, E. (2020). Marriage, divorce, and social safety net policy. *Southern Economic Journal*, 86(4):1576–1612.
- Jiao, Y., Lu, Y., and Zhang, X. (2022). Racial disparities in labor market responses to the etc. *Available at SSRN*.

- Jones, M. R. and Ziliak, J. P. (2022). The antipoverty impact of the eitc: New estimates from survey and administrative tax records. *National Tax Journal*, 75(3):451–479.
- Karademir, S., Laliberté, J.-W. P., and Staubli, S. (2024). The multigenerational impact of children and childcare policies. Technical report, National Bureau of Economic Research.
- Kleven, H. (2019). The eitc and the extensive margin: A reappraisal. Technical report, National Bureau of Economic Research.
- Kleven, H. J. (2016). Bunching. *Annual Review of Economics*, 8:435–464.
- Kuka, E. and Shenhav, N. (2020). Long-Run Effects of Incentivizing Work After Childbirth. Technical report, National Bureau of Economic Research.
- Leung, J. H. and Seo, H. K. (2023). How do government transfer payments affect retail prices and welfare? evidence from snap. *Journal of Public Economics*, 217:104760.
- Manning, W. D. (1990). Parenting employed teenagers. *Youth & Society*, 22(2):184–200.
- Manoli, D. and Turner, N. (2018). Cash-on-hand and college enrollment: Evidence from population tax data and the earned income tax credit. *American Economic Journal: Economic Policy*, 10(2):242–71.
- McInnis, N. S., Micheltore, K., and Pilkauskas, N. (2023). The intergenerational transmission of poverty and public assistance: Evidence from the earned income tax credit. Technical report, National Bureau of Economic Research.
- Meyer, B. D. (2002). Labor supply at the extensive and intensive margins: The EITC, welfare, and hours worked. *American Economic Review*, 92(2):373–379.
- Micheltore, K. and Lopoo, L. M. (2021a). The effect of eitc exposure in childhood on marriage and early childbearing. *Demography*, 58(6):2365–2394.

- Micheltmore, K. and Lopoo, L. M. (2021b). Exposure to the earned income tax credit in early childhood and family wealth. *RSF: The Russell Sage Foundation Journal of the Social Sciences*, 7(3):196–215.
- Micheltmore, K. and Pilkauskas, N. (2021). Tots and teens: How does child’s age influence maternal labor supply and child care response to the earned income tax credit? *Journal of Labor Economics*, 39(4):895–929.
- Mincer, J. A. (1974). Age and experience profiles of earnings. In *Schooling, Experience, and Earnings*, pages 64–82. NBER.
- Mortenson, J. A. and Whitten, A. (2020). Bunching to maximize tax credits: Evidence from kinks in the us tax schedule. *American Economic Journal: Economic Policy*, 12(3):402–432.
- Neumark, D. and Shirley, P. (2020). The long-run effects of the Earned Income Tax Credit on women’s labor market outcomes. *Labour Economics*, 66:101878.
- Noble, K. G. (2021). Early Childhood: The Opportunity to Untap Human Potential. *The Scale-Up Effect in Early Childhood and Public Policy: Why Interventions Lose Impact at Scale and What We Can Do About It*, page 2.
- Noble, K. G., Houston, S. M., Brito, N. H., Bartsch, H., Kan, E., Kuperman, J. M., Akshoomoff, N., Amaral, D. G., Bloss, C. S., Libiger, O., et al. (2015). Family income, parental education and brain structure in children and adolescents. *Nature Neuroscience*, 18(5):773.
- Pilkauskas, N. and Micheltmore, K. (2019). The effect of the earned income tax credit on housing and living arrangements. *Demography*, 56(4):1303–1326.
- Pilkauskas, N. V. and Micheltmore, K. (2023). Who’s caring for the kids? the earned income tax credit and childcare arrangements. *The ANNALS of the American Academy of Political and Social Science*, 706(1):37–64.

- Robinson, S. and Tazhitdinova, A. (2023). What drives tax policy? political, institutional and economic determinants of state tax policy. Technical report, National Bureau of Economic Research.
- Rodgers, L. P. (2018). Give credit where? the incidence of child care tax credits. *Journal of Urban Economics*, 108:51–71.
- Rodgers, L. P. (2023). Tax credit refundability and child care prices: Evidence from california. *National Tax Journal*, 76(1):95–116.
- Romeo, R. R., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., Rowe, M. L., and Gabrieli, J. D. (2018). Beyond the 30-million-word gap: Children’s conversational exposure is associated with language-related brain function. *Psychological science*, 29(5):700–710.
- Saez, E. (2010). Do taxpayers bunch at kink points? *American Economic Journal: Economic Policy*, 2(3):180–212.
- Sanderson, E. and Windmeijer, F. (2016). A weak instrument f-test in linear iv models with multiple endogenous variables. *Journal of Econometrics*, 190(2):212–221.
- Schanzenbach, D. W. and Strain, M. R. (2021). Employment effects of the earned income tax credit: Taking the long view. *Tax Policy and the Economy*, 35(1):87–129.
- Social Security Administration (2021). SSI Annual Statistical Report, 2020. https://www.ssa.gov/policy/docs/statcomps/ssi_asr/.
- Suskind, D., Suskind, B., and Lewinter-Suskind, L. (2015). *Thirty million words: Building a child’s brain*. Dutton.

Urban Institute (2022). Housing and Community Development Expenditures.
<https://www.urban.org/policy-centers/cross-center-initiatives/state-and-local-finance-initiative/state-and-local-backgrounders/housing-and-community-development-expenditures>.

Wolfers, J. (2006). Did unilateral divorce laws raise divorce rates? a reconciliation and new results. *American Economic Review*, 96(5):1802–1820.

CHAPTER 3

Business as usual? Firm Responses to the EITC

The Earned Income Tax Credit (EITC) represents one of the largest antipoverty programs in the United States, providing refundable tax credits to low-income working families. The literature has thoroughly debated the role of the program in incentivizing employment and supplementing incomes for qualifying households (Schanzenbach and Strain, 2021; Nichols and Rothstein, 2015; Eissa and Hoynes, 2004). However, the interaction between social programs and labor markets poses key questions about economic incidence. Standard economic theory implies that increasing the labor supply of low-skill workers through such transfers should exert downward pressure on equilibrium wages (Neumark and Wascher, 2011; Leigh, 2010; Rothstein, 2010; Dickert-Conlin and Holtz-Eakin, 1999). Previous empirical assessments of the EITC’s effects have been limited to household survey data, leaving open questions about how the program shapes outcomes observed at the firm and market level. This paper presents novel evidence of the tax credit effects on wages, employment, and the number of establishments across industries and occupations. Furthermore, I estimate the first industry and occupation-specific estimates of the tax credit, showing that the labor impact of the program is concentrated in certain industry groups and occupations.

This paper bridges a gap in the literature on the labor market effects of the EITC, sparked by recent doubts on otherwise established findings. The general consensus maintains that the program causes extensive margin responses in maternal labor supply (Schanzenbach and Strain, 2021; Neumark and Shirley, 2020; Nichols and Rothstein, 2015; Eissa and Hoynes, 2004). While less striking, evidence exists that reported work hours also increase (Bastian and Lochner, 2020; Meyer, 2002; Meyer and Rosenbaum, 2001). The shifting labor supply should exert influence on wages and produce negative spillovers to non-targetted groups. In that sense, a small subset of studies finds that a 10% increase in state generosity in the tax credit decreases wages and salaries by 5% (Leigh, 2010). Wages tend to decrease for non-targeted groups, such as childless women (Rothstein, 2010) and workers without a college degree (Neumark and Wascher, 2011; Leigh, 2010). In principle, changes in

wages substantiate the hypothesis of labor supply shifts. Despite the overwhelming consensus, [Kleven \(2023\)](#) argues that studies on the labor market impact of the tax credit confound other social programs expanded in the 1990 Welfare Reform. The thorough critique casts doubts on the labor market effects of the EITC and the best methodologies to evaluate the program.

Evidence on the effectiveness of the tax credit relies uniquely on household responses in the Current Population Survey, American Community Survey, and panel surveys. Inference from household surveys is challenged by omitted variables, simultaneity between household behavior and eligibility criteria, and undefined treatment groups among recipients because take-up is unobserved in most data. Program evaluations typically employ two designs: one approach uses variation in the maximum cash transfer available according to the number of children in a household in a given year and state, a measure of generosity labelled as EITC exposure ([Micheltore and Lopoo, 2021a](#); [Bastian and Lochner, 2020](#)); and another approach is a difference-in-difference strategy using time variation after the 1996 Welfare Reform and differences in the treatment status of mothers defined by the number of children and a proxy for take-up (for example, cognitive scores ([Isaac, 2020](#)) or academic achievement ([Eissa and Hoynes, 2004](#))). As an unexplored data source, establishment surveys and industry-level exposure contribute novel insights to the literature.

This paper documents the relationship between state expansions in the EITC and labor markets using industry and occupational data. The analysis exploits variation in state's roll-out of EITC supplements to estimate their impact on wages, employment, and the number of establishments across industries and occupations. The research design leverages establishment surveys by the Bureau of Labor Statistics: The Quarterly Census of Employment and Wages (QCEW) and the Occupational Employment and Wage Statistics Research Estimates (OEWS). These establishment surveys record labor data across industries and occupations, respectively, aggregated at the state level. My empirical framework exploits differences in exposure across industries and occupations, computed with household data on the number of

program eligible individuals working in this industry or occupation in the American Community Survey (ACS). In a nutshell, the triple difference estimator uses variation across states with EITC supplements, differences in roll-out timing, and differences in national exposure across industries and occupations.

The identifying assumptions are tested with a battery of robustness tests. First, event studies show that outcome measures had similar pre-trends across industries and occupations before enacting the program. Second, the TWFE literature warns of comparisons with inadequate baselines when there is variation in timing ([Goodman-Bacon, 2021](#)). My results are consistent with robust methodologies ([Sun and Abraham, 2021](#)). Third, identification requires that the policy timing is independent of industry and state characteristics. Following studies with similar sources of variation ([East, 2020](#)), I test the correlation between treatment timing and state and policy variables.

The results indicate that, after the enactment of state EITC programs, wages decrease between 2 and 8 percent across exposed industries and occupations. However, the impact on wages appears to be transitory, and wages stabilize over time. There is little evidence of employment effects, as firms in exposed industries do not report an increase in the number of workers. Furthermore, the number of establishments remains stable post-program implementation.

This paradox —decreasing wages and stable employment— suggests that workfare programs may increase willingness to work without shifting equilibrium quantities, such that increased competition in these markets depresses wages without improving employment. Self-employment, the elasticity of labor demand, or tight labor markets in these industries may be critical factors; in cases where demand is inelastic, industries can absorb the benefits of the policy through lower labor costs without adjustments in employment. The analysis reveals that wage adjustments are primarily concentrated in service industries and occupations, which are negatively correlated with EITC exposure, whereas there is a negligible correlation between exposure and changes in employment and the number of establishments. While employment effects

are inconclusive, the significant wage reductions underscore the program’s influence on labor markets, with pronounced effects in the service economy.

This paper contributes to three strands of the literature. First, it provides novel estimates of the program’s labor market impacts based on new data and unique sources of variation. This study adds evidence to a contentious literature with mixed results on the labor market impact of one of the largest social programs in the US (Kleven, 2023; Nichols and Rothstein, 2015; Eissa and Hoynes, 2004; Eissa and Liebman, 1996). The mixed results may be driven by elasticities in affected industries, an unexplored margin that would benefit from further research. Second, this paper offers the first estimates of heterogeneous effects of the EITC on industries and occupations. Previous studies have documented that the program drives wages down, finding spillovers to non-targeted workers (Neumark and Wascher, 2011; Leigh, 2010; Rothstein, 2010). These studies rely on variation in state generosity, while this paper exploits variation in roll-out and rich industry and occupational data. Third, this paper complements one study on the economic incidence of conditional cash transfers exploiting firms’ data. In particular, Zurla (2024) studies an EITC-like program in Italy dubbed the monthly EUR 80 bonus. However, the institutional background differs in this settings, leading to differences in the magnitude of results. The Italian program is managed directly by firms that observe who the recipients of the program are, allowing them to capture up to 30% of the subsidy. US firms are unable to identify treated households, and they cannot leverage participants’ information to push for lower wages.

The underlying economics highlights open questions about the economic incidence and benefits of the program. The EITC is an anti-poverty program that has lifted millions of households out of poverty (Jones and Ziliak, 2022; Hoynes and Patel, 2018; Neumark and Wascher, 2001), with positive effects on health outcomes (Braga et al., 2020; Markowitz et al., 2017; Hoynes et al., 2015; Evans and Garthwaite, 2014), education (Bastian and Micheltore, 2018; Manoli and Turner, 2018; Micheltore, 2013; Chetty et al., 2011), and wealth and earnings (Micheltore and Lopoo, 2021b;

Chetty et al., 2013). The evidence of falling wages and unresponsive employment poses an anomaly as to whether increased willingness to work led to more competitive low-skill markets without additional employment opportunities. Some possible explanations relate to tight labor markets with few vacancies, inelastic labor demand in these industries, and concentrated employment effects on self-employed mothers (Lim and Micheltore, 2018; Kuka, 2014; LaLumia, 2009; Saez, 2010). These are important avenues for future research to understand the economic incidence and capture by firms.

3.1 The EITC supplement programs

The federal government initiated the EITC program in 1979. Since its inception, nine reforms have expanded the targeted population. In principle, the program was part of a broader debate aimed at targeting entitlements for worker families and reducing welfare dependence.¹ The federal EITC is a refundable cash transfer, meaning that an eligible household with no tax liabilities receives a transfer from the Internal Revenue Service (IRS) after filing their taxes.² The EITC targets low-income families, who may apply upon filing their regular taxes (using line 27a in Form 1040 and an additional form disclosing the number of qualifying children). The EITC schedule determines the magnitude of the transfer and the brackets that deem a taxpayer as eligible.³

¹In contrast to a Negative Income Tax, the tax credit proposed by Senator Russell Long introduced a "work bonus" aiming to reduce welfare dependence and increase employment.

²The federal cash transfer is annual. Since 1978, the option existed for individuals to receive monthly disbursements instead of annual. Between 2002 and 2004, the Advanced Earned Income Tax Credit (AEITC) had a low take-up (around 3% of eligible taxpayers) and 80% noncompliance (GAO, 2007). The AEITC program was eliminated in 2010. Among states, most programs are annual. The District of Columbia disburses approximately 20% of the state credit monthly. In 2014, the Center for Economic Progress, Government of Chicago, the Chicago Housing Authority and the University of Illinois at Urbana-Champaign designed a year-long experiment, the Periodic Earned Income Tax Credit Payment Pilot Project, that disbursed up to 50% of the EITC in advance installments to a cohort of 343 participants (Andrade et al., 2019).

³The generosity of the cash transfer is a function of the taxpayer's adjusted gross income and the number of eligible children. While the brackets and maximum eligible adjusted gross income depend on family size and marital status. Every January, the IRS publishes the most recent tax schedule. The program is designed with different incentives across the schedule. In the phase-in, the

In addition to the federal government, states have passed their own supplementary EITC programs. Table 3.1 lists the roll-out year of all EITC state supplement programs.⁴ Rhode Island was the first state to offer a nonrefundable EITC since 1975, but it was not enacted into law until 2001 and, instead, was part of the unique income tax structure in the state (LaLumia, 2009). In 1984, Wisconsin instituted the first supplementary tax credit. Today, thirty-one states and the District of Columbia offer state tax credits matching a portion of the federal EITC.

These programs are heterogeneous and often marked by a contentious legislative history. The programs tend to be refundable except in the cases of Missouri, Ohio, South Carolina, and Utah, where the tax credit only offsets the taxpayer’s liabilities but does not produce an additional transfer.⁵ Most states set their supplement programs as a percent of the federal EITC, while few others have their own income-based criteria to determine the magnitude of the credit. Eligibility conditions also vary: California has a narrower income cap than the federal program, and the program is not available to childless taxpayers in Wisconsin. California, Colorado, Maryland, Minnesota, Vermont, and Washington allow individuals with a valid taxpayer identification number to receive the benefit. Indiana, Ohio, and Oregon

cash transfer increases linearly subsidizing pre-tax earnings for any income above zero to incentivize non-working taxpayers to seek employment. The plateau is the area where the benefit is at its maximum, providing the largest amount of support to a low-income household. Lastly, the schedule includes a phase-out that gradually reduces the subsidy to move households out of welfare.

⁴The non-partisan initiative, Tax Credits for Workers and Families (TCWF), offers a legislative history of state and local tax credits (TCWF, 2023). In this section, I draw from their detailed resources, state legislation, the National Bureau of Economic Research state EITC provisions (NBER, 2023), the Center on Budget and Policy Priorities (Waxman and Hinh, 2023), the Tax Policy Center (TPC, 2023), the Institute on Taxation and Economic Policy (ITEP, 2019), and other sources to summarize the history of these programs.

⁵Progressivity may be hindered by nonrefundability. For example, the case of the nonrefundable Child Tax Credit (CTC) is well-documented in Goldin and Michelmore (2022). The bottom 10 percent of income is ineligible for the CTC, as they do not have sufficient tax liabilities to offset with the program. Black children represent 22 percent of all children but only 14 of eligible children. Simulations show that 99 percent of Black and Hispanic children would be eligible if the CTC was fully refundable. These disparities sparked an important debate on the program’s incidence (Brehm and Malkova, 2023).

introduced the EITC while offsetting part of the program with tax increases. In these cases, state tax reforms increased the total tax rate on single women with children.⁶

State supplements also exhibit tumultuous politics. Political landscapes have driven some states to repeal or reinstate the policy. For example, the Wisconsin tax credit had a short lifespan, beginning in 1984 with the 1983 Wisconsin Act 27 and lasting only two years before being repealed. In 1986, Wisconsin re-enacted the program, and, this time, it became permanent. North Carolina implemented a refundable tax credit program in 2007 as a 3.5% supplement of the federal transfer, which was soon expanded to 5% in 2008. However, the state eliminated the tax credit in 2014. Another relevant example is Colorado, which enacted a supplementary EITC in 1999 contingent on the state's revenue. Colorado did not disburse the credit due to lacking funds from 2001 until 2013 when lawmakers reformed the program, making it permanent.

State variation in the tax credit has been a key feature to evaluate the program. Cross-state variation is pivotal in calculating measures of family exposure based on the combined federal and state tax credits at the plateau of the schedule ([Agan and Makowsky, 2023](#); [Micheltmore and Lopoo, 2021a,b](#); [Braga et al., 2020](#); [Bastian and Micheltmore, 2018](#); [Leigh, 2010](#); [Neumark and Wascher, 2001](#)). Other state-level approaches and specifications use the phase-in rate, indicators on the policy availability or variation in other parameters ([Agan and Makowsky, 2023](#); [Neumark and Shirley, 2017](#)). Reevaluating the impact on the extensive margin of the labor supply, [Kleven \(2023\)](#) informs the literature by examining federal and state reforms using differences in marital status and fertility across states with and without supplementary programs. The paper finds that, other than the 1993 reform, federal and state expansions in the tax credit had limited or insignificant effects on the labor supply of women ([Kleven, 2023](#)).

⁶In addition to state tax credits, local policies have emerged over time. Some geographies in the US offer local EITC programs: Denver, Chicago, Maryland, New York City, San Francisco, and Philadelphia. Their political success has been limited due to lacking funds.

Table 3.1: Roll-out year in state supplementary EITC programs.

State	Non-Refundable	Refundable
California		2015
Colorado		1999
Connecticut		2011
Delaware	2006	2020
D.C.		2001
Hawaii	2017	2020
Illinois	2000	2003
Indiana		1999
Iowa	1990	2007
Kansas		1998
Louisiana		2007
Maine	2000	2016
Maryland	1987	1998
Massachusetts		1997
Michigan		2006
Minnesota		1991
Montana		2019
Nebraska		2006
New Jersey		2000
New Mexico		2007
New York	1994	1994
North Carolina		2008
Ohio	2013	
Oklahoma	2016	2002
Oregon	1997	2006
Rhode Island	2001 (1975)	2015
South Carolina	2018	
Utah	2022	
Vermont		1988
Virginia	2000	2022
Washington		2022
Wisconsin	1984	1989

Notes: The California EITC is estimated on the basis of income instead of a percentage of the Federal EITC. Colorado enacted the law in 1999 but it was contingent on budget surplus. The credit was not paid between 2001 and 2014. Indiana’s EITC was created in 1999 but was not based on the federal credit. In 2003, Indiana switched to an EITC model as a percentage of the federal credit. Michigan passed the EITC in 2006 but the legislature compromised reduced the rate after almost eliminating the tax credit in 2011. It was recently increased to 30% of the Federal credit in 2023. Minnesota tax credit is the Working Family Credit (WFC). The credit is computed as a percentage of income instead of based on the Federal EITC since 1998. New York offers a fully refundable tax credit to full year residents, and a partially refundable EITC to part-year residents. North Carolina overturned the tax credit after 2014 and a tax cut was approved instead. Rhode Island was the first state to offer a nonrefundable state EITC at 17%. However, it was not enacted into law until 2001. Washington state is the only state with an EITC program and without an income tax. The Washington tax credit is a lump-sum amount conditional on family size, automatically available to taxpayers who qualify for the federal credit.

The labor effects of the EITC program are an empirical question, as theory is ambiguous on the program's impact. The program is expected to raise wages above reservation wages of non-workers, driving them into the labor force. The impact on workers is nuanced and depends on income and substitution effects, which may lead affected workers to reduce their hours of work. For these reasons, the literature has empirically differentiated between extensive and intensive margin effects of the program (Eissa et al., 2006; Meyer, 2002). The theoretical shift in the labor supply following expansions in the tax credit program is theoretically expected to decrease wages (Neumark and Wascher, 2011; Rothstein, 2010; Leigh, 2010; Dickert-Conlin and Holtz-Eakin, 1999). While the tax credit program is targeted to low-income families with children, labor market distortions affect targeted and non-targeted workers in low-paying jobs, generating negative spillovers (Rothstein, 2010; Leigh, 2010). These considerations spark concerns about the best policy design supporting low-income workers, such as employee or employer-based subsidies (Dickert-Conlin and Holtz-Eakin, 1999).

The impact of the tax credit subsidy on firms has seldom been documented. Zurla (2024) is one exception, which leverages rich administrative data on Italian firms paying an EITC-like work bonus. In this setting, firms respond to the subsidy by decreasing the earnings of individuals by 2%. In this Italian setting, government authorities channel the cash transfer through firms, allowing them to leverage the information and appropriate part of the subsidy. This is a natural consequence of the role of firms in wage setting (Card et al., 2018, 2013). The role of the tax credit on establishments and firm dynamics is also unexplored, although several channels may lead the work subsidy to affect firm dynamics. First, the EITC increases the incentives to generate income, but mothers may prefer flexible work arrangements derived from self-employment (Kuka, 2014; LaLumia, 2009). Alternatively, similar to immigration and other labor supply shifts, tax credits may boost job creators and raise initial capital and start new businesses (Azoulay et al., 2022; Deming, 2022).

3.2 Empirical Strategy

This paper seeks to estimate a setting where variation in firm responses is explained by variation in the EITC across states and years. The outcome Y captures the number of employees, mean wages, and the number of establishments in an industry j , state s , and year t . On the right-hand side, the Two-Way Fixed Effects (TWFE) approach uses the timing of enactment in state supplements $D_{st} = (\text{EITC}_{st} > 0)$. The estimating equation is:

$$Y_{jst} = \delta D_{st} + \beta X_{jst} + \phi_j + \phi_s + \phi_t + \epsilon_{jst} \quad (3.1)$$

The parameter of interest is δ , representing the association between firm responses and the enactment of a state EITC program. The specification controls for industry-specific trends using the fixed effect ϕ_j , while state fixed effects ϕ_s and time fixed effects ϕ_t control for idiosyncratic heterogeneities across states and macroeconomic shocks, respectively. To mitigate omitted variable bias, the vector X_{jst} includes time-varying sociodemographic characteristics from survey data in each state, and policy variables including the unemployment rate, state per capita GDP, minimum wage, and state marginal tax rate.⁷

One issue with Equation 3.1, central to the research question in this paper, is that treatment effects are not homogeneous across industries. Some industries are expected to have a larger proportion of working low-income mothers or welfare recipients than others. The literature, focused on household surveys, does not have this problem; as treatment is defined according to policy criteria: low-educated mothers (Eissa and Hoynes, 2004), AFQT scores (Isaac, 2020), fertility (Micheltmore and Lopoo, 2021a; Bastian and Micheltmore, 2018), or other characteristic that captures likely welfare eligibility. Therefore, a new measurement of industry exposure must be constructed.

⁷Socioeconomic characteristics include the percentage of female workers, average race and ethnicity, educational achievement, number of children, median family income, and age.

Let T_j represent an indicator variable that captures whether the percentage of workers in an industry or occupation eligible to the program is above a threshold.⁸ This threshold is computed as a percentile in the distribution of EITC eligible workers across the country. Exposure at aggregate levels avoids the influence of confounders related to state generosity in the program and allows for estimates that are independent of whether workers in that industry and state were included in the census. In addition, assume that exposure is time invariant for simplicity.⁹ We therefore develop the following framework:

$$Y_{jst} = \gamma D_{st} \times T_j + \zeta D_{st} + \rho X_{jst} + \theta_j + \theta_s + \theta_t + \mu_{jst} \quad (3.2)$$

Causal identification relies on several assumptions. First, the parallel trend assumption argues that non-exposed industries in untreated states would have similar trends in their outcomes than exposed industries in treated states. I test this with event study designs. Additionally, identification requires the timing of enactment to be as good as random, conditional on year and state fixed effects that account for shocks that may drive policy roll-out. This assumption may be tested, as in previous research studying the state roll-out of other social programs such as the Supplemental Nutrition Assistance Program ([East, 2020](#)). To do so, I estimate the relationship between the probability of enacting the policy and a series of state and state-policy controls. This provides evidence on whether policymakers strategically rely on the policy to ameliorate weathering economic trends or, alternatively, policymakers pass the tax credit in wealthier geographies. Lastly, heterogeneous timing in treatment

⁸This section describes the setting based on industry exposure but estimates exploiting occupational variation are also calculated with this methodology.

⁹The assumption that industry exposure is time-invariant may be unrealistic. First, the EITC program has experienced important reforms across time aimed at expanding eligibility. For example, the inclusion of childless workers in the early 1990s. Second, female and maternal labor supply have experienced substantial changes in this Century. The EITC mainly targets low-income mothers. Therefore, industry exposure may have changed most for industries and occupations prone to the expansion of female labor supply during the 1990s. However, the assumption of time-invariant exposure is required due to data limitations, as household surveys do not track the sociodemographic characteristics of workers in all industries. In Appendix Figure [B.9](#), I show results are stable when using the percentage of workers in an industry across different years.

may suffer from biases if treatment effects are reliant on comparisons between treated units or units that will be treated during the sample (Goodman-Bacon, 2021). Therefore, I estimate weighted regressions robust to staggered treatment (Sun and Abraham, 2021).

3.3 Data

The Bureau of Labor Statistics (BLS) publishes the two sources of establishment data in this study: the Quarterly Census of Employment and Wages (QCEW) and the Occupational Employment and Wage Statistics Research Estimates (OEWS). The QCEW microdata is constructed using the unemployment insurance (UI) accounting system, which receives the Quarterly Contributions Report from all employers covered under the UI program. This dataset records a quarterly count of employment, wages, and establishments at the industry level in each state since 1990.¹⁰ The OEWS data is collected semiannually by the BLS and state workforce agencies, surveying a random sample of establishments by electronic means and personal visits. It provides annual wage and employment statistics, recording the distribution of hourly wages paid in each state across occupations since 1999.¹¹

I merge the BLS datasets with the 1% American Community Survey (ACS) using concordances between industry codes and occupation codes. Initially, the process involves matching *ind1990* codes with *sic4* using concordances listed in Lake and Millimet (2016). Then, I match SIC codes with NAICS-6 codes following concordances in Schaller and DeCelles (2021).¹² The matching process results in an 81% match of industries in the QCEW with workers in the ACS. In the case of OEWS,

¹⁰The QCEW has several aggregation levels. The sample is restricted to Statewide 4-digit NAICS code (56) and privately owned firms (ownership code 5). NAICS organizes businesses into economic sectors (2-digit), sub-sectors (3-digits), industry groups (4-digits), industries (5-digits), and national industries (6-digits). This empirical investigation is limited to industry groups and uses the terms “industry groups” and “industries” synonymously.

¹¹OEWS has been available since 1997 with different occupational codes.

¹²The sample is restricted to ACS observations with industry code or occupation code. Among observations with industry code, the process yields a full match to NAICS code.

only 48% occupations are matched with the ACS using the Standard Occupational Classification (SOC). The analytical sample comprises 387,000 observations at the industry level from 1990 to 2019 and 442,000 observations from 1999 to 2018 at the occupational level.¹³

Descriptive statistics at the industry and occupational level are presented in Table 3.2. The final analytical sample comprises 13,000 industries per year (250 industry groups in each state), with an average of five hundred establishments. On average, industries employ seven thousand individuals with a mean salary of \$860 per week. Approximately two thousand industries pay zero wages in a given year, and seven hundred with no establishments.¹⁴ Covariates from household characteristics in each industry are summarized in the table. On average, workers in each industry tend to be male, college-educated, and white, with a family income of \$72,000.¹⁵

Occupational characteristics are presented at the bottom of Table 3.2. There are 22 thousand occupations registered per year in all states (approximately four hundred occupations). On average, each occupation has five thousand workers and pays \$22 an hour. The OEWS publishes key characteristics in the wage distribution, such as the median wage paid by firms and selected percentiles. Each year, 3,500 observations are censored at zero. In comparison with the QCWE, the OEWS displays lower weekly wages and similar socioeconomic characteristics.

EITC exposure is estimated with socioeconomic characteristics in the ACS. To calculate exposure, I first compute the federal and state tax liabilities per person using TAXSIM and the family income, marital status, number of children, and state in household data (Feenberg and Coutts, 1993). Next, the average eligibility to the tax credit is estimated for every industry and occupation. Therefore, my measurement

¹³The sample is restricted to avoid the influence of COVID.

¹⁴BLS withholds employment and wage data to protect the privacy of employers and employees. Therefore, state-level records depend on the size of the industry. Missing values are coded as zero in the dataset, regardless of whether the industry or occupation exists in each state. The sensitivity of results to excluding observations equal to zero is tested in Appendix Figure B.6.

¹⁵The ACS is unbalanced in these characteristics as the survey is not a representative sample of workers across industries and occupations. For this reason, I use the average of these values in missing years, states, and industries.

of EITC exposure captures the percent of workers eligible to the program in a given industry or occupation from 1980 to 2021.¹⁶

Economic industries and occupations are categorized into major groups. In the case of industries, industry groups are aggregated as economic sectors. Figure 3.1 underscores relevant patterns in the average EITC exposure in economic sectors. With more than 20% exposure, eligible workers take jobs in the accommodation and food services, and the transportation and warehousing sectors. Drinking establishments, recreational vehicles, and traveler accommodations are highly exposed industry groups in the first sector. In the transportation and warehousing sector, courier and delivery services, local messengers and freight, taxis and limousines are examples of exposed industries. Other industry groups with high exposure are animal slaughtering and processing, childcare services, fruit and vegetable preserving and specialty food manufacturing, greenhouses, nurseries, and floriculture, motion picture and video industries, nursing care facilities, office furniture manufacture, mental health and substance abuse facilities, among others. Consistent with expectations, management is a sector with no exposure to the program. This novel measurement shows a clear picture of the sensitivity of certain industries to social programs that was undocumented to this date.

Similarly, Figure 3.2 illustrates the average exposure among major occupations. The most affected occupations are related to food preparation and serving; farming, fishing, and forestry; building and grounds cleaning and maintenance; personal care and service; transportation and material moving; healthcare support; and education, training, and library. Consistently with industry data, occupational data shows that exposure is high among waiters and waitresses, telemarketers, taxi drivers, service station attendants, roofers, personal and home care aides, food preparation workers, laundry and dry-cleaning workers, maids and housekeeping cleaners, and packers and packagers. Over 25% of workers in these occupations are eligible to the program.

¹⁶I use occupational codes from 1990 to 2021, as SOC codes are not available in 1980. Estimates are robust to using different years of exposure.

Table 3.2: Descriptive Statistics in 2018.

	Mean	SD	Min	Max
Industry Outcomes				
Establishments	547.34	4691.95	0.00	500604.00
Employees	7101.84	20372.65	0.00	656709.00
Weekly Wages 2018	860.80	736.85	0.00	27737.00
Industry/State/Year Covariates				
Female	0.42	0.18	0.00	1.00
Hispanic	0.09	0.10	0.00	1.00
College educated	0.53	0.16	0.00	1.00
High school degree	0.40	0.15	0.00	1.00
No school degree	0.07	0.06	0.00	1.00
White	0.82	0.13	0.00	1.00
Black	0.08	0.08	0.00	1.00
Married	0.59	0.13	0.00	1.00
Number of Children	0.79	0.28	0.00	8.00
Age	43.18	3.74	19.00	65.00
Median Family Income	72393.07	25080.09	0.00	645500.00
Observations	12903			
Occupation Outcomes				
Total Employment	5430.79	17329.18	30.00	558350.00
Mean Hourly Wages	22.28	16.04	0.00	121.13
Median Hourly Wages	20.85	14.81	0.00	99.31
Hourly Wages percentile 10	13.53	9.17	0.00	83.39
Hourly Wages percentile 25	16.60	11.64	0.00	90.58
Hourly Wages percentile 75	26.02	18.59	0.00	99.97
Hourly Wages percentile 90	30.55	22.09	0.00	99.97
Occupation/State/Year Covariates				
Female	0.47	0.28	0.00	1.00
Hispanic	0.08	0.13	0.00	1.00
College educated	0.61	0.25	0.00	1.00
High school degree	0.34	0.23	0.00	1.00
No school degree	0.05	0.09	0.00	1.00
White	0.82	0.19	0.00	1.00
Black	0.08	0.13	0.00	1.00
Married	0.58	0.22	0.00	1.00
Number of Children	0.77	0.45	0.00	8.00
Age	42.86	5.98	19.00	65.00
Median Family Income	80552.35	40950.80	0.00	860000.00
Observations	22134			

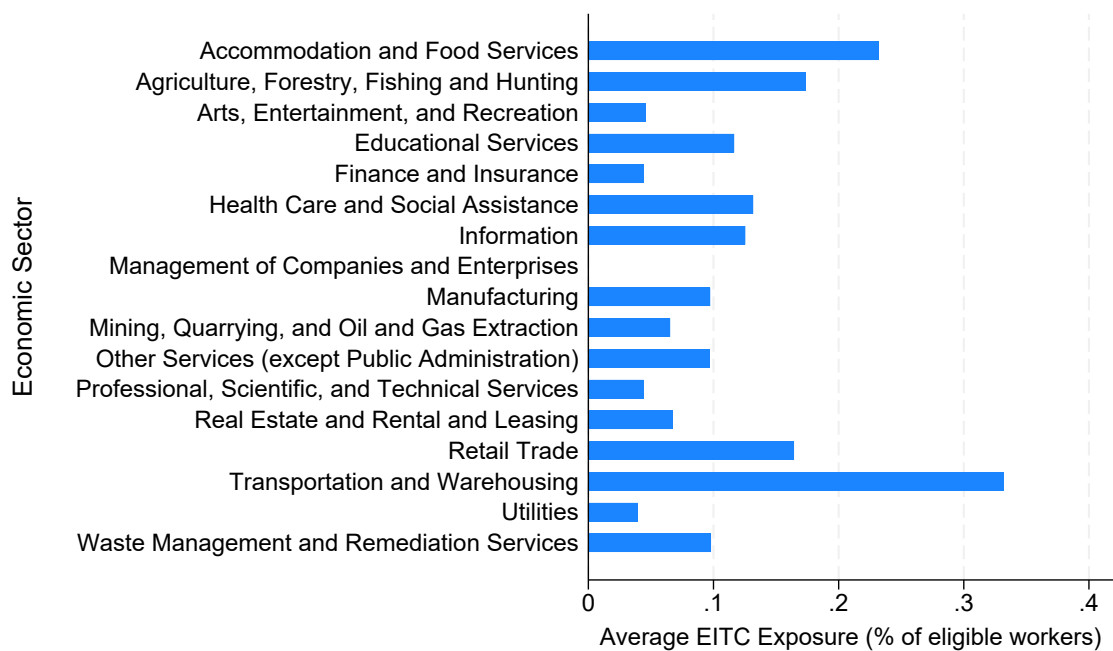


Figure 3.1: EITC exposure across economic sectors.

Notes: NAICS codes are grouped at the economic sector level (2 digit NAICS). EITC exposure is the average percent of eligible workers eligible to the tax credit between 1990 and 2018 in a given industry (4 digit NAICS).

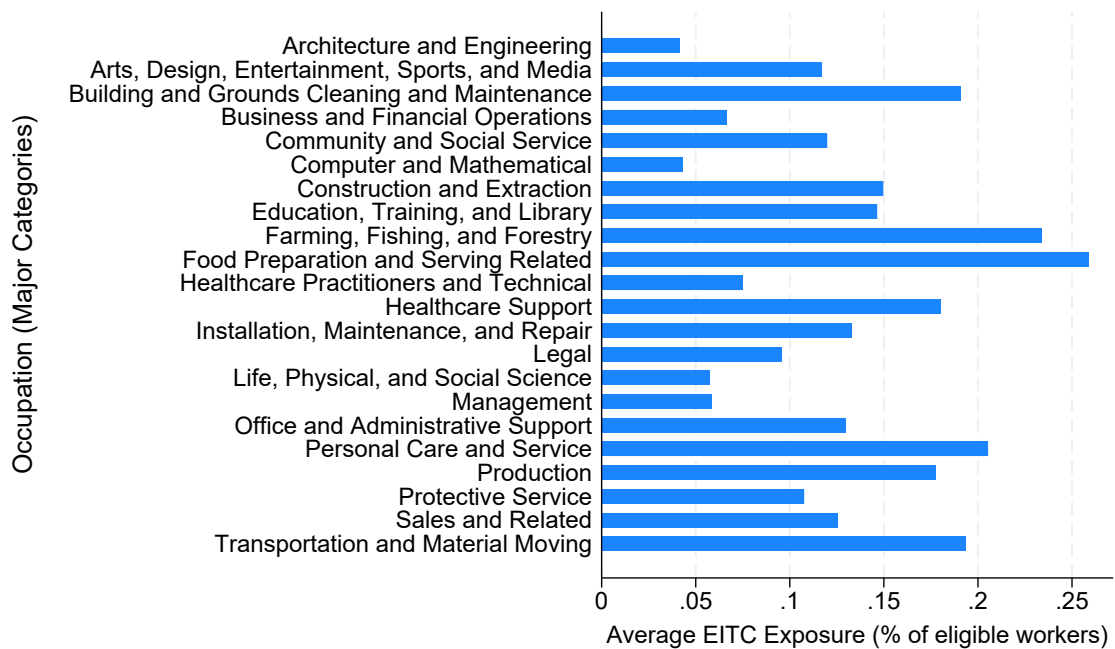


Figure 3.2: EITC exposure across major occupations.

Notes: Economic occupations are grouped as major occupations using the first 2 digits in the SOC code. EITC exposure is the average percent of eligible workers eligible to the tax credit between 1990 and 2018 in a given economic occupation.

3.4 Results

The main findings are displayed in Table 3.3. The table presents the TWFE estimates of the EITC’s impact on industry-level outcomes. Columns 1 to 4 are based on the interaction between the state having an EITC supplement program and the staggered roll-out (Equation 3.1). Columns 5 to 8 exhibit estimates calculated with variation in the state’s tax credit enactment and differences in industries’ exposure to the EITC computed with the 75th percentile of the distribution of workers eligible for the tax credit (Equation 3.2).¹⁷ Columns 3 and 7 show estimates on a restricted sample excluding states that passed only a non-refundable tax credit. Columns 4 and 8 restrict the sample to states without a refundable tax credit.

There are no effects on the number of establishments or employment. However, the findings reveal that the EITC program decreased real weekly wages between \$69 and \$74, or a decrease of 8-8.5% relative to the average. The effects of the program are observable in models exploiting industry exposure (columns 5 to 7), but coefficients are insignificant in the TWFE estimator that relies only in state roll-out. In Appendix Table B.1, I test the sensitivity of results to a logarithmic functional form in outcomes. The semi-elasticities show a more moderate effect, as wages drop between 1.2% and 1.7%. These figures are closer to those in previous research (Zurla, 2024; Leigh, 2010).

The effect of the program appears driven by refundable tax credits. Column 7 shows that enactment of the refundable EITC program decreased wages by \$74 while the nonrefundable program, in column 8, is associated with a weaker and statistically insignificant decrease. The null effects in non-refundable tax credits are consistent with the institutional characteristics of refundability. The refundable program fundamentally acts as a work subsidy program to low-income households, increasing the wage level above reservation wages. Nonrefundable programs are not direct cash transfers and are likely to reach a smaller subset of the population (Brehm and

¹⁷The results in Appendix Figure B.9 are robust to different thresholds in the percentile of the distribution.

Malkova, 2023; Goldin and Michelmore, 2022). The results are therefore consistent with nonrefundable programs having no labor market effects, while the refundable state supplement decreases wages without a positive impact on employment.

The results highlight that market prices respond to the program but not the quantities in the labor market. This seemingly inconsistent result may find an explanation in the features of the labor demand which require further research. First, the labor demand may be inelastic in the specific industries exposed to the program. However, research suggests that the labor demand of low-skilled work is elastic in the long run in low-skill jobs (Lichter et al., 2015; Slaughter, 2001; Rich, 1990; Hamermesh, 1986, 1976). In the short run, these industries may be constrained by capital limitations. Alternatively, the tax credit program may increase willingness to work and, if markets are tight, introduce prospective workers to compete for the same number of job vacancies.¹⁸

Results at the occupational level are listed in Table 3.4. All estimates are computed using the TWFE estimator with program exposure and controlling for all covariates.¹⁹ ²⁰ The first column presents the coefficient of passing an EITC supplement on employment. The second and third columns display the mean and median of the wage distribution in each occupation and state. The remaining columns have the coefficients of models on real wages at different percentiles of the wage distribution.

The negative employment effects in the first column in Table 3.4 are unexpected in their sign and significance. However, employment effects in occupational the data are insignificant in logarithmic form (See Appendix Table B.2) and in a sample restricted

¹⁸Results are robust to weighting by the population. See the Appendix Table B.5.

¹⁹The TWFE estimates without program exposure are displayed in Appendix Table B.3. These coefficients are insignificant except for a positive association at the bottom of the wage distribution.

²⁰The table reports results in a full sample of refundable and nonrefundable tax credits. The restricted sample results are presented in Table B.4, limiting the sample to refundable and non-refundable tax credit states. The restricted samples are consistent with the general finding, as all coefficients on wages are negative and statistically significant.

to refundable or non-refundable tax credits (Appendix Table B.4). The coefficient is also sensitive to varying the threshold in exposure (Appendix Figure B.9).

The overall results highlight a negative effect of the program on hourly wages, seen in Table 3.4. The effect is consistently negative across the distribution of wages. The EITC led to a reduction of \$1.55 in the mean real wages per hour in the distribution, or 8% with respect to the baseline. The effect oscillates and finds its highest relative impact on median wages (8.5%) and in the first quintile of the distribution (7.6%). However, the logarithmic functional form shows that hourly wages only decrease between 1.1 and 2% (see Appendix Table B.2). Furthermore, only the bottom of the wage distribution is affected when using the logarithmic functional form in the outcome.²¹

While the EITC should have no effect on labor markets paying higher wages, there are two possibilities for decreases at the top of the wage distribution. First, wage decreases at the bottom of the distribution may increase the positive skewness in the log-normal distribution of wages, thereby reducing statistical measurements across upper ranges in the distribution. Second, the average 90th percentile wages in the sample are below \$30 per hour and \$42 thousand in annual wages. Therefore, even workers at the top of this distribution may be eligible to the tax credit depending on the tax credit brackets, family characteristics, and hours worked. On a similar note, the literature has argued that the tax credit program targets low-income families instead of low-skill workers, documenting that the program decreases wages for non-targeted groups (Leigh, 2010; Rothstein, 2010; Dickert-Conlin and Holtz-Eakin, 1999). The heterogeneous impact across industries and occupations remains of interest.

In terms of incidence, a back of the envelope calculation implies that firms capture a third of the average tax credit received by families. Nationwide, the average family receives \$2,541 in EITC transfers per year. Given the lower-bound estimate of 1.7%

²¹Unlike weekly wages, occupational data of hourly wages has no zeroes or missing values. Therefore, log-level models can accurately estimate semi-elasticities.

or a \$17.2 decrease in weekly wages, a third of the subsidy is captured by firms. My results are consistent, with computed values in Rothstein (2010) and Zurla (2024).

3.4.1 Industry Heterogeneities

I document industry heterogeneities estimating Equation 3.1 in a subsample of each industry group. The estimator captures heterogeneities across industries using variation in state roll-out, computing 253 estimates. The baseline comparison is industries in states without a supplementary program. To ease interpretability, results are grouped by economic sector and presented as a box plot in Figure ??.²² In the box plot, the median value reveals whether most industries show a positive or negative coefficient. The plot shows a high degree of variability in estimates.

The upper left panel displays the correlation between program roll-out and establishments in most economic sectors, with large decreases in establishments in the accommodation and food industry, management, and retail sectors. With regards to employment, entertainment, healthcare, and educational services report mostly positive coefficients. Most industries report drops in wages in the accommodation and food services sector, transportation, real estate, retail, utilities, and in mining. This suggests that these are the most impacted economic sectors. Arts and entertainment, manufacturing, and finance are sectors with mostly positive correlations. However, these results only provide the distribution of estimates and no confidence intervals.

I recover all estimates and document their association with industry exposure. The results show that heterogeneities in the program's effects across industries are related to the EITC exposure of industries. Figure 3.6 plots the results in all three outcomes. The employment coefficients show a small positive correlation with industry exposure. This may suggest that industries with over 40% eligible workers had larger positive effects on employment albeit the correlation is weak. Lastly, industry exposure has a negative correlation with the program's impact on real wages.

²²The large amount of coefficients is best organized in economic sectors. Plotted results do not include confidence intervals.

Table 3.3: Average impact of the EITC on industry outcomes.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of Establishments								
EITC Enactment	33.888 (56.481)	9.376 (30.310)	14.220 (38.677)	-23.303 (16.318)	-39.625 (28.867)	-54.203 (41.415)	-53.570 (47.901)	-19.972 (34.764)
Employment								
EITC Enactment	-255.308 (202.305)	-380.450** (154.989)	-417.465** (182.047)	-413.261** (202.824)	-960.891 (599.351)	-957.839 (593.516)	-801.550 (588.136)	-848.818 (801.576)
Real Wages (\$2018 per week)								
EITC Enactment	24.484 (16.761)	17.248 (11.275)	24.664** (12.227)	4.432 (17.583)	-69.048*** (18.813)	-58.462*** (17.503)	-74.071*** (18.990)	-50.532* (28.151)
Obs.	387090	387090	355718	309166	387090	387090	355718	309166
R2	0.46	0.47	0.47	0.51	0.47	0.47	0.47	0.52
Covariates	N	Y	Y	Y	N	Y	Y	Y
Sample	All	All	Refundable	Non-Refundable	All	All	Refundable	Non-Refundable
Exposure	N	Y	N	N	Y	Y	Y	Y

Notes: The table reports TWFE estimates of the EITC's impact on the number of establishments, employment, and real weekly wages across industries. The treatment variable captures whether the state passed the EITC. Estimators with exposure use the interaction between the enactment of the tax credit and an indicator for EITC exposure above the 75th percentile of exposure. Estimates based on the enactment of a refundable (or non-refundable) EITC supplement exclude observations from states that ever enacted a non-refundable (refundable) program. All models include state fixed effects, year fixed effects, and industry fixed effects. Models with covariates control for the unemployment rate, state per capita GDP, minimum wage, state marginal tax rate, the percentage of female workers, average race and ethnicity, educational achievement, number of children, median family income, and the age of workers in the industry. Standard errors are clustered at the state level. $p < 0.01$ $p < 0.05$ $p < 0.1$.

Table 3.4: Average impact of the EITC on occupational outcomes.

	(1) Employment	(2) Mean Wages	(3) Median Wages	(4) 10th Pct	(5) 25th Pct	(6) 75th Pct	(7) 90th Pct
EITC Enactment	-326.319*** (83.018)	-1.551*** (0.029)	-1.503*** (0.027)	-0.894*** (0.019)	-1.191*** (0.022)	-1.752*** (0.034)	-1.825*** (0.041)
Obs.	353100	358546	358106	358544	358517	356264	348839
R2	0.50	0.88	0.88	0.85	0.87	0.88	0.87
Covariates	Y	Y	Y	Y	Y	Y	Y
Sample	All	All	All	All	All	All	All
Exposure	Y	Y	Y	Y	Y	Y	Y
Outcome baseline	4797.58	19.37	18.17	11.69	14.41	22.77	27.19

Notes: The table reports TWFE estimates of the EITC's impact on the number of employees and real hourly wages across (average, median, and selected percentiles) occupations. The treatment variable captures whether the state passed the EITC. Estimators with exposure use the interaction between the enactment of the tax credit and an indicator for EITC exposure above the 75th percentile of exposure. All models include state fixed effects, year fixed effects, and occupation fixed effects. Models control for the unemployment rate, state per capita GDP, minimum wage, state marginal tax rate, the percentage of female workers, average race and ethnicity, educational achievement, number of children, median family income, and the age of workers in the industry. Standard errors are clustered at the state level. $p < 0.01$ $p < 0.05$ $p < 0.1$.

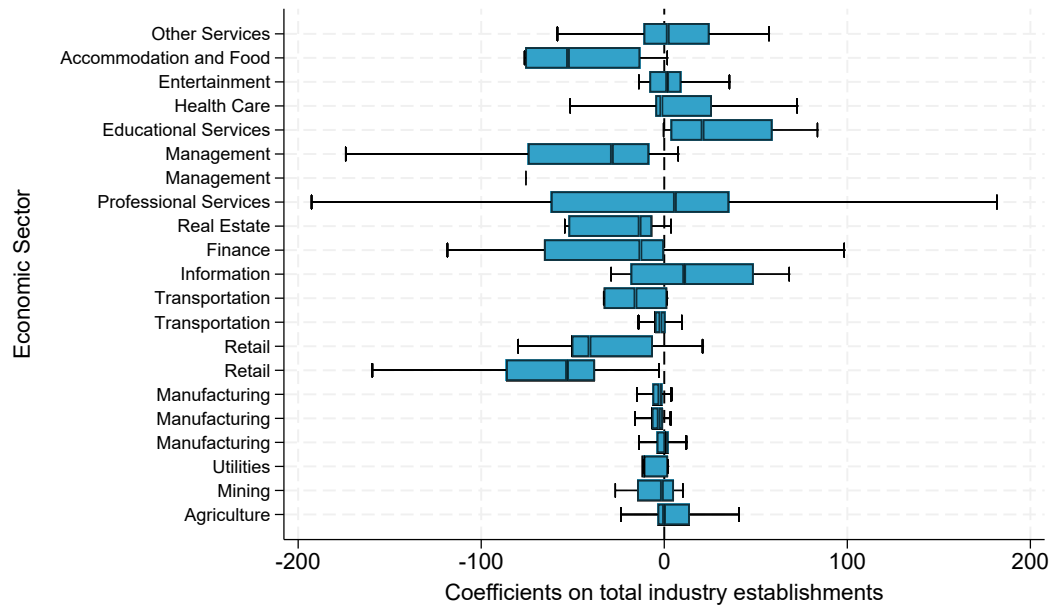


Figure 3.3: Heterogeneous effects of the EITC across economic sectors: Establishments.

Notes: The figure presents box plots for the distribution of estimates of the impact of EITC roll-out across industries, grouped by economic sector. Models control for state and year fixed effects, the unemployment rate, state per capita GDP, minimum wage, state marginal tax rate, the percent of female workers, average race and ethnicity, educational achievement, number of children, median family income, and the age of workers in the industry. Standard errors are clustered at the state level.

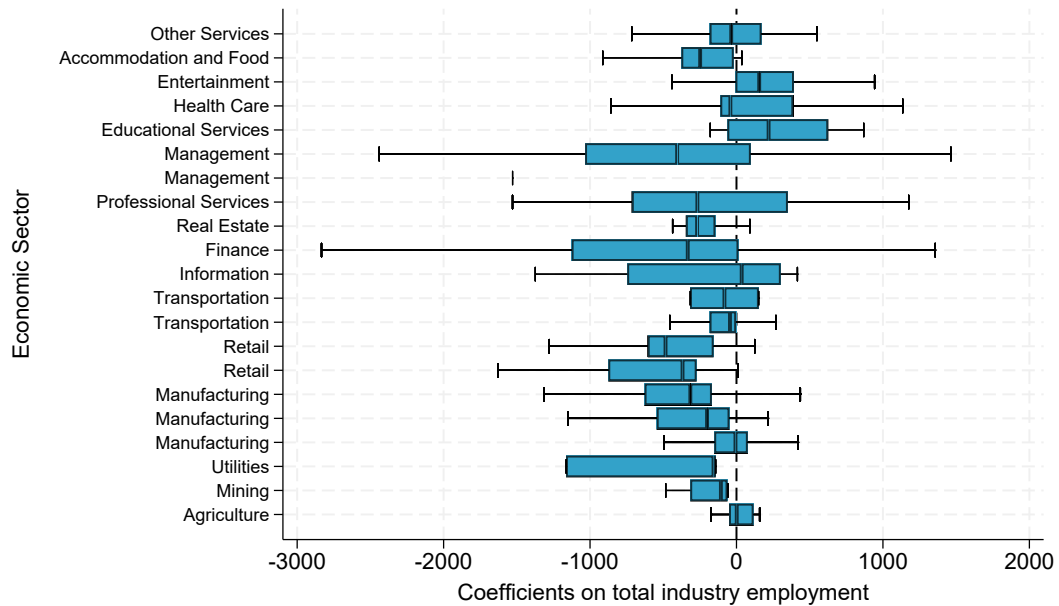


Figure 3.4: Heterogeneous effects of the EITC across economic sectors: Employment.

Notes: The figure presents box plots for the distribution of estimates of the impact of EITC roll-out across industries, grouped by economic sector. Models control for state and year fixed effects, the unemployment rate, state per capita GDP, minimum wage, state marginal tax rate, the percent of female workers, average race and ethnicity, educational achievement, number of children, median family income, and the age of workers in the industry. Standard errors are clustered at the state level.

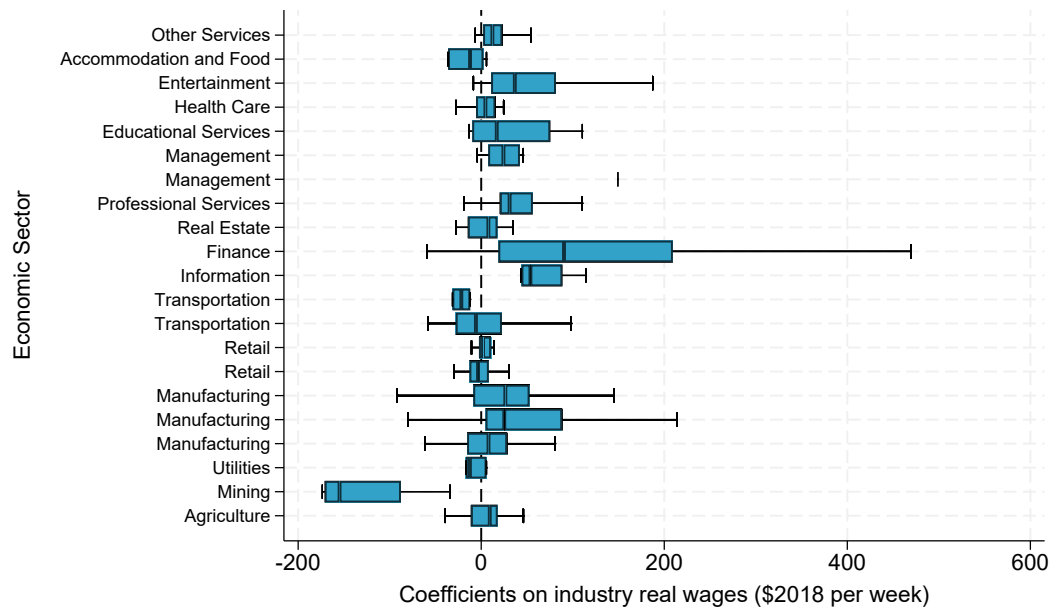


Figure 3.5: Heterogeneous effects of the EITC across economic sectors: Wages.

Notes: The figure presents box plots for the distribution of estimates of the impact of EITC roll-out across industries, grouped by economic sector. Models control for state and year fixed effects, the unemployment rate, state per capita GDP, minimum wage, state marginal tax rate, the percent of female workers, average race and ethnicity, educational achievement, number of children, median family income, and the age of workers in the industry. Standard errors are clustered at the state level.

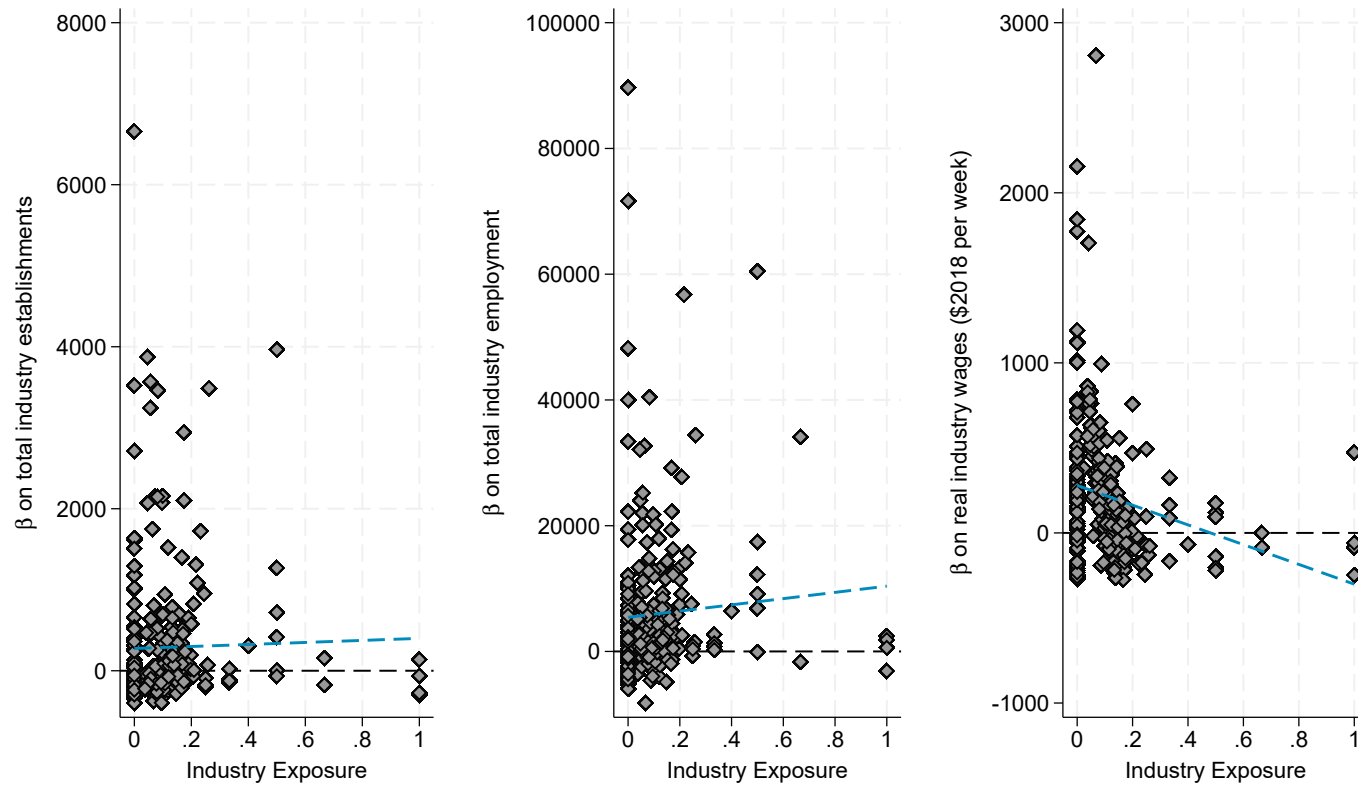


Figure 3.6: Heterogeneous effects of the EITC and industry exposure.

Notes: The figure presents the correlation between industry exposure (x-axis) and estimates of the impact of EITC roll-out across industries (y-axis). The left panel are the coefficients on the number of establishments in an industry and state; the middle panel are the coefficients on employment; and the right panel are the coefficients on real wages. Models control for state and year fixed effects, the unemployment rate, state per capita GDP, minimum wage, state marginal tax rate, the percent of female workers, average race and ethnicity, educational achievement, number of children, median family income, and the age of workers in the industry. Standard errors are clustered at the state level.

3.4.2 Occupational Heterogeneities

I follow a similar approach in the analysis of heterogeneities across occupations. The process yields over four hundred estimates, reported in Appendix Figure ???. The largest negative wage effects are found for jobs in the food preparation major occupation, dining room and cafeteria attendants and bartender helpers, ushers and lobby attendants, counter attendants in cafeteria or coffee shops, and other jobs in the food preparation and serving major occupational categories.

Figure 3.7 illustrates the correlation between employment and wage coefficients with occupation exposure. The left panel includes all occupational coefficients across the horizontal axis of occupational exposure. In general, employment effects have a flat correlation with exposure, although some occupations have large positive correlations between the EITC and employment. The right panel shows a clear negative correlation between exposure and the impact of the program on real wages. Occupations where a high proportion of workers are eligible for the EITC program experience the largest decreases, up to \$5.13 per hour.

One caveat in these heterogeneities is the implication that wages may decrease below minimum wages, highlighting the importance of policy interactions (Neumark and Wascher, 2011, 2001). Theoretically, the minimum wage limits the incidence of negative wage effects, as wages will only reduce up to the price floor (Rothstein and Zipperer, 2020). Empirical evidence suggests that, while the interaction between these policies enhances effectiveness for single mothers, there are negative spillovers for less-skilled men and childless women (Neumark and Wascher, 2011). My empirical evidence is consistent with this literature.

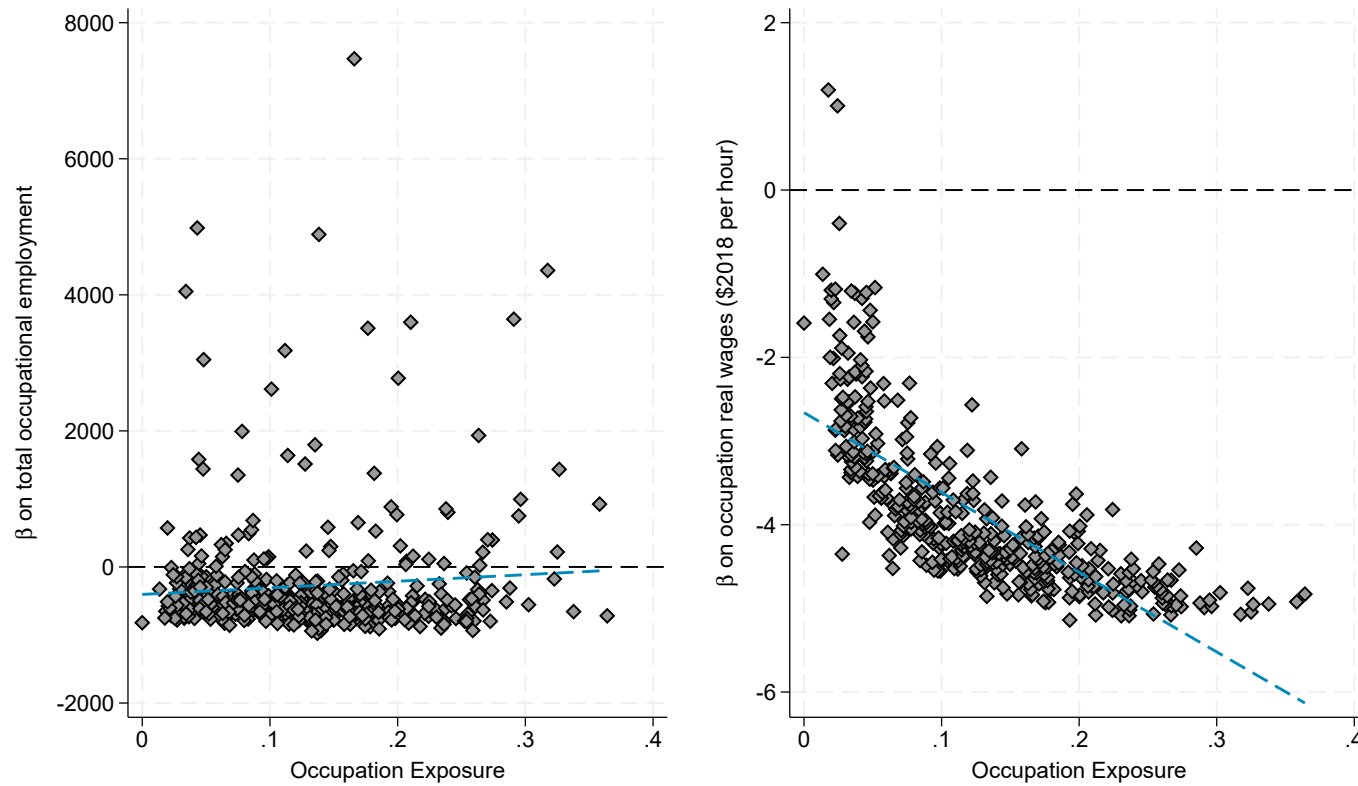


Figure 3.7: Heterogeneous effects of the EITC and occupation exposure.

Notes: The figure presents the correlation between occupation exposure (x-axis) and estimates of the impact of EITC roll-out across occupations (y-axis). The left panel are the coefficients on the number of establishments in an occupation and state; the middle panel are the coefficients on employment; and the right panel are the coefficients on real wages. Models control for state and year fixed effects, the unemployment rate, state per capita GDP, minimum wage, state marginal tax rate, the percent of female workers, average race and ethnicity, educational achievement, number of children, median family income, and the age of workers in the industry. Standard errors are clustered at the state level.

3.5 Robustness

3.5.1 Event Studies

This section presents event studies comparing the TWFE estimator and the [Sun and Abraham \(2021\)](#) robust estimator. These event studies highlight time heterogeneities in the labor market effects of the program. Furthermore, the evidence is indicative of parallel pre-trends, mitigating concerns that difference-in-difference assumptions are violated.²³

Figure [3.8](#) illustrates the event study with the time heterogeneous impact of EITC roll-out on employment. While I do not find any effects on employment using a standard TWFE estimator, the weighted estimator shows a large transitive response in employment. According to the event study, the EITC roll-out program may have been followed by increases in employment of up to 3,000 workers five years after the program was enacted. The effect appears transitory, followed by a re-adjustment in industries. Occupational data displays similar patterns in Appendix Figure [B.4](#). The delayed response in labor markets may be interpreted as learning or knowledge delays, which may be congruent with substantial variability in knowledge about the program ([Wilson, 2022](#); [Chetty et al., 2013](#)).

Evidence on the time heterogeneities in the impact of the program on wages is displayed in Figure [3.9](#). The figure shows a consistent pre-trend in both methodologies, although the weighted estimator has a large significant peak five years before the program is enacted. The TWFE estimates exhibit a statistically insignificant surge in wages immediately after the program is enacted but trending downwards until wages fall below the baseline. These estimates are all negative. The robust estimator shows instead that wages progressively decrease in the long run. This evidence is highly consistent with occupational data, where wages have

²³The event study on establishments is available in Appendix Figure [B.1](#). Event-studies based on occupational data are available in Appendix Figures [B.4-B.5](#).

an initial increase and fall soon after until they are below the control (see Appendix Figure B.5).

3.5.2 Timing in policy enactment

One major threat to identification is the plausible correlation between time-varying state characteristics and the adoption of supplementary tax credit programs. This section analyzes whether a state's introduction of the EITC program is predicted by observable characteristics: the state's total tax revenue, unemployment rate, spending in higher education, the minimum wage, the maximum welfare benefits for a family of three, and the GDP per capita. I find little evidence of correlation with observable policy covariates. Appendix Figure B.7 shows no correlation between state-policy variables and the introduction of a state matching EITC, except in the case of minimum wage policies which are controlled for in the main specification. A larger minimum wage appears to be correlated with passing an EITC supplement.

Similarly, sociodemographic characteristics in states may be correlated with program roll-out. There is no systematic evidence of correlation, as shown in Appendix Figure B.8. Some variables exhibit a positive correlation with state supplementary programs: the percent of college-educated population, the percent of never-married population, and total family income. Most variables are uncorrelated, such as race and ethnicity, sex, the percent of the population with a high school degree, family size, and employment levels.

3.5.3 Variations in the exposure threshold

Industry exposure to the program is defined as an indicator variable based on whether the percentage of eligible workers is above a threshold. Results are based on the 75th percentile of the distribution, which may be considered an ad hoc definition of exposure. However, as shown in Appendix Figure B.9, all estimated coefficients are consistent when using various thresholds of exposure. Particularly, treatment effects

on wages are negative regardless of defining industry exposure above the 40th or above the 95th percentile in the distribution of eligible workers in the industry or occupation. The impact on establishments and employment in industries fluctuates from significant for low levels of exposure to insignificant at high levels of exposure. Employment levels in occupations have stable coefficients.

Another plausible source of bias relates to the decision to define EITC exposure using the percentage of workers in the ACS. The EITC substantially expanded eligibility in 1993 and 1996 during the Welfare Reform, likely affecting the composition of workers in industries and occupations. This may imply unstable treatment and control groups as reforms affect eligibility and labor markets. I test the robustness of my findings by re-defining industry exposure with data preceding the major EITC expansions in 1993 and 1996. Treatment effects are re-estimated using the percentage of workers eligible to the program in industries and occupations across individual years. Coefficients are remarkably stable independently of the household survey used to define exposure, as seen in the Appendix Figure [B.10](#). Results for industries are consistent with the main estimates when restricting exposure to pre-reform years in 1980 and 1990, before industries were affected by any EITC expansion.²⁴

3.6 Discussion

This paper presents new estimates of the EITC's impact on firm dynamics and labor market outcomes, based on establishment survey data. The empirical investigation leverages differences in state roll-out of supplementary tax credit programs and differences in industry and occupational exposure to the program. The measurement of exposure is constructed with the percent of workers eligible for the program according to the American Community Survey. The paper contributes to a large body of work studying the labor market effects of the program on employment ([Kleven](#),

²⁴Occupational data is unavailable before 2000, as the SOC codes are not available in 1980 and 1990. Furthermore, because the ACS is not a representative survey of industries and occupations, observations are lost whenever surveyed individuals do not work in a particular industry.

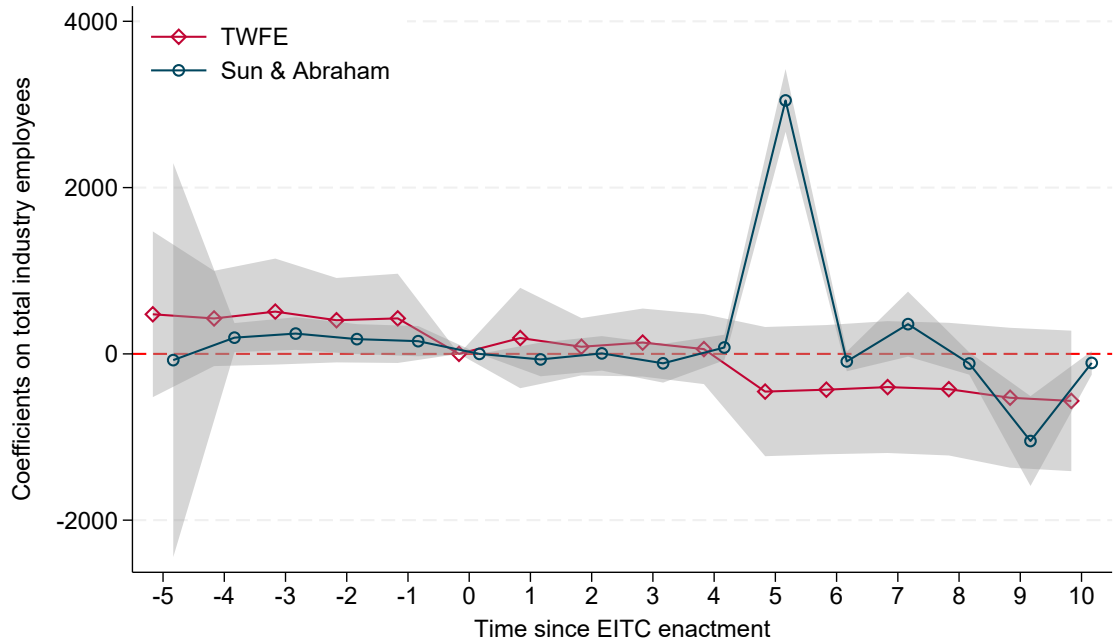


Figure 3.8: Event study of EITC roll-out and employment across industries.

Notes: The figure presents standard event studies using the Two-Way Fixed Effects (TWFE) estimator and the [Sun and Abraham \(2021\)](#) estimator. The outcome is the number of employed workers in each industry. The treatment variable captures the interaction between the enactment of the EITC and an indicator for EITC exposure above the 75th percentile of exposure. Models control for occupation fixed effects, state fixed effects, year fixed effects, the unemployment rate, state per capita GDP, minimum wage, state marginal tax rate, the percent of female workers, average race and ethnicity, educational achievement, number of children, median family income, and the age of workers in the industry. The 95 percent confidence intervals are estimated with cluster-robust standard errors at the state-level.

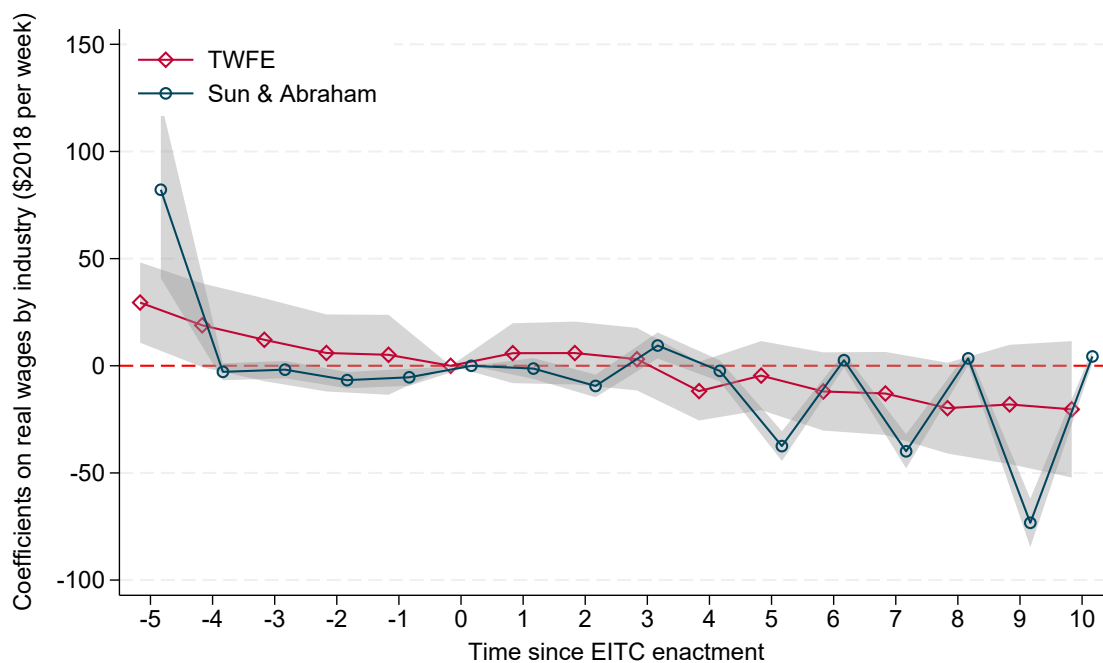


Figure 3.9: Event study of EITC roll-out and wages across industries.

Notes: The figure presents standard event studies using the Two-Way Fixed Effects (TWFE) estimator and the [Sun and Abraham \(2021\)](#) estimator. The outcome is the real wage paid in each industry. The treatment variable captures the interaction between the enactment of the EITC and an indicator for EITC exposure above the 75th percentile of exposure. Models control for occupation fixed effects, state fixed effects, year fixed effects, the unemployment rate, state per capita GDP, minimum wage, state marginal tax rate, the percent of female workers, average race and ethnicity, educational achievement, number of children, median family income, and the age of workers in the industry. The 95 percent confidence intervals are estimated with cluster-robust standard errors at the state-level.

2023; Nichols and Rothstein, 2015; Eissa and Hoynes, 2004; Eissa and Liebman, 1996) and wages (Zurla, 2024; Neumark and Wascher, 2011; Leigh, 2010; Rothstein, 2010). Moreover, the paper provides the first industry and occupational estimates of the program’s effect on labor markets.

Using establishment data and a triple difference approach, this study documents that wages decrease by over 8 percent, all else equal, after the enactment of state-level EITC supplement programs. However, employment and the number of establishments reported by firms remain flat. My findings are consistent across industry and occupational data from two different establishment surveys. Heterogeneities suggest that these effects are highly concentrated in exposed industries and occupations related to the service economy, such as accommodations and food services. The findings highlight the distributional incidence of the program, consistent with estimates of negative spillovers to non-targeted populations (Neumark and Wascher, 2011; Leigh, 2010; Rothstein, 2010). Furthermore, there is evidence that parallel trends are robust in a traditional two-way fixed effects estimator and using Sun and Abraham (2021)’s estimator. The event studies also document that the impact of the EITC may be lagged three years in the future and transitory, as labor markets react to increased willingness to work. This is similar to evidence in Zurla (2024), and may be related to changes in the growth rate of wages.

This study has several limitations. First, establishment surveys have well-documented deficiencies in capturing double employment or self-employment (U.S. Bureau of Labor Statistics, 2024). Self-employment may be a large driver of differences across households and establishment surveys, and it is likely that the EITC’s labor market effects concentrate on the self-employed (Lim and Micheltore, 2018; Kuka, 2014; LaLumia, 2009). Establishment surveys may double-count workers with several jobs. On the other hand, establishment surveys are constructed with administrative payroll data. Despite their limitations, establishment surveys have advantages in addressing identification challenges given the endogenous responses of households. Second, while lacking in employment counts and unable to capture

willingness to work, establishment surveys accurately record hourly wages in labor markets and the wage distribution.

The research design is also limited by the available measurements of exposure. Policy changes in the tax credit program have substantially expanded program eligibility. Changes in eligibility may cause unobserved changes in trends in specific industries and occupations. This research presents new evidence on the industries and occupations impacted by the program, but future research is required on the relationship between work incentives from social programs and occupational choice and workers' self-selection into industries. Researchers have found strong effects of the EITC on self-employment, but there remains a gap in understanding job choices ([Lim and Micheltore, 2018](#); [Kuka, 2014](#); [LaLumia, 2009](#)). For example, single mothers may choose to work in occupations and industries with flexibility in their work schedules. An additional limitation is that the roll-out design does not allow computing dollar-to-dollar estimates of the tax credit impact, as we are unable to observe the transfer received by workers.

There are additional areas for future research on the labor market decisions of mothers and work subsidies. Employment-based conditions imply distortions in the time allocation of mothers ([Bastian and Lochner, 2022](#)). The maternal labor choice weighs complex opportunity costs in the household production function and childcare. There has been substantial interest among economists in self-employment, as early research found bunching in the distribution of earnings among this subpopulation ([Lim and Micheltore, 2018](#); [Kuka, 2014](#); [LaLumia, 2009](#); [Saez, 2010](#)). The heterogeneous effects of the EITC on industries and occupations hint at a concentrated impact on service jobs. Future research should aim at understanding the impact of the EITC on work schedules and occupational choice, which the literature has yet to document and may explain stagnant employment after these programs are enacted.

Finally, these results may help policymakers understand the interactions between workfare programs and labor markets in affected occupations and industries. This

paper further evidences the difference between economic incidence and statutory incidence, documenting that tax credits aimed at increasing the workforce may decrease wages in affected industries and occupations. Employment conditions to welfare have vast popularity in the design of public assistance to the poor, but work incentives distort labor markets in affected sectors. Consequently, these policies may push individuals to work more for less pay, particularly if minimum wage policies are not binding.

References

- Agan, A. Y. and Makowsky, M. D. (2023). The minimum wage, eitc, and criminal recidivism. *Journal of Human Resources*, 58(5):1712–1751.
- Andrade, F. C. D., Kramer, K. Z., Greenlee, A., Williams, A. N., and Mendenhall, R. (2019). Impact of the chicago earned income tax periodic payment intervention on food security. *Preventive Medicine Reports*, 16:100993.
- Azoulay, P., Jones, B. F., Kim, J. D., and Miranda, J. (2022). Immigration and entrepreneurship in the united states. *American Economic Review: Insights*, 4(1):71–88.
- Bastian, J. and Lochner, L. (2020). The eitc and maternal time use: more time working and less time with kids? Technical report, National Bureau of Economic Research.
- Bastian, J. and Lochner, L. (2022). 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.
- Bastian, J. and Michelmore, K. (2018). The long-term impact of the earned income tax credit on children’s education and employment outcomes. *Journal of Labor Economics*, 36(4):1127–1163.
- Braga, B., Blavin, F., and Gangopadhyaya, A. (2020). The long-term effects of childhood exposure to the earned income tax credit on health outcomes. *Journal of Public Economics*, 190:104249.

- Brehm, M. E. and Malkova, O. (2023). The Child Tax Credit over Time by Family Type: Benefit Eligibility and Poverty. *National Tax Journal*, 76(3):000–000.
- Card, D., Cardoso, A. R., Heining, J., and Kline, P. (2018). Firms and labor market inequality: Evidence and some theory. *Journal of Labor Economics*, 36(S1):S13–S70.
- Card, D., Heining, J., and Kline, P. (2013). Workplace heterogeneity and the rise of west german wage inequality. *The Quarterly journal of economics*, 128(3):967–1015.
- Chetty, R., Friedman, J. N., and Rockoff, J. (2011). New evidence on the long-term impacts of tax credits. In *Proceedings. Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association*, volume 104, pages 116–124. JSTOR.
- Chetty, R., Friedman, J. N., and Saez, E. (2013). Using differences in knowledge across neighborhoods to uncover the impacts of the eitc on earnings. *American Economic Review*, 103(7):2683–2721.
- Deming, K. (2022). The long-term impact of the earned income tax credit on entrepreneurship. In *2022 APPAM Fall Research Conference*. APPAM.
- Dickert-Conlin, S. and Holtz-Eakin, D. (1999). Employee-based versus employer-based subsidies to low-wage workers: A public finance perspective. Center for Policy Research, 395.
- East, C. N. (2020). The effect of food stamps on children’s health: Evidence from immigrants’ changing eligibility. *Journal of Human Resources*, 55(2):387–427.
- Eissa, N. and Hoynes, H. W. (2004). Taxes and the labor market participation of married couples: the earned income tax credit. *Journal of Public Economics*, 88(9-10):1931–1958.

- Eissa, N., Kleven, H. J., and Kreiner, C. T. (2006). Welfare effects of tax reform and labor supply at the intensive and extensive margins. *Tax policy and labor market performance*, pages 147–86.
- Eissa, N. and Liebman, J. B. (1996). Labor supply response to the earned income tax credit. *The quarterly journal of economics*, 111(2):605–637.
- Evans, W. N. and Garthwaite, C. L. (2014). Giving mom a break: The impact of higher eitc payments on maternal health. *American Economic Journal: Economic Policy*, 6(2):258–90.
- Feenberg, D. and Coutts, E. (1993). An introduction to the TAXSIM model. *Journal of Policy Analysis and Management*, 12(1):189–194.
- GAO (2007). Advance Earned Income Tax Credit: Low Use and Small Dollars Paid Impede IRS’s Efforts to Reduce High Noncompliance. U.S. Government Accountability Office <https://www.gao.gov/products/gao-07-1110>. [Online; accessed 2023-12-14].
- Goldin, J. and Michelmore, K. (2022). Who benefits from the child tax credit? *National Tax Journal*, 75(1):123–147.
- Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277.
- Hamermesh, D. S. (1976). Econometric studies of labor demand and their application to policy analysis. *Journal of Human Resources*, pages 507–525.
- Hamermesh, D. S. (1986). The demand for labor in the long run. *Handbook of labor economics*, 1:429–471.
- Hoynes, H., Miller, D., and Simon, D. (2015). Income, the earned income tax credit, and infant health. *American Economic Journal: Economic Policy*, 7(1):172–211.

- Hoynes, H. W. and Patel, A. J. (2018). 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.
- Isaac, E. (2020). Marriage, divorce, and social safety net policy. *Southern Economic Journal*, 86(4):1576–1612.
- ITEP (2019). When Did Your State Enact its EITC? <https://itep.org/when-did-your-state-enact-its-eitc/>.
- Jones, M. R. and Ziliak, J. P. (2022). The antipoverty impact of the eitc: New estimates from survey and administrative tax records. *National Tax Journal*, 75(3):451–479.
- Kleven, H. (2023). The EITC and the Extensive Margin: A Reappraisal. Technical report, National Bureau of Economic Research.
- Kuka, E. (2014). Eitc and the self-employed: Real or reporting effects? *Public Finance Review*, 42(6):691–719.
- Lake, J. and Millimet, D. L. (2016). Good jobs, bad jobs: What’s trade got to do with it?
- LaLumia, S. (2009). The earned income tax credit and reported self-employment income. *National Tax Journal*, 62(2):191–217.
- Leigh, A. (2010). Who benefits from the earned income tax credit? incidence among recipients, coworkers and firms. *The BE Journal of Economic Analysis & Policy*, 10(1).
- Lichter, A., Peichl, A., and Siegloch, S. (2015). The own-wage elasticity of labor demand: A meta-regression analysis. *European Economic Review*, 80:94–119.
- Lim, K. and Michelmore, K. (2018). The eitc and self-employment among married mothers. *Labour Economics*, 55:98–115.

- Manoli, D. and Turner, N. (2018). Cash-on-hand and college enrollment: Evidence from population tax data and the earned income tax credit. *American Economic Journal: Economic Policy*, 10(2):242–71.
- Markowitz, S., Komro, K. A., Livingston, M. D., Lenhart, O., and Wagenaar, A. C. (2017). Effects of state-level earned income tax credit laws in the us on maternal health behaviors and infant health outcomes. *Social Science & Medicine*, 194:67–75.
- Meyer, B. D. (2002). Labor Supply at the Extensive and Intensive Margins: The EITC, Welfare, and Hours Worked . *American Economic Review*, 92(2):373–379.
- Meyer, B. D. and Rosenbaum, D. T. (2001). Welfare, the earned income tax credit, and the labor supply of single mothers. *The quarterly journal of economics*, 116(3):1063–1114.
- Micheltmore, K. (2013). The effect of income on educational attainment: Evidence from state earned income tax credit expansions. *Available at SSRN 2356444*.
- Micheltmore, K. and Lopoo, L. M. (2021a). The effect of eitc exposure in childhood on marriage and early childbearing. *Demography*, 58(6):2365–2394.
- Micheltmore, K. and Lopoo, L. M. (2021b). Exposure to the earned income tax credit in early childhood and family wealth. *RSF: The Russell Sage Foundation Journal of the Social Sciences*, 7(3):196–215.
- NBER (2023). State EITC provisions 1977-2018 . <https://taxsim.nber.org/state-eitc.html>.
- Neumark, D. and Shirley, P. (2017). The long-run effects of the earned income tax credit on women’s earnings. Technical report, National Bureau of Economic Research.

- Neumark, D. and Shirley, P. (2020). The long-run effects of the earned income tax credit on women’s labor market outcomes. *Labour Economics*, 66:101878.
- Neumark, D. and Wascher, W. (2001). Using the eitc to help poor families: New evidence and a comparison with the minimum wage. *National Tax Journal*, 54(2):281–317.
- Neumark, D. and Wascher, W. (2011). Does a higher minimum wage enhance the effectiveness of the earned income tax credit? *ILR Review*, 64(4):712–746.
- Nichols, A. and Rothstein, J. (2015). The Earned Income Tax Credit (EITC). Technical report, National Bureau of Economic Research.
- Rich, D. P. (1990). On the elasticity of labor demand. *Quarterly Review of Economics and Business*, 30(3):31–42.
- Rothstein, J. (2010). Is the EITC as Good as an NIT? Conditional Cash Transfers and Tax Incidence. *American economic Journal: economic policy*, 2(1):177–208.
- Rothstein, J. and Zipperer, B. (2020). The eitc and minimum wage work together to reduce poverty and raise incomes.
- Saez, E. (2010). Do taxpayers bunch at kink points? *American economic Journal: economic policy*, 2(3):180–212.
- Schaller, Z. and DeCelles, P. (2021). Weighted crosswalks for naics and sic industry codes. *Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]*, pages 44–76.
- Schanzenbach, D. W. and Strain, M. R. (2021). Employment effects of the Earned Income Tax Credit: Taking the long view. *Tax Policy and the Economy*, 35(1):87–129.
- Slaughter, M. J. (2001). International trade and labor–demand elasticities. *Journal of international Economics*, 54(1):27–56.

- Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of econometrics*, 225(2):175–199.
- TCWF (2023). Tax Credits for Workers and Their Families: State Tax Credits –State Resources. <https://www.taxcreditsforworkersandfamilies.org/#>.
- TPC (2023). Tax Policy Center – State EITC as Percentage of the Federal EITC. <https://www.taxpolicycenter.org/statistics/state-eitc-percentage-federal-eitc>.
- U.S. Bureau of Labor Statistics (2024). Employment situation technical note. <https://www.bls.gov/news.release/empsit.tn.htm>. Accessed: 2024-02-01.
- Waxman, S. and Hinh, I. (2023). States Can Enact or Expand Child Tax Credits and Earned Income Tax Credits to Build Equitable, Inclusive Communities and Economies. <https://www.cbpp.org/research/state-budget-and-tax/states-can-enact-or-expand-child-tax-credits-and-earned-income-tax>.
- Wilson, R. (2022). The impact of social networks on eitc claiming behavior. *Review of Economics and Statistics*, 104(5):929–945.
- Zurla, V. (2024). Firm Responses to Earned Income Tax Credits: Evidence from Italy. Unpublished Manuscript.

CHAPTER 4

Not Just About Optics: Adverse Selection and Moral Hazard in Police Militarization and Citizens Deaths

The role of police in our society, including ‘police militarization,’ continues to be a controversial topic among scholars and policy makers. Police militarization is complex, involving culture, personnel, operating procedures, and also tactical equipment. While law enforcement agencies can and do acquire their own tactical equipment, a non-trivial share of tactical equipment acquired by local law enforcement agencies comes from the U.S. Government. Since 1997, the U.S. Department of Defense has supplied state and local law enforcement agencies (LEAs) with approximately \$7.6 billion dollars of military equipment through the 1033 program. The economic implications of these transfers are *ex ante* unclear. On the one hand, these military items can be viewed as inputs to public safety, purchased with tax dollars. If these items are used appropriately by police, there may be some efficiency gains compared to just discarding them. If these items are used injudiciously against the public, however, there is clear potential for this program to do more harm than good.

Evidence on the relationship between 1033 participation and police violence against civilians has been mixed. Research has extensively analyzed the program’s effectiveness on reducing violent crime (Masera, 2021a; Bove and Gavrilova, 2017; Steidley and Ramey, 2019; Ramey and Steidley, 2018; Mummolo, 2018), and questioned whether it increases violence both by and against police officers (Masera, 2021a; Harris et al., 2017). Research outside economics showed associations between military hardware and an increase in civilian deaths by police, suggesting that the 1033 program transfers altered the material and operational capabilities of police forces, adversely affecting their cultural and organizational norms (Lawson Jr, 2019; Delehanty et al., 2017). However, the first studies with research designs intended to identify causal relationships found no link between the acquisition of military equipment and civilian fatalities (Harris et al., 2017; Bove and Gavrilova, 2017). Subsequent work in Masera (2021a) has found that tactical gear has led to increased civilian deaths at the hands of police. The conflicting results have been attributed to many factors, including research design and choice of instrument (Gunderson et al.,

2021), unit of observation (agency vs county), and the construction of measures of units received through the 1033 program (Koslicki, 2022). However, the single most important difference between the studies that do and do not find linkages between receipts of tactical hardware and civilian killings by police is the data source for the outcome variable. Early work with causal research designs, including Harris et al. (2017) and Bove and Gavrilova (2017), used homicide statistics *self-reported* by police departments to the FBI’s Uniform Crime Reporting (UCR) data. Work that has found relationships between tactical receipts and killings Masera (2021a) based their analysis on data collected by third-party, the Fatal Encounters dataset.

In this paper, we compare these data sources to reconcile mixed results from causal studies in the economics literature, and subsequently examine an understudied margin of police behavior: killing civilians and not disclosing these killings to the federal government. We use instrumental variables methods that are well established in this literature, but match self-reported statistics of felons killed by police in the UCR dataset with third-party documented killings in the Fatal Encounters dataset at the agency-year level. We show that even with higher quality data, alternative instruments, and a finer geographic unit of observation, there is no relationship between acquisitions of tactical hardware and civilian deaths by police. These results are consistent with Harris et al. (2017) and Bove and Gavrilova (2017). When using the same research design, but relying on Fatal Encounters data for the data on civilian deaths, we find evidence consistent with Masera (2021a,b) that acquiring tactical hardware is causally linked to increased citizen deaths. We then use the same research design to show that acquisition of military equipment is causally linked to killing civilians, but *not* reporting them to the FBI. The probability of under-reporting increases for departments that acquire a larger stock of 1033 items, and the effect is larger for departments acquiring tactical gear. Departments that acquire weapons and ammunition are 9 percentage points more likely to not report homicides by police officers. The effect is larger in specific categories of tactical gear such as optics (24 percentage points), military rifles (26 percentage points), and ammunition

(15 percentage points). Non-tactical gear has smaller effects below 3 percentage points. Among departments that have killed a civilian, there is a lower likelihood of responding to UCR, suggesting that departments do not necessarily misrepresent the tally but stop responding to UCR. We then show that there are stark differences in both the descriptive statistics and regression results along racial lines. The majority of police killings of white civilians are captured in the self-reported UCR data, while very few of the killings of people of color documented in Fatal Encounters are reported in UCR.

We show that our results are robust to several specification changes and alternative assumptions. The literature has implemented several variations of instruments in the first stage explaining the acquisition of military gear, all of which revolve around cross-sectional proximity to available equipment interacted with time series variation on the availability of said equipment. We show that our results are stable across several alternative specifications in the first stage, using alternative sources of time variation such as boots on the ground in Afghanistan and Iraq (Masera, 2021a,b), real military expenditure (Bove and Gavrilova, 2017), and total items distributed in the nation (Bruce et al., 2019; Harris et al., 2017).

Second, event studies are consistent with our findings and suggest that tactical gear increased under-reporting in the long-term. These event studies have robust pre-trends, although the methodology is not robust to simultaneity bias and expectations by law enforcement agencies. The adverse selection hypothesis is challenging to test, as police culture is not observed and there is no third-party homicide data prior to the start of the federal program. We show that transition matrices are consistent when using civilian assets seized by LEAs before the program, which may be interpreted as an imperfect measurement of negative pre-program behavior.¹

¹The evidence is only suggestive. DOJ asset forfeiture laws had the clear motivation to increase anti-drug policing (Baicker and Jacobson, 2007). Given that the 1033 Program favored law enforcement agencies in high intensity drug trafficking area, it is plausible that the higher levels of asset forfeiture capture LEAs facing higher criminal activity. However, asset seizure remains a controversial policy and research has found that police departments tend to change their arrest patterns under these laws (Baicker and Jacobson, 2007).

Furthermore, parametric results corroborate these hypothesis, as lagged under-reports are a predictive variable in the acquisition of vehicles and optics through the 1033 program.

These results raise questions about both the interpretation of the “causal” relationship between acquiring military equipment and civilian deaths, but also the cultural relationship in law enforcement between militarization and transparency. Regarding the former, the clearest interpretation of a causal relationship between acquiring military hardware and civilian deaths is that of tactical equipment as a production to civilian deaths. However, when the same research design indicates a causal relationship between killing *and* not reporting the deaths of civilians, the implication is that the acquisition of tactical equipment is indicative of shifts towards more aggressive policing in general. Even more troubling, the relationship between militarization and *non-disclosed* killings of civilians indicates that increases in police aggression are accompanied by decreased transparency.

Having shown that receipts of tactical hardware predicts an increased likelihood of a non-disclosed killing of a civilian, we next investigate whether there is concrete evidence of adverse selection, or whether these results are most consistent with moral hazard. To unpack the ‘chicken and egg’ dynamics of militarization and aggressive, non-transparent policing, we focus on police activities prior to 2009, when the number of items available through 1033 increased by an order of magnitude due to the start of the withdrawal from Iraq and Afghanistan.² We examine two distinct sources of variation for evidence that aggressive policing practices precede receipts of tactical hardware. First, we examine whether departments with a prior history of killing civilians and not disclosing those events in UCR are more likely to heavily militarize from 2009-2014. Second, we incorporate data on Civil Asset Forfeitures from 1990-1997 from [Baicker and Jacobson \(2007\)](#).³ Civil Asset Forfeiture is a controversial practice by which law enforcement agencies can seize the assets of anyone *suspected*

²In addition, the 1033 Program rarely denied item requests before 2009 ([Masera, 2021a,b](#)). Therefore, the underlying characteristics of departments may be different before and after 2009.

³We are grateful to the authors for helping us access their data from the NBER archive.

of drug trafficking activity. A conviction is not necessary to seize an individual's assets, nor is the return of those assets in any way assured in the absence of a conviction. Although CAF was originally intended as a way to reduce the effectiveness of drug-enterprises, its widespread abuse has led to characterizations of profiteering and "bandit behavior" (Baicker and Jacobson, 2007). We therefore treat CAF as a measure of aggressive and non-transparent policing. Results indicate that law enforcement agencies with histories of undisclosed civilian killings and/or high levels of civil asset forfeiture are more likely to heavily militarize when larger numbers of tactical items become available in the withdrawal from Iraq. We interpret these results as evidence of adverse selection.

These results have multiple implications for policy going forward. Generally, this study adds evidence to support the calls for greater transparency on the operations of police departments and sheriff's offices. Although the disclosure of killing civilians is voluntary in the UCR surveys, the fact that in recent years, police only choose to disclose roughly a quarter of civilian deaths is alarming. As our results link increased militarization to increased civilian deaths *and* increased probability of non-disclosure, the overall situation will only get worse in the absence of some check on the power of local law enforcement agencies.

More specifically to the 1033 program, these results indicate there has been failure to screen –and in many cases– failure to monitor the activities of recipients of lethal tactical items by the federal government. While several jurisdictions have had their 1033 privileges suspended or revoked, those suspensions have come from mishandling inventory rather than using the items against the populace. However, the federal government cannot be blamed outright for negligence, as Department of Justice has extremely limited oversight authority when it comes to state and local police. Nevertheless, our results indicate that repeated requests for tactical items, or requests for a large number of tactical items, should be grounds for an examination of *why* those items are deemed necessary. The linkage between receipts of tactical items and

undisclosed civilian deaths indicates that weapons and optics are not just tools of the trade, but indicative of how the “mission” is carried out.

4.1 Institutional Background

4.1.1 Police Killings of Citizens

High-profile homicides by police officers have spurred intense policy debates and social movements, underlining the lack of transparency in use of force by law enforcement. To this date, the FBI’s Uniform Crime Reporting (UCR) is the most comprehensive disclosure by law enforcement agencies of cases resulting in civilians killed by officers. However, UCR is a voluntary survey administered to police departments. Although response rates are quite high (usually over 80% ([Kaplan, 2021](#))), there are reasons to suspect the usual problems with survey data (including social desirability bias) apply. Federal authorities acknowledge these limitations and the Department of Justice (DoJ) has estimated that only half of the killings by police are reported, leading the FBI to undertake the surveying of new comprehensive data ([Banks et al., 2015](#)).⁴ In the absence of unbiased and comprehensive administrative data on civilian deaths, civilian organizations and non-for-profits have taken action to record and track police violence against civilians. There are at least three third-party sources of homicides by police: Mapping Police Violence, Fatal Force, and Fatal Encounters. The differences in statistics immediately stand out in the data. See, for example, [Figure 4.1](#) which compares self-reported data in the UCR with these three alternative sources, documenting the large and increasing gap in reported homicides.

⁴Since 2015, the FBI works with state and local law enforcement in developing the National Use of Force Data Collection (UOF). The dataset offers aggregated statistics at the national and state-level of incidents that resulted in the death or serious bodily injury of a person, or the discharge of a firearm at or in the direction of a person. UOF, similarly to UCR, relies on voluntary self-reports by law enforcement and recently achieved 63 percent participation rate among law enforcement agencies. The FBI releases aggregated frequencies of responses to questions and lists all law enforcement agencies participating. There is no public access to statistics of deaths or injuries neither at the national or agency level, but only the percent of agencies reporting these events. See <https://www.fbi.gov/how-we-can-help-you/more-fbi-services-and-information/ucr/use-of-force>

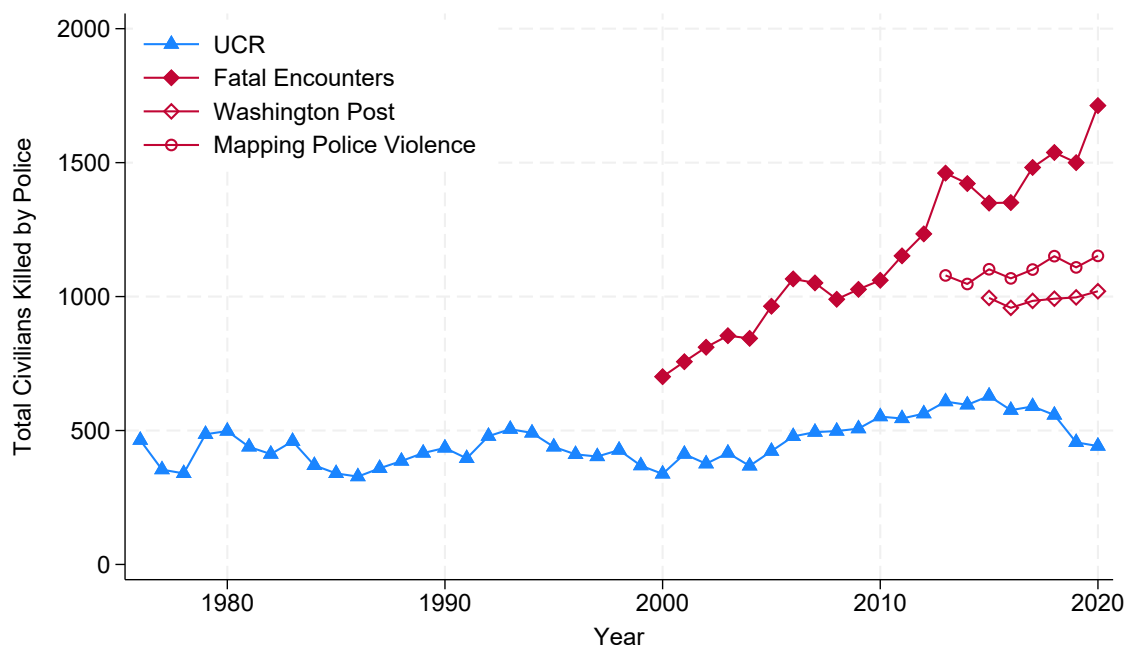


Figure 4.1: Civilians killed by police recorded in UCR and Fatal Encounters datasets.

Notes: The figure plots the total number of civilians killed by police officers according to four different datasets. In Fatal Encounters data, we drop all instances involving more than one agency, and drop all observations where police officers are off-duty.

It is hard to overstate the difference between the self-reported numbers of police killings and those reported by third parties. In 2020, the Federal Bureau of Investigation (FBI) reported that law enforcement officers killed 428 ‘felons.’ The actual number of civilians killed in 2020 was approximately 1,750, according to fatal encounters. The discrepancy has its origins in the non-mandatory reporting of homicides to federal authorities. This allows departments to strategically abstain from reporting when homicides are committed.⁵ Figure 4.1 illustrates how the gap between killings documented in Fatal Encounters and killings reported in UCR has grown since 2000. The underlying trends indicate that as recently as the mid-to-late 1990s, UCR may have been reasonably accurate.

4.1.2 Institutional Details on the 1033 Program

The National Defense Authorization Act of 1997 authorized the Secretary of Defense to transfer excess Department of Defense (DoD) equipment to police departments. Section 1033 of that Act, commonly known as the 1033 Program, is administered by the DoD’s Law Enforcement Support Office (LESO), a section of the Defense and Logistics Agency (DLA). The 1033 Program is essentially an in-kind federal grant to local governments designed to supply surplus defense items to state and local law enforcement agencies (LEAs). The DLA oversees the distribution of new equipment and the decommissioning/processing of items that military units no longer want or require. Excess items are returned to the DLA’s disposition centers, where they are inspected and catalogued. After a 21 day screening period, The DLA makes decommissioned items available for claims by entities that have use for those items. Other military units have first rights, followed by the coast guard, then other federal agencies. Military units from foreign allies have fourth preference and state and local law enforcement can claim what has been passed over by those other groups.

⁵One could argue that the difference is semantics - that the self-reported data are correct, if only 428TK of those were killed. However, it is also true that non-felons should not be killed, and if non-felons are killed, those killings should be disclosed.

Participation into 1033 is voluntary. Agencies must request to join the program and their application must be approved by both the 1033 Program state coordinator and the DLA. To be eligible, agencies are to be responsible for the enforcement of federal, state, and local laws, and their officers must have powers of arrest and apprehension. Agencies do not pay a nominal price for these items, but do incur all logistical costs, including the cost of evaluating items in person and any shipping or transportation costs. It is these costs that have been central to the instrumental variables strategies employed in prior work (Harris et al., 2017; Masera, 2021a). Regulation across states may require flat fees, but some states charge fees depending on the size of the agency or the type of equipment.⁶

LESO reports that 8,200 LEAs from 49 states and territories have participated in the program at least once. If multiple claims are made on the same items, the Program gives favorable consideration to areas that are designated High Intensity Drug Trafficking Areas (HIDTA). The term HIDTA was defined by the Anti-Drug Abuse Act of 1988 and assists federal, state, local, and tribal LEAs operating in critical drug-trafficking regions. Currently 33 regional HDTAs have been designated in all 50 states and territories comprising over 900 initiatives to share investigations, interdiction, and intelligence. The local law enforcement agency must ask to be designated as a HIDTA, and the designation reflects greater willingness and enthusiasm to collaborate with federal drug interdiction efforts more so than any thing else.⁷

While we would generally expect that the extent to which law enforcement agencies participate in the 1033 program is driven by median voter preferences, the reality is a bit different. In practice, acquisitions are opaque to local authorities and the

⁶For example, the South Carolina Department of Administration requires annual fees conditional on the department size to participate in the Program. See <https://www.admin.sc.gov/surplus/1033-Program>

⁷Eligibility requires that an area must be a center of illegal drug production, manufacturing, importation, or distribution; state, local, and tribal LEAs must allocate resources to counteract drug trafficking; drug-related activities must have harmful impact in the area; and federal resources are deemed as necessary to respond to drug trafficking in the area.

public ([Bruce et al., 2019](#)) as there is no signoff on 1033 receipts required from local government outside of law enforcement.

Finally, although the 1033 program is most known for transferring tactical equipment, over 80 percent of transfers are non-tactical in nature. 1033 participants may choose to request armored vehicles, aircrafts, assault weapons, and military clothing, but can also request office supplies, vehicle parts, and medical supplies. In fact, non-tactical transfers represent over 70 percent of items ([Bruce et al., 2019](#)). From 1997 to 2021, the program contributed \$7.6 billion dollars in military equipment to police departments.⁸ However, Law Enforcement Support Office (LESO)’s website states that less than 2 percent of equipment transferred are small arms and less than 1 percent are tactical vehicles. Nevertheless, it is these tactical items that are likely to have the closest ties to changing operational practices and agency culture.

4.1.3 Mixed Evidence on Police Militarization and Civilian Deaths

Previous research has examined the impact of receiving military hardware on violent crime, and violence both by and against police officers. Despite mixed findings ([Steidley and Ramey, 2019](#); [Ramey and Steidley, 2018](#); [Mummolo, 2018](#)), evidence suggests that the procurement of military gear by police departments may contribute to a reduction in violent crime ([Masera, 2021b](#); [Harris et al., 2017](#); [Bove and Gavrilova, 2017](#)). Advances in the quality and availability of 1033 program data led to inquiries on these effects ([Gunderson et al., 2021](#)) but recent analyses indicate that while the program may reduce violent crime, it potentially creates negative spillover effects in neighboring regions, leading to unaddressed adverse externalities ([Masera, 2021b](#)). Previous work has also found the acquisition of military equipment is associated with a heightened risk of assault or injury to officers ([Masera, 2021a](#)), with these effects

⁸Estimates are based on the items’ Original Acquisition Value (OAV). In Fiscal Year 2021, \$196 Million (OAV) worth of equipment was conditionally transferred to participating law enforcement agencies ([DLA, NA](#)).

more pronounced for specific types of equipment ([Carriere and Encinosa, 2017](#)). In a related paper, [Mummolo \(2018\)](#) indicates that the deployment of militarized SWAT teams does not impact police safety. Nonetheless, militarized deployment diminishes trust in law enforcement, particularly among African-American communities, which are disproportionately affected by these deployments.

Violent police behavior is also a focal point in policy discussions. Public interest surged after the 2014 fatal shooting of Michael Brown and the subsequent riots in Ferguson, Missouri.⁹ Early research indicated a correlation between military hardware and an increase in civilian deaths by police, suggesting that the 1033 program transfers altered the material and operational capabilities of police forces, adversely affecting their cultural and organizational norms ([Lawson Jr, 2019](#); [Delehanty et al., 2017](#)). These findings, however, were contested by causal studies, which found no link between the acquisition of military equipment and civilian fatalities ([Harris et al., 2017](#); [Bove and Gavrilova, 2017](#)). However, both of these papers relied on self-reported killings of civilians in the UCR data. Recent findings using Fatal Encounters data [Masera \(2021a\)](#) demonstrate that tactical gear has led to an uptick in civilian deaths at the hands of police. While there are differences in the identification strategy, the treatment of the durable nature of tactical items received, and the unit of observation, the divergence in these studies' conclusions can largely be attributed to their data sources: while [Harris et al. \(2017\)](#) used homicide statistics reported by police departments to the UCR, [Masera \(2021a\)](#) based their analysis on third-party data (Fatal Encounters dataset), bypassing voluntary disclosure mechanisms. This discrepancy highlights an aspect of police behavior —reporting practices— that remains unexplored.

⁹This event prompted the federal government to initiate measures aimed at overseeing and minimizing the lethal capabilities available to local police. Both the House of Representatives and the Senate Committee on Homeland Security conducted independent hearings to scrutinize the equipment and practices of police departments. In January 2015, President Obama enacted Executive Order 13688 to oversee and limit the transfer of military hardware to law enforcement agencies. Despite these efforts, policy responses have been divided, and President Trump repealed this order in 2017, intending to reinstate the equipment and capabilities provided to state, tribal, and local law enforcement agencies.

4.2 Data

We use four different datasets: the FBI’s Uniform Crime Reporting (UCR) Supplementary Homicide Report (SHR), the Fatal Encounters dataset, the Census of State and Local Law Enforcement Agencies (CSLLEA), and a longitudinal panel of 1033 shipments to LEAs.

The first data source comprises self-reported crime statistics in the UCR available from 1976 to 2020.¹⁰ UCR encompasses 12,888 unique Originating Agency Identifier (ORI) codes recording 756 thousand crimes resulting in the injury or homicide of a “victim” or “offender.” The data documents demographic characteristics of all victims and offenders, the location and characteristics of the law enforcement agency, type of crime, weapon, and circumstances involving the killing of an offender by the police or by another civilian.¹¹ In total, 22,836 civilians (out of 892 thousand offenders) are characterized as felons killed by the police in UCR since the beginning of the dataset, and 12,087 since 2000. It is lawful for agencies to abstain from reporting to the FBI, such that the data feeds from self-enrollment by LEAs.

The second dataset is Fatal Encounters, compiled by the University of Southern California. Fatal Encounters is compiled by a team of researchers who fact-check publicly available sources and news media, public record requests to obtain data from federal, state, and local law enforcement, and reports verifying information through

¹⁰Kaplan (2021) comprehensively describes and examines the UCR and SHR data. Chapter 6 raises the issue of voluntary disclosure and miscounted homicides in SHR. The voluntarily disclosed data in UCR is compared with death certificates in the CDC wonder dataset. Murders have a similar trend in both datasets, but UCR under-represents homicides by approximately 2,500 homicides per year. Unlike general murders, which may be studied using death certificates, there are no other public government records of homicides caused by the police. As described in the text, alternative sources are more representative of these instances.

¹¹We code a killing by the police as an event circumstance stating “felon killed by police.” However, technically the person killed does not need to have committed a felony (Kaplan, 2021). Whether the killing of the individual was justified or not, or whether the individual was a criminal, is not observed in the data. There is a sub-circumstance variable that classifies the event on whether the felon was killed in the commission of a crime, the felon attacked a police officer, the felon attacked a civilian, the felon resisted arrest, the felon attempted to flight, or there was not enough information. The use of the word “felon,” however, may be a misnomer as other crimes may be involved or no crime.

published media and public records. The dataset includes 31,497 death records since 2000 and is one of scant sources enhancing quality of police killings and use-of-force statistics. Fatal Encounters records the name and sociodemographic characteristics of everyone killed by the police, the name of departments involved, and a series of qualitative information regarding the event, such as whether the death was suicidal or an accident, whether there was a vehicle, violence, drugs, or guns involved, if the police officer was off-duty, and whether the case was justified. Analysis suggest that, among non-official sources, Fatal Encounters is the most reliable database for police killings (Finch et al., 2019; Ozkan et al., 2018).¹²

The third dataset comprises 1033 records in a longitudinal panel of shipments to LEAs. The 1033 Program publishes “snapshots” corresponding to a cross-section with the stock of transfers from LESO to LEAs in the quarter when data is released (Koslicki, 2022). We build on the replication files in Connealy et al. (2023) and construct a longitudinal panel based on shipments of 1033 equipment transferred to police departments. To construct the data, we identify each unique police agency across all quarterly snapshots, and each unique shipment defined as the correlate between agencies, the shipment date, and the NATO Stock Number (NSN). In the final sample, there are 5,946 LEAs participating in the program with information on the quantity, value and items transferred to each department.¹³ The panel shows that LESO transferred 3.3 billion dollars of initial acquisition value in excess military gear.¹⁴

¹²Finch et al. (2019) compares Fatal Encounters to a sample of full records from 11 states and 328 agencies. They find Fatal Encounters has incomplete records in only two states: Connecticut (92 percent completion) and Florida (95 percent).

¹³Originally the dataset includes 7,060 LEAs in Section 1033. The LESO website places the number in 8,200 police departments, including federal agencies. Because the 1033 Program starts in 1997, we drop 1,114 agencies whose shipments occur before the onset of the program. See appendix Figure C.5 with the distribution of program take-up.

¹⁴This is an important disparity with official records from LESO, as the program has transferred \$7.6 billion since 1990. There are several explanations for disparities, and they highlight the limitations of publicly available 1033 data. For example, items returned to LESO are removed from the dataset, and they are not captured in the data if they were returned before the first Snapshot (last quarter of 2014), although Masera (2021a) asserts that no item is returned before 2014. In addition, items classified as low-risk are removed from the dataset after an agency owns

Table 4.1 reports the total quantity of items transferred to LEAs.¹⁵ Following the literature, we use the NSN to group transfers in five categories: Weapons and Ammunition, Vehicles, Optics, Clothing, and others (Koslicki, 2022; Harris et al., 2017). The first panel in Table 4.1 shows total shipments in each of these categories. Between 1990 and 2020, LESO shipped 3 million items and 431 thousand of these were weapons and ammunition, 23 thousand were vehicles, and 144 thousand were optics. There are alternative classifications of equipment across research (Koslicki, 2022). The second panel in Table 4.1 follows Masera (2021a) with an alternative classification, dividing transfers in tactical gear, non-tactical items, and military vehicles. The last panel in Table 4.1 shows transfers of rifles, machine guns, bayonets, and pistols; and magazines and ammunition.¹⁶

Lastly, we use the law enforcement census –CSLLEA (United States Bureau of Justice Statistics, 2023, 2011b,a, 2009). CSLLEA records the location, budget, number of sworn officers, number of employees, and other characteristics for 17,690

them for a year (DEMIL A code is used to designate low-risk clothing and wearable items by the DoD). Another explanation relates to nuances in aggregating the value of items: the value may be recorded per unit or per shipment across snapshots, without clear identification. For this reason, we base our study in quantities transferred instead of the value.

¹⁵Appendix Table C.1 is the correlate in the OAV of shipments.

¹⁶Koslicki (2022) has a thorough discussion of categories. We use the NATO Stock Number or National Stock Number (NSN) to define categories. The NSN is a 13-digit alpha-numeric code standardized across NATO members. The first and second digits define supply groups: 10–Weapons, 11–Nuclear Ordnance, 12–Fire Control, 13–Ammunition and Explosives, 14–Guided Missiles, 15–Aircraft and Airframe Structural Components, 16–Aircraft Components and Accessories and 17–Aircraft Launching, Landing, and Ground Handling Equipment, 18–Space Vehicles, 19–Ships, Small Craft, Pontoons, and Floating Docks and 20–Ship and Marine Equipment, 22–Railway Equipment, 23–Ground Effect Vehicles, Motor Vehicles, Trailers, and Cycles, 24–Tractors, 26–Tires, and others (28 to 99). Masera (2021a) uses a three-way category consisting of military vehicles (NSN 23–Ground Effect Vehicles and 15–Aircrafts), tactical items and weapons (10–Weapons, 1240–Optical Gear, 5855–Night Vision Equipment, 13–Ammunition), and non-tactical (e.g. 71–Furniture, 75–Office Supplies, 65–Medical, and more). We use a five-way classification in line with Harris et al. (2017): Weapons and Ammunition (10–Weapons and 13–Ammunition), all vehicles (15–Aircraft, 18–Space Vehicles, 19–Ships, 23–Ground Effect Vehicles, 24–Tractors), Optics (1240–Optical Sighting and Ranging Equipment and 5855–Night Vision Equipment, Emitted and Reflected Radiation), Clothing (84–Clothing, Individual Equipment, and Insignia), and all other items. We generate two additional categories, “Rifles, Guns” includes all transfers of rifles, M240, M4, M5, machine guns, pistols, and bayonets; and “ammunition” includes magazines and ammunition.

police departments.¹⁷ Census data is a pooled cross-section constructed from matching files in 2000, 2005, 2012, and 2018 using the Law Enforcement Agency Identifiers Crosswalk Series ([United States Bureau of Justice Statistics , 2004](#); [United States Bureau of Justice Statistics, 2007, 2018](#)).

We find 10,777 UCR respondents in the CSLLEA using their ORI code (93.59 percent of all agencies in UCR). In the case of Fatal Encounters, we match agencies using the state and name of the department, finding 5,058 police agencies from Fatal Encounters in the CSLLEA (96.41 percent of agencies in Fatal Encounters).¹⁸ 85% of agencies in the 1033 data are found in the analytical dataset, for a total of 5,946 police departments receiving 1033 transfers.¹⁹

The analytical sample has 12,462 agencies in the CSLLEA collapsed into a yearly panel of civilian killings. Data is aggregated at yearly levels, mitigating lags in official reports ([Kaplan, 2021](#)). We restrict the sample period from 2000 to 2020, due to availability in both Fatal Encounters and UCR.²⁰ 9 percent of all UCR respondents self-report killing a civilian at any point in time. Of the six thousand agencies that take-up military gear, 81 percent are UCR respondents but only 13 percent report killing a civilian. In stark contrast, Fatal Encounters records that 43 percent of 1033 participants have killed a civilian. Figure 4.2 displays the increasing trend in non-disclosed homicides by police, computed as the total killings by police in Fatal

¹⁷There are over 18 thousand departments recorded. However, 9 percent of the data have no county available. We impute the FIPS code to these counties allocating a random county within the ZIP code of these departments. 296 agencies remain unidentified and are dropped from the sample. Lastly, all observations without data on population size in the jurisdiction and total sworn officers are dropped.

¹⁸Following [Masera \(2021a\)](#), we restrict the sample of killings in Fatal Encounters by dropping all instances involving more than one agency (3,818 or 12 percent of total killings by police), and dropping all observations where police officers are off-duty (802 observations). Special characters are removed from the department name, and all names are converted to lowercase. Department names are simplified to account for common abbreviations, such as “dept,” “dep,” or “department.” Finally, since the Fatal Encounters data is based on the location of the death rather than the location of the agency, regular expressions are used to match parentheses describing cases where state or police departments crossed borders. We then examine the matching process on a case-by-case basis and compared the data with information press reports.

¹⁹As explained above, however, 1,114 agencies are dropped as their transfers begin before 1997.

²⁰Following the literature, we restrict the sample to the continental US ([Masera, 2021a](#); [Harris et al., 2017](#)).

Encounters minus those recorded in UCR, and the total stock of 1033 items acquired by LEAs.

4.2.1 Characterizing reports of homicides by police

We define under-reporting as the positive difference between the number of homicides by police reported in the UCR and the number of homicides recorded in Fatal Encounters. Under-reports capture the gap between self-reported and actual homicides illustrated in Figure 4.1 and 4.2.²¹ Our analytical database tracks the yearly correlate of this gap per agency, thereby documenting police participation responses in UCR, the total number of self-reported homicides by the police, and the total number of civilians killed by the police as recorded in Fatal Encounters.²²

Under-reporting is a nationwide phenomenon. The geographic distribution shown in Figure 4.3 reveals state variation in the percent of agencies that do not fully disclose the tally of homicides. South Carolina, Louisiana, Florida, Arkansas, and West Virginia report the highest per capita fatalities. However, police departments across all states have reported incomplete figures: Hawaii (100%), Nevada (71%), California (69%), Arizona (56%), and Florida (52%) have the highest percentage of police departments that under-reported during the period. Conversely, there are

²¹Appendix Figure C.1 illustrates the total count of civilians killed by police by the race of victims. The racial background is recorded in the UCR data as a characteristic of the offender, while Fatal Encounters uses demographic characteristics that are publicly available. When race is unspecified or withheld by authorities, Fatal Encounters generates an imputed value. This figure plots the annual sum of killings, comparing three variables: self-reported counts in the UCR (blue), the count of deaths in Fatal Encounters (red), and the count in Fatal Encounters using imputed racial backgrounds (green). The main insight from Figure C.1 is that the increasing gap in reported killings is driven by the homicides of underrepresented minorities. Specifically, in the left panel, the killings of white civilians by law enforcement show an upward trend in all series, including self-reported police data from the UCR, which converges with Fatal Encounters data. However, the middle and right panels diverge, showing that racial minorities are experiencing an upward trend only in third-party statistics, while the self-reported UCR figures have remained constant since 2000.

²²The distribution of this gap is illustrated in Appendix Figure C.2. Less than 1% of all records are negative, that is only in 1,225 instances Fatal Encounters documents fewer homicides than UCR. Our study concentrates on the under-represented statistics by local law enforcement rather than the accuracy of Fatal Encounters, setting it apart from other studies (Finch et al., 2019; Ozkan et al., 2018). Thus, to analyze police under-reporting, we code the gap as zero in instances where the police reported more deaths than those documented in public records.

less than 15% of agencies under-reporting in South Dakota, North Dakota, and New Hampshire. Total undisclosed homicides by police range from 2 non-reported homicides to 112.3 homicides per 100,000 inhabitants.

Figure 4.4 explores trends in the total and average under-reports conditional on agencies having killed a civilian and their disclosure choice. The left panel in Figure 4.4 shows total under-reported homicides among never-respondent agencies (blue), previous UCR respondents (red), and current respondents (green). One counter-intuitive finding is that agencies that never report to UCR are not the main drivers of the upward trend in civilian deaths, but current respondents have the largest tally. In the right panel, however, the trend reverses when looking at average under-reports per agency adjusted by population. In a year, on average, police officers are responsible for 0.43 homicides per 100,000 when their department does not respond to UCR, 0.22 civilians in agencies that previously reported, and 0.10 civilians per 100,000 among respondents. This may point out the importance of accountability but also suggests that there are different features among UCR respondents.

Trends discussed above suggest differences among UCR respondents and non-respondents. Table 4.2 presents the non-parametric statistical differences in the characteristics of UCR respondents and non-respondents. The first two rows show that UCR respondents are more likely to have officers involved in the homicide of civilians, both in the Fatal Encounters data and in self-reported records. There are no differences in the population size in the jurisdiction of these agencies or the budget per capita (or per officer). However, UCR Respondents have 3 times the total budget of non-reporters, and 3 times more police officers. There are also differences in the administrative jurisdiction: sheriff's offices are more likely to respond, while local police departments tend to be non-respondents. Lastly, reporting agencies are more likely to investigate homicides, have specialized task forces to combat gangs, anti-explosive departments, K9 Units, Crime Analysis Units, SWAT teams, Basic Training, and are more likely to perform Firearm Background Checks.

Table 4.1: Total quantity of 1033 items shipped by category.

Period	Total	Weapons	Vehicles	Optics	Clothing	Others
1990-1994	2,726	2,659	16	48	-	3
1995-1999	14,589	13,182	224	1,083	-	100
2000-2004	22,687	21,264	273	1,097	-	53
2005-2009	80,505	70,242	591	5,837	2,413	1,422
2010-2014	1,185,691	227,361	10,270	78,978	203,400	665,682
2015-2020	1,819,511	97,110	11,738	57,314	444,134	1,209,215
Total	3,125,710	431,818	23,113	144,357	649,947	1,876,475

	Total	Tactical	Non-Tactical	Military Vehicles
1990-1994	2,726	2,707	3	16
1995-1999	14,589	14,265	100	224
2000-2004	22,687	22,361	53	273
2005-2009	80,505	76,079	3,860	566
2010-2014	1,185,691	306,339	869,388	9,964
2015-2020	1,819,511	154,424	1,654,469	10,618
Total	3,125,710	576,175	2,527,873	21,662

	Guns, Rifles	Magazines
1990-1994	2,656	-
1995-1999	12,343	-
2000-2004	20,853	-
2005-2009	65,317	1,776
2010-2014	90,314	108,928
2015-2020	20,040	73,801
Total	211,523	184,505

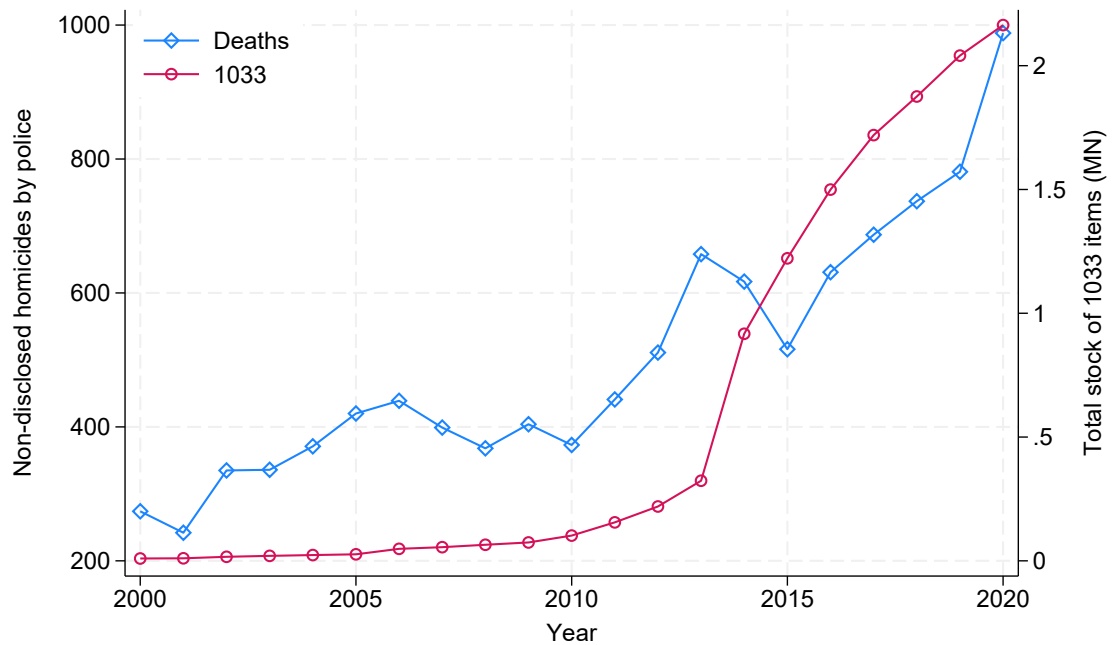


Figure 4.2: Total non-disclosed homicides by police and the stock of military items.

Notes: The figure plots the relationship between non-disclosed homicides by police and the total stock of military items held by police departments in a given year. In blue, the series is the difference between the total number of homicides by police reported in Fatal Encounter and the sum of ‘felons’ killed by police in UCR. In red, the series shows the cumulative items shipped through the 1033 program in millions.

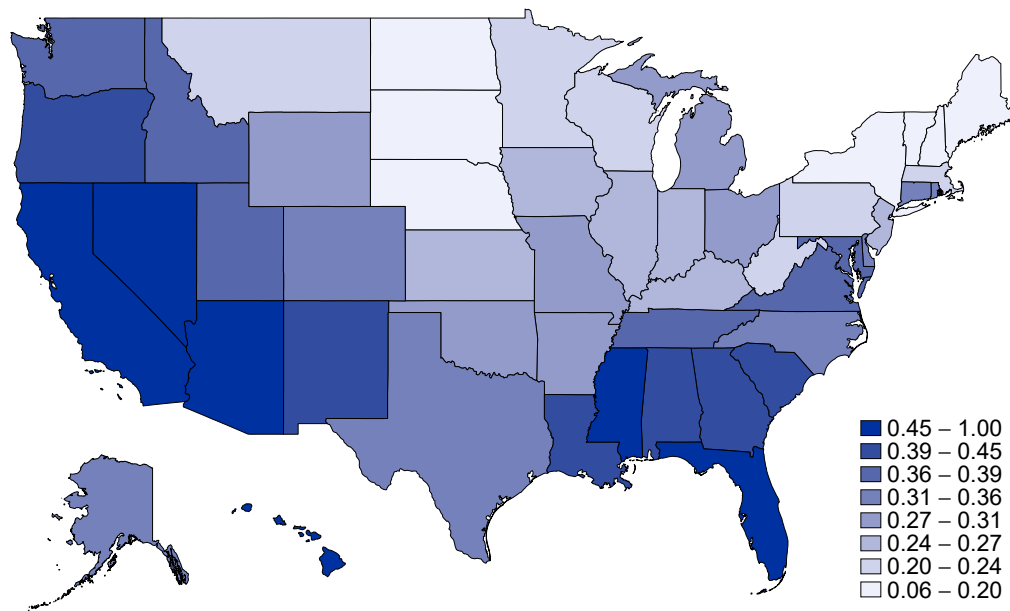


Figure 4.3: Percent of agencies under-reporting per state (2000-2020).

Notes: The map plots the percent of agencies that under-report at least one homicide. Under-reported killings are defined as the yearly difference between total killings in Fatal Encounters and the total killings recorded in UCR.

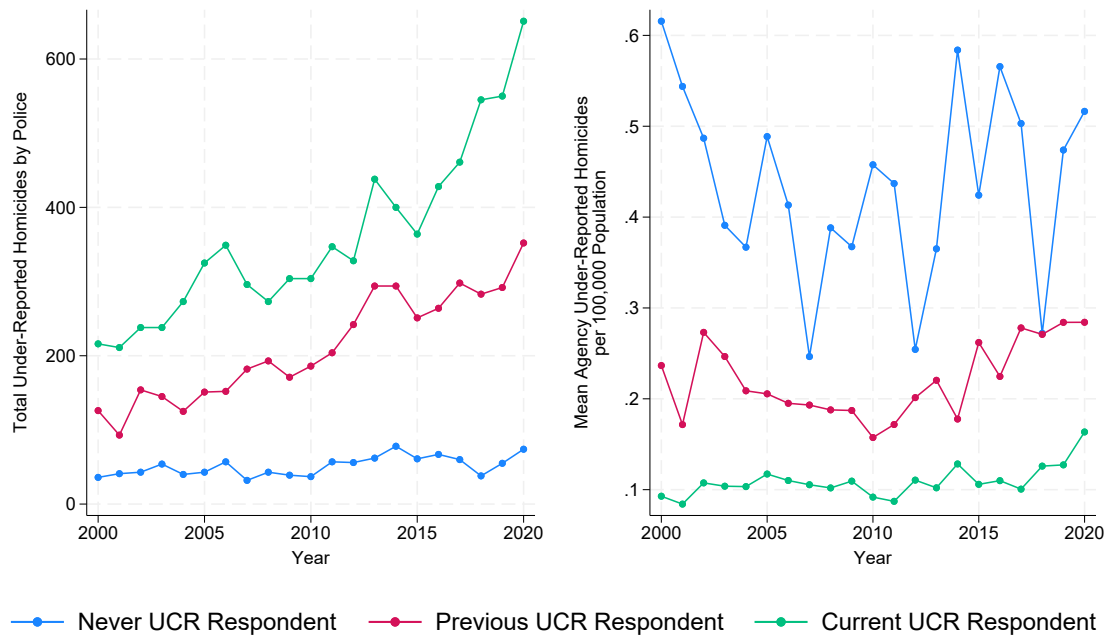


Figure 4.4: Gap between Fatal Encounters and UCR among agencies that killed a civilian.

Notes: The figure plots the positive difference between homicides by police recorded in Fatal Encounters and UCR. The left panel shows total under-reported homicides per year. The right panel shows the mean under-reported homicides by police per 100,000 population. Three types of departments are considered: departments that never responded to UCR, departments that previously responded to UCR but did not during that year, and current respondents.

Table 4.2: Balance test of agency characteristics and disclosure of civilian homicides by police.

	Non-Respondents	UCR Respondents	Difference	p-value
Fatal Encounters total 2000-2020	0.50	2.34	-1.85	0.00
UCR total 2000-2020	0.00	1.12	-1.12	0.00
Jurisdiction Population (000s)	42.91	56.27	-13.36	0.21
Agency Budget (100,000)	45.28	162.56	-117.28	0.00
Total budget per 100,000 population	495.88	722.48	-226.60	0.17
Agency Budget per officer (000s)	205.11	401.54	-196.42	0.02
Full-time sworn officers	20.21	58.57	-38.36	0.00
Part-time sworn officers	3.15	2.80	0.34	0.01
Total full-time employees	32.44	85.92	-53.48	0.00
Sheriff	0.11	0.24	-0.13	0.00
Local/County/Regional Police	0.89	0.76	0.13	0.00
State Police	0.00	0.00	0.00	0.02
Investigates Homicides in 2018	0.40	0.77	-0.36	0.00
Task Force for Gangs 2018	0.22	0.42	-0.20	0.00
Bomb/Explosives Disposal	0.03	0.07	-0.04	0.00
K9 Unit	0.23	0.53	-0.30	0.00
Crime Analysis Unit	0.24	0.46	-0.22	0.00
Firearm Background Check Unit	0.33	0.47	-0.14	0.00
Tactical Operations (SWAT)	0.18	0.49	-0.31	0.00
Basic Training Academy Unit	0.02	0.06	-0.04	0.00
N Non-Respondents	4463			
N Respondents	6820			

Notes: The Table reports the non-parametric t-test of the difference in the mean agency characteristics of UCR respondents and non-respondents. Agency characteristics are recorded in the CSLLEA 2018.

4.2.2 Characterizing 1033 participation

Table 4.3 summarizes the non-parametric differences among 1033 takers and never-takers. 1033 participants have larger tally of homicides by police in both Fatal Encounters and UCR. Furthermore, 1033 takers are 2 percentage points more likely to have under-reported homicides by police captured by a larger tally in Fatal Encounters than in the UCR data. Police departments are larger among 1033 Program participants, with three times the population size, twice the budget, and more sworn officers. However, the budget per capita and the budget per officer are statistically identical across participants and non-participants.²³ 1033 take-up is predominant among sheriff's offices, while local police departments tend to be non-participants.

1033 takers tend to be responsible for more homicides, but there are qualitative differences in the transfers received by departments. UCR respondents are more likely to participate in Section 1033, but they tend to take-up less weapons and vehicles. Furthermore, UCR respondents and non-respondents do not differ in their take-up of clothing, optics, and total transfers. Table 4.4 documents statistical differences in the quantity and OAV adjusted by the number of sworn officer in each department.²⁴ There are statistical differences in the quantity of vehicles and weapons, as UCR non-respondents are more likely to acquire these items. The Table showcases a counter-intuitive phenomena: UCR respondents are more likely to participate but they acquire fewer tactical items.

Take-up year varies across departments such that we observe under-reports before and after the first 1033 shipment by category. We posit that upward trends in under-reporting before and after take-up may suggest adverse selection, as departments with growing cover up culture are likely to under-report. We also advance the hypothesis

²³These statistics are from the 2018 Census. Because 1033 transfers predate agency characteristics, departments may leverage in-kind transfers to reduce their capital investments (Bruce et al., 2019).

²⁴These patterns remain using transfers per 100,000 population. UCR respondents have statistically larger take-up when using unadjusted statistics for population or department size.

that structural breaks in the trend after first shipment are indicative of moral hazard. These trends are explored in Figure 4.5, where the horizontal axes captures the first year in which an agency received 1033 shipments ($t = 0$) and, in the vertical axes, the probability that a department will not disclose the total number of killings by police. Behavioral changes after the first acquisition are suggestive of moral hazard, while growing trends before and after acquisition may be suggestive of adverse selection. Each panel in Figure 4.5 shows under-reporting before and after the acquisition of a different category of items: all 1033 transfers, weapons, vehicles, optics, clothes, and the rest.²⁵

Summarizing Figure 4.5, we find suggestive evidence of adverse culture in agencies that take-up 1033 transfers of all kinds. The upward trend implies that 1033 participants had a growing level of homicide under-reporting prior to receiving 1033 transfers and this upward trend continued after transfers. The idea that 1033 participant agencies had increasing levels of under-reporting prior to take-up suggests that LEAs had some inclination toward cover-up culture regardless of the 1033 Program. Further evidence requires an appropriate counterfactual analysis, which we develop in the upcoming section. In addition, Figure 4.5 hints at moral hazard following the shipment of weapons. The discontinuous jump in under-reporting signals that police departments their under-reporting behavior after receiving 1033 transfers. The trend in increasing under-reports becomes steeper in all categories, including clothing and other shipments.

4.3 Empirical Strategy

To examine how military grade equipment affects reporting behavior, let Δreport_{it} be the difference between killings by police recorded by a third-party source and self-reported killings in agency i and time t . The estimating equation is:

²⁵See Appendix Figure C.6 with the distribution of take-up by category.

$$1(\Delta\text{report}_{it} > 0) = \delta\text{items}_{it} + \beta X_{it} + \lambda_i + \lambda_t + \epsilon_{it} \quad (4.1)$$

The dependent variable $1(\Delta\text{report}_{it} > 0)$ captures the extensive margin decision of not disclosing homicides by police.²⁶ Departments select whether to participate in the UCR or not, such that underreports in non-participants may highlight different underlying behavior. Therefore, we test the robustness of the results in a sample of UCR respondents, where the dependent variable underscores the omission of homicides by police officers from the dataset.

The parameter of interest δ is the marginal effect of one additional item on the probability of having a positive difference between actual homicides and reported homicides.²⁷ We define the number of items as the stock of quantities acquired by the police department up to year t using the shipment date across all snapshots.²⁸ The specification controls for idiosyncratic characteristics in the department σ_i and unobserved yearly shocks σ_t . We control for cross-county and time variation in the population size, labor force, and unemployment. The specification includes time-varying agency characteristics: the operating budget, number of sworn officers, and total number of employees.²⁹

We tackle several concerns in the literature. First, previous studies have been criticized for an ecological fallacy related to the difference in treatment units (police departments) and county-level outcomes (Gunderson et al., 2021). The original 1033 files were aggregated by state and county until the release of agency-level information

²⁶The intensive margin decision is presented in the Appendix Tables C.4-C.5, where $\Delta\text{report}_{it} = \text{report}_{it}^{\text{Fatal Enc}} - \text{report}_{it}^{\text{UCR}}$ and $\text{report}_{it}^{\text{Fatal Enc}} - \text{report}_{it}^{\text{UCR}} < 0 \rightarrow \Delta\text{report}_{it}$.

²⁷Most police departments do not participate in 1033. Therefore, we report a lin-log functional form and show our results remain robust in Appendix Table C.5.

²⁸We avoid double counting items repeated in snapshots by identifying unique shipments per department and NSN. Therefore, a transfer of items is considered as a unique shipment whenever the shipment date and NSN match in different snapshots. The procedure follows from Connealy et al. (2023) and Koslicki (2022).

²⁹CSLLEA data is available in 2000, 2004, 2008, and 2018. We use the Law Enforcement Agency Identifiers Crosswalk to match each census (United States Bureau of Justice Statistics, 2018, 2007; United States Bureau of Justice Statistics, 2004). However, the 2004 CSLLEA does not have agency identifiers to match agencies across the census. Missing years are replaced with the police department's average.

in 2015. Therefore, our analysis is entirely based on police departments, contrasting the behavioral decision of LEAs and treatment at the same unit of analysis. Second, we measure treatment in the quantity of items held by the department as there have been documented inconsistencies in the OAV of items, with at least 50,000 items listed with zero value (Gunderson et al., 2021).³⁰ Items with zero value may be conspicuous (e.g., wrenches and wires), or controlled items including weapons, optics and night vision gear, and tractors. The quantity of items also has a straightforward economic meaning: δ captures the marginal effect of an additional item in the hands of the police department on the probability of under-reporting homicides.

Our identification strategy exploits a Bartik-type instrument inspired in Masera (2021a). Time variation is derived from the number of boots on the ground in Afghanistan and Iraq during the period. The idea is that new military gear acquired when deploying units may generate excess military gear. Similarly, the withdrawal of troops leads to used and unused equipment available to police departments after returning to the US. Cross-sectional variation is derived from the distance between the police department and the nearest DLA distribution center (Masera, 2021a; Harris et al., 2017) and the distance to any military base (Masera, 2021a).³¹ The first stage is:

$$\text{items}_{it} = \gamma (\text{eqp}_t \times \text{close}(\text{Disposition})_i) + \psi (\text{eqp}_t \times \text{close}(\text{Mil Base})_i) + \phi X_{it} + \xi_i + \xi_t + \mu_{it} \quad (4.2)$$

³⁰Data inconsistencies were larger before 2017. In the first quarterly datasets, up to 6,614 shipments are recorded with no acquisition value. The latter datasets have less than one hundred such observations. However, the total quantities of items listed have decreased from 1.79 million in the 2014 snapshots to 450 thousand in 2013. This may be because low-risk items are not tracked after a year in possession of police departments. Our results are consistent with using the OAV of items in Appendix Table C.6.

³¹We depart from Masera (2021a), who constructs a binary indicator for departments closer than 20kms from a storage facility. Instead, we use the continuous distance between DLA distribution centers and police departments. Our results are robust to a non-linear function of the distance, but not to the binary indicator which produces a weak first stage.

where eqp_t measures the difference between boots on the ground in Afghanistan and Iraq in 2009 and subsequent years, normalized to 1 in 2014 as in [Masera \(2021a\)](#). The index captures the availability of 1033 items as a function of deployed boots on the ground and is defined according to the equation:

$$\text{eqp}_t = \begin{cases} 0 & t < 2010 \\ \frac{\text{boots}_{2009} - \text{boots}_t}{\text{boots}_{2009} - \text{boots}_{2014}} & t \geq 2010 \end{cases}$$

The coefficient γ captures a shift in the availability and the relative cost of items for departments. Cross-time availability of items is as good as random, independent of any agency's needs as it responds to military operations abroad. Similarly, departments close to a disposition center face fewer shipping cost, measured by the distance. The interaction between boots on the ground and distance to the nearest military base, captured by ψ , allows for comparing police departments that are close to a disposition center with respect to those that are close to military bases that are not disposition centers. Therefore, it mitigates concerns related to the strategical location of disposition centers in selected military bases, where police departments near these bases may have unobservable differences that threaten causal inference. Previous studies exploit time-variation in boots on the ground interacted with a binary indicator for police departments closer than 20kms to a disposition center, and year fixed effects interacted with a binary indicator capturing closeness to a military base ([Masera, 2021a](#)). We depart from this specification by interacting a continuous measurement of closeness to a military base with the normalized boots on the ground variable. Other studies use similar cross-sectional variation but interact the distance to disposition centers with the total national transfers in 1033 equipment as a measurement of items' availability ([Bruce et al., 2019](#); [Harris et al., 2017](#)). On the other hand, [Bove and Gavrilova \(2017\)](#) relies on time variation in the US department of defense budget and cross-sectional variation in the between-counties variation in the probability of receiving military aid.

Table 4.3: Balance test of agency characteristics and 1033 take-up.

	Never-Takers	1033 Takers	Difference	p-value
Fatal Encounters total 2000-2020	0.75	2.56	-1.81	0.00
UCR total 2000-2020	0.26	1.14	-0.87	0.00
Under-reporting	0.01	0.03	-0.02	0.00
Jurisdiction Population (000s)	23.29	73.80	-50.51	0.00
Agency Budget (100,000\$)	64.21	171.67	-107.46	0.00
Total budget per 100,000 population	599.29	669.75	-70.47	0.67
Agency Budget per officer (000s\$)	294.56	358.53	-63.97	0.42
Full-time sworn officers	24.01	63.85	-39.84	0.00
Part-time sworn officers	2.79	3.12	-0.33	0.01
Total full-time employees	33.23	97.40	-64.17	0.00
Sheriff	0.12	0.26	-0.14	0.00
Local/County/Regional Police	0.88	0.73	0.15	0.00
State Police	0.00	0.00	-0.00	0.00
N Never-Takers	5791			
N Takers	5404			

Notes: The Table reports the non-parametric t-test of the difference in the mean agency characteristics of 1033 participants and non-participants. Agency characteristics are recorded in the CSLLEA 2018.

Table 4.4: Balance test of 1033 take-up per sworn officer and disclosure of civilian homicides by police.

	Non-Respondents	UCR Respondents	Difference	P-value
1033 Take-Up	0.38	0.55	-0.17	0.00
1033 Quantities				
Total Transfers	7.40	6.90	0.50	0.72
Weapons	1.14	0.91	0.24	0.01
Vehicles	0.15	0.10	0.05	0.00
Clothes	1.28	1.36	-0.08	0.77
Optics	0.26	0.32	-0.05	0.12
Others	4.56	4.22	0.34	0.78
Rifles, Guns	0.70	0.50	0.20	0.00
Ammunition	0.34	0.42	-0.08	0.27
1033 OAV (000\$)				
Total OAV	13.08	9.95	3.13	0.02
Vehicles	6.85	5.81	1.04	0.12
Weapons	0.30	0.28	0.02	0.49
Clothes	0.08	0.09	-0.01	0.43
Optics	0.65	0.66	-0.00	1.00
Others	5.19	3.11	2.08	0.00
Rifles, Guns	0.18	0.13	0.05	0.00
Ammunition	6.41	7.02	-0.61	0.67
N Non-Respondents	4463			
N Respondents	6820			

Notes: The Table reports the non-parametric t-test of the difference in the mean acquisition of 1033 transfers between UCR respondents and non-respondents. The quantities and original acquisition value (OAV) are the total value between 2000 and 2020.

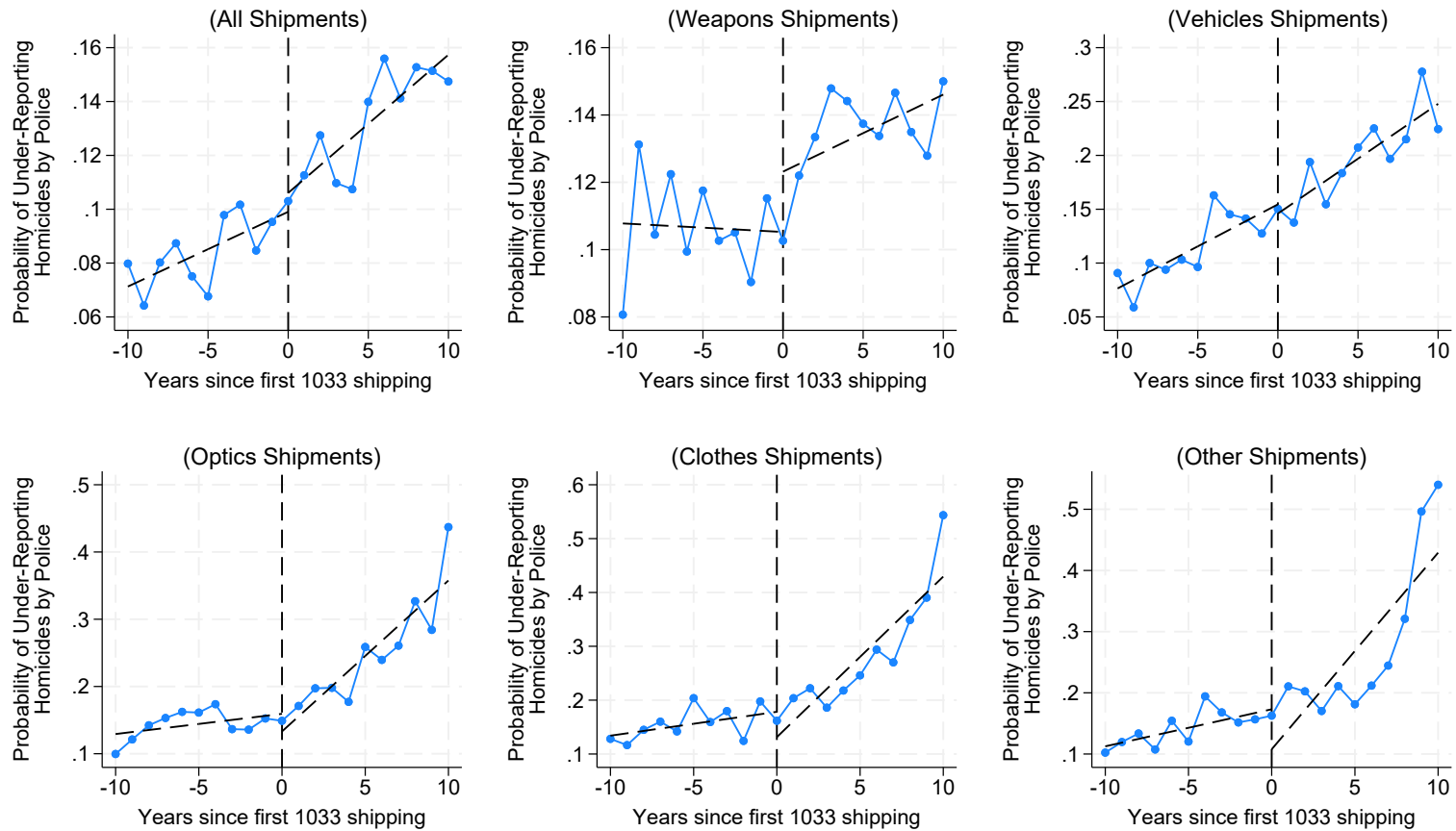


Figure 4.5: Under-reported homicides by police and years since 1033 take up, by shipment type.

Notes: The Figure plots the population-adjusted mean of under-reporting before and after 1033 shipments across different categories of shipments. Under-reporting is defined as a binary indicator for the agency having a larger count of killings by police in Fatal Encounters than in the UCR data.

We complement the analysis with evidence on the behavioral decisions of police departments prior to the acquisition of military hardware. Under-reporting and the take-up of 1033 items are marked by a simultaneity bias, which is mitigated by instrumenting 1033 participation. However, we seek to document whether LEAs with negative underlying behaviors have been more prone to acquire items. To analyze this relationship, our methodological strategy relies on a Markov chain or transitional matrix methodology. We estimate the past under-reporting behavior of police departments as they choose to acquire additional gear. The approach aims at documenting whether the departments that increase their stock of military gear differ in their probability of under-reporting. The Markov chain process is modeled with the equation:

$$\Pr(\Delta \text{reports}_{i,t-1} \mid \text{items}_{it} \in Q_p, \text{items}_{i,t+1} \in Q_k) \quad (4.3)$$

The matrix is constructed by computing whether the police department has under-reported during a window of time $t - 1$. Transition states are defined by whether the stock of items is at the p th quintile of the item per sworn officer's distribution in the initial period t , transitioning to the q th quintile of the distribution in final period $t + 1$. This non-parametric process captures suggestive evidence on whether police departments that choose to take-up gear exhibit larger instances of misreporting behavior. If the probability of under-reporting is larger when departments transition between being light to heavy acquisition of items.

This study is limited, as we cannot casually link initial under-reports to 1033 take-up due to simultaneity bias and lacking data prior to the program. Nevertheless, we document two alternative approaches in Appendix C.0.5. First, we study whether the reporting-behavior of police departments in previous periods correlates with current year variation in the acquisition of military gear. Second, we present a series of two-way fixed effects event-studies that show the dynamic relationship between under-reports and take-up. The parallel trends assumption allows us to distinguish whether

police departments that do not participate in the 1033 program exhibit significant deviations from the participants.

4.4 Results

4.4.1 First stage

Independent of the distance to DLA facilities, there has been an increase in the acquisition of military gear after 2009. Figure 4.6 shows the time variation in the Bartik instrument and the endogenous stock of 1033 items.³² The stock of items doubled by 2004 but was nearly eight times the starting quantities by 2009. By 2014, police departments had 95 times the number of items they had in 2000, and this trend in the acquisition of military gear increased with the withdrawal of troops abroad. The first wave of 2,500 troops was sent to Afghanistan in 2001 and 141 thousand troops sent to Iraq in 2003. The series peaks with 203,600 combined boots on the ground in 2009. The normalized series captures the withdrawal of troops, and its increase matches a growing trend in the stock of items acquired by police departments.

The distance between police departments and disposition centers is a frequent instrument in this literature (Masera, 2021a; Gunderson et al., 2021; Harris et al., 2017). The relevance of this instrument is studied with the following equation:

$$\text{items}_{it} = \rho_i + \rho_t + \chi \text{eqp}_t \times f_1(\text{distance storage}_i) + \phi \text{eqp}_t \times f_2(\text{distance military}_i) + u_{it} \quad (4.4)$$

³²The normalized instrument oscillates between 0 and 1.11 because our analytical sample ends in 2020, while the sample in Masera (2021a) ends in 2014. See appendix figure C.7 for a comparison between boots on the ground and the normalized index. Our results are robust to alternative measure that captures the cumulative absolute variation in boots on the ground in appendix table C.8.

Following Masera (2021a), equation 4.4 analyzes nonlinearities in the relationship between distance to a disposition center and take-up. f_1 and f_2 are fourth-order polynomials of the distance to the closest storage facility and military base, respectively. The outcome variable is the number of items acquired by police departments since the program's start. Figure 4.7 plots the predicted stock of items as a function of distance to a disposition center, bounded at 200 kilometers. Accumulation of items is larger for departments that are close to facilities and decreases with distance. Therefore, consistently with the literature, distance has a negative effect on take-up.

4.4.2 The impact of military hardware on homicides by police

We start by reconciling mixed results in previous studies. We separately estimate a regression capturing the impact of 1033 program transfers on civilians killed by police using self-reported data in UCR, and a second regression for total homicides by police using Fatal Encounters data. In each instance, we use high-quality data matched at the agency level and eight variations in the construction of the first stage instrument. First, the main model uses the preferred specification where the stock of items is instrumented with normalized boots on the ground interacted to the closest DLA facility and a second interaction between boots on the ground and the distance to the closest military facility. The second model includes the interaction between boots on the ground and the closest DLA storage facility, omitting additional interactions. The third model includes a quadratic interaction between the closest DLA facility and boots on the ground. The fourth controls for the cubic distance to the closest DLA facility. The fifth specification includes the cubic distance to a DLA facility and the interaction between boots on the ground and the closest military facility. The sixth model replicates the first specification but uses the cumulative absolute change in boots on the ground instead of a normalized index. The next

specifications partly emulates the identification strategy behind previous studies. The seventh model exploits time variation in the cumulative real military expenditure, similar to [Bove and Gavrilova \(2017\)](#), interacted with distance to the closest storage facility and the closest military facility. The last model uses time variation from the cumulative items transferred through the 1033 program, as in [Harris et al. \(2017\)](#), interacted with the distance to a DLA storage facility and the closest military base.³³

Figure 4.8 displays sixteen estimates.³⁴ The first eight estimates are computed with self-reported UCR data and consistently show null effects, regardless of the first stage specification. The next eight models, in the right panel in Figure 4.8, show that one hundred additional items acquired by police departments increase the number of civilians killed by police, according to the Fatal Encounters data. The results consistently support the null results in [Harris et al. \(2017\)](#) and [Bove and Gavrilova \(2017\)](#), and the positive results in [Gunderson et al. \(2021\)](#) and [Masera \(2021a\)](#). Two different datasets yield substantially different results, underscoring that military hardware may affect the reporting behavior of law enforcement agencies.

³³These specifications are loosely based on previous studies. [Masera \(2021a,b\)](#) uses similar sources of time variation but interacts closeness to a military base with time dummies. [Bove and Gavrilova \(2017\)](#) interacts time variation in real military expenditure with cross-sectional variation in the between-counties variation in the probability of receiving military aid. [Harris et al. \(2017\)](#) exploits time variation in the national aggregate of items transferred through the 1033 Program with the inverse distance to the closest military base, the inverse distance to the sixth closest military base, an indicator on whether the county is designated as HIDTA, and land area. Our unit of analysis also differs from earlier studies based on county data ([Harris et al., 2017](#); [Bove and Gavrilova, 2017](#)). We also use a construction of the 1033 Program data following the most up to date understanding of limitations in the snapshots ([Connealy et al., 2023](#); [Koslicki, 2022](#)). Lastly, standard errors are clustered at the commuting zone by year level ([Masera, 2021a,b](#)).

³⁴Appendix Table C.2 presents these results with the respective first stage statistics and coefficients.

4.4.3 The impact of military hardware on the disclosure of killings by police

The main findings evidence that, at the agency level, the acquisition of military items through the 1033 program affects the reporting behavior of police departments.³⁵ Table 4.5 shows the effect of one hundred additional military items acquired by police departments on the extensive margin decision to not disclose civilians killed by police.³⁶ The left-hand variable is an indicator for whether the department had a higher tally of homicides by police in Fatal Encounters than those reported in UCR each year. Each column in Table 4.5 uses the total stock of items acquired according to the category of items: total items transferred through the 1033 program in the first column, the total number of weapons and ammunition in the column 2, vehicles in column 3, optics in column 4, clothing items in column 5, and all others in the sixth column. The last two columns distinguish between rifles and ammunition, respectively.

Table 4.5 reveals that the probability of not disclosing the total number of homicides by police increases by 0.5 percentage points for every one hundred additional items transferred to police departments. However, the effect is larger for tactical gear than non-tactical equipment. One hundred weapons and optics acquired by the police department increase the likelihood of under-reporting by 9.4 pp. and 23.9 pp., respectively. The large response in optical gear may indicate that this hardware helps the deployment officers in night operations and targeted lethal actions. We also find a remarkably high effect for vehicles but the F-statistic shows a weak first stage.³⁷ In all other instances, the instrument is strong, and the first stage coefficient has the expected sign: distance to a military storage facility increases the

³⁵We observe this association at the national level as well. Appendix Figure 4.2 documents the total difference between civilians killed by police in Fatal Encounters and in UCR. The gap between reported statistics has a large uptick in 2010, while the transfers of military equipment was increasing through the 1033 program.

³⁶Appendix Table C.3 reports results using standardized stock of items.

³⁷The weak first stage for vehicles is also seen in previous research (Masera, 2021b).

costs of acquiring gear, decreasing take-up. Non-tactical gear such as clothes and others have a modest impact on the decision to disclose statistics. Rifles have large effects of 26.3 pp., and 14.8 pp. for one hundred additional units of ammunition.³⁸

The heterogeneous effects across the type of items support the hypothesis that tactical military gear affects behavioral choices, while non-tactical gear has little effect in LEAs behavioral decisions. Tactical items have the largest coefficients, with rifles, optics, ammunition, and weapons increasing the likelihood of under-reports above 10 pp. The effect is moderate, below 3 pp., for non-tactical items such as clothes and others. The differences in magnitudes suggest that police departments respond to some transfers more than others: tactical gear makes departments more likely to under-report the tally of homicides by police.

We take this as evidence that military hardware does not only affect killings by police, but the decision to disclose these homicides. First, transfers consistently have a positive and precise effect on the likelihood to under-report homicides by police and the intensive margin count. Second, the largest effects are concentrated in tactical items, which have been shown to increase violent engagement with civilians (Masera, 2021a,b; Gunderson et al., 2021). As expected, non-tactical items have smaller or insignificant statistical associations. Furthermore, the results are stable across various specifications in the first stage equation.³⁹

³⁸The results are consistent at the intensive margin. Appendix Table C.4 presents the relationship between one hundred additional items transferred to police departments and the number of under-reported homicides by police. The result is also consistent in a lin-log functional form in Appendix C.5.

³⁹See Appendix Table C.8 which shows results with eight different first-stage specifications. In this test we leverage the different variations in the instrument that have been used in the literature. We first test different nonlinear interactions with distance to storage facilities. Next, we exploit time variation in the total absolute change in boots on the ground instead of the normalized index. We also show the result is robust to alternative time variation, such as cumulative military expenditure in constant prices (Bove and Gavrilova, 2017) and the cumulative increase in the national stock of items (Harris et al., 2017). These specifications define the instrument in cumulative terms, matching the data generation process in the stock of items. In addition, Appendix Table C.6 shows that results remain consistent when using endogenous OAV of items instead of quantities. In the latter case, vehicles have a weaker effect than weapons and optics, but the instrument is strong.

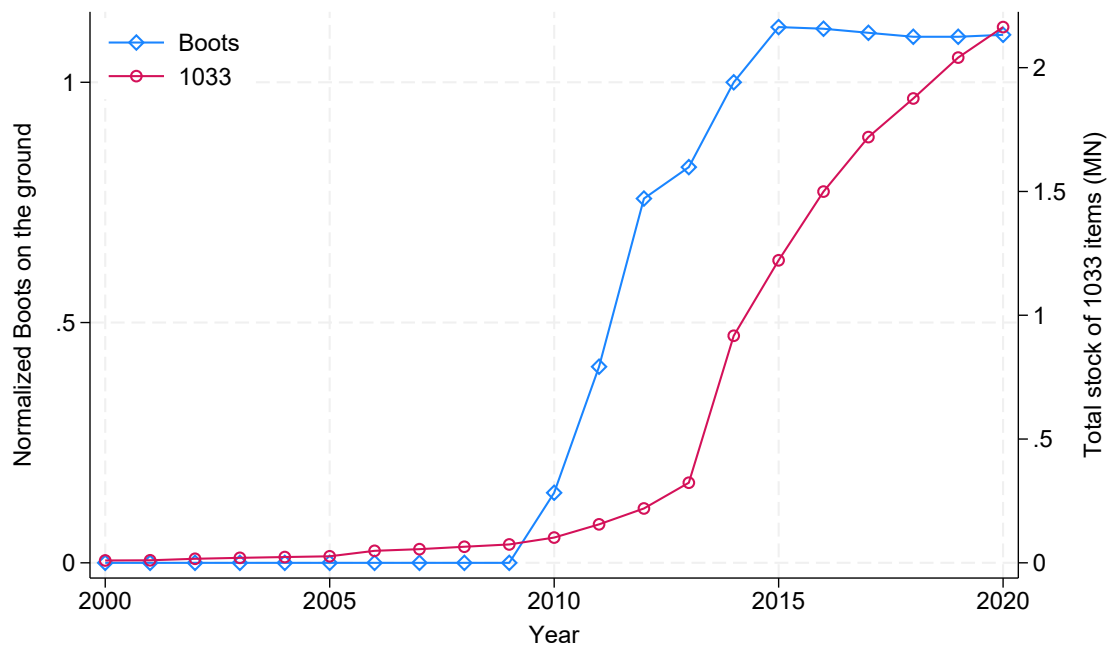


Figure 4.6: Boots on the ground and the national stock of items.

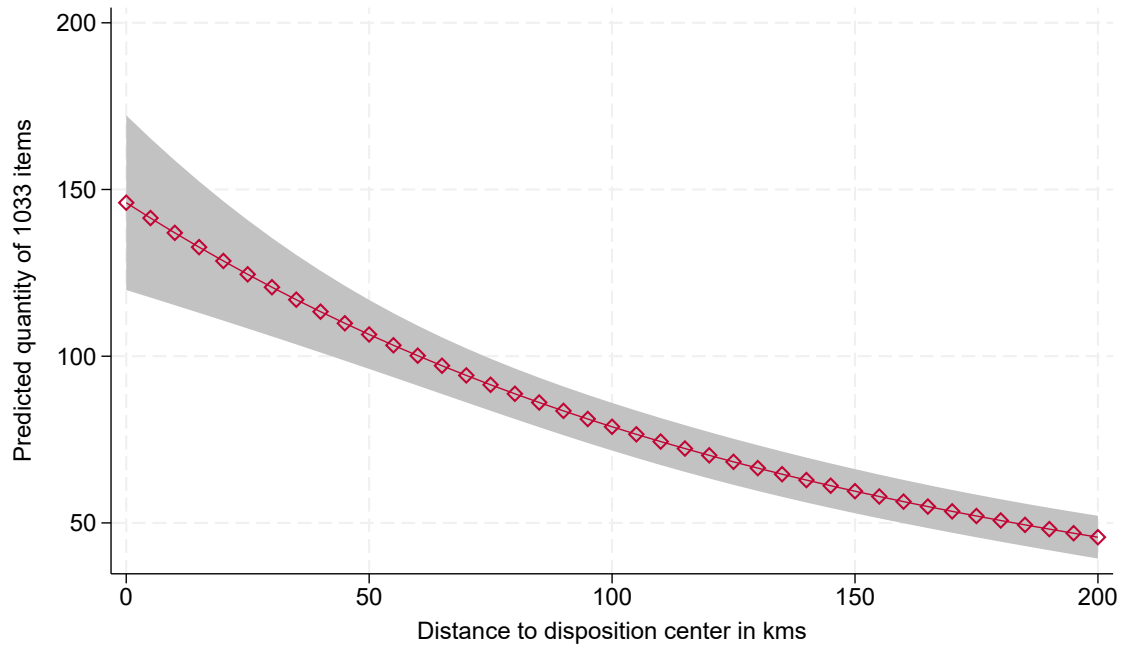


Figure 4.7: Distance to the closest disposition center and predicted quantity of items.

Notes: The figure displays the predicted quantity of items acquired by police departments as a non-linear function of the distance to disposition center, *ceteris paribus*. The regression controls for the interaction between normalized boots on the ground and distance to the closest military base, year fixed effects, and agency fixed effects. Predicted values are estimated at the mean covariates. 95 percent confidence intervals in grey.

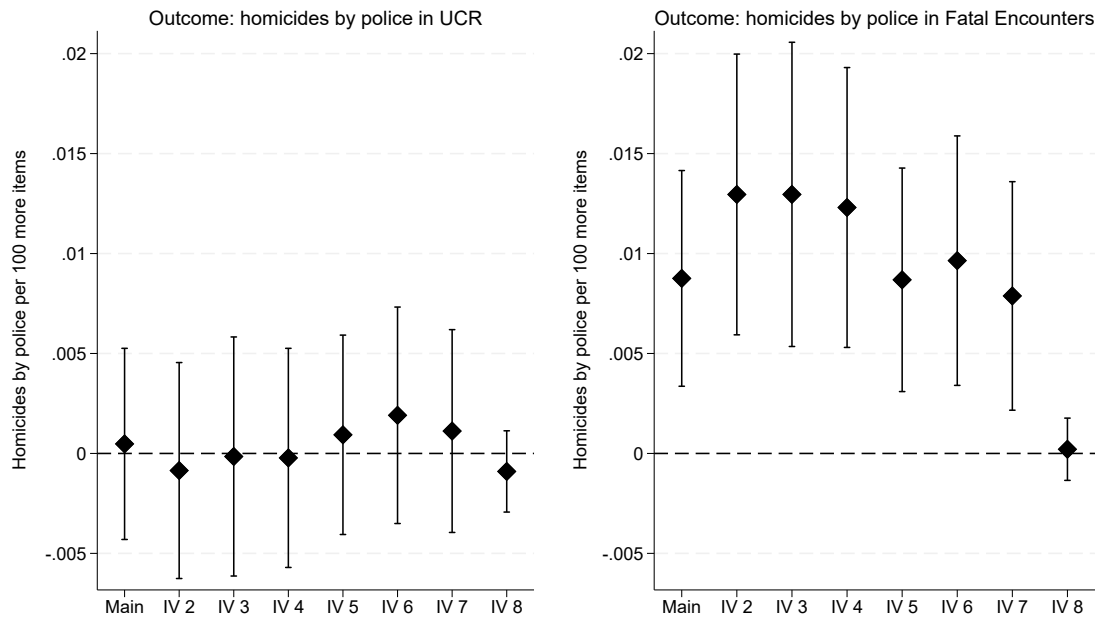


Figure 4.8: The effect of military items on homicides using different instruments and datasets.

Notes: The table shows the effect of one hundred additional items on the number of civilians killed by police in two datasets (UCR and Fatal Encounters). Each estimate uses a different specification in the first stage: **Main model:** the instrument uses time variation in normalized boots in the ground (EQP) interacted with distance to a DLA storage facility and interacted with the distance to the closest military base (Masera, 2021a); **Model 2:** the first stage includes time variation in normalized boots in the ground (EQP) interacted with distance to a DLA storage facility; **Model 3:** the first stage includes time variation in EQP interacted with distance to the closest DLA storage facility and a second interaction between EQP and quadratic distance to the closest DLA facility; **Model 4:** the first stage includes time variation in EQP interacted with distance to the closest DLA storage facility, an interaction between EQP and quadratic distance, and an interaction between EQP and cubic distance; **Model 5:** the first stage includes time variation in EQP interacted with distance to a DLA storage facility, quadratic distance, and cubic distance, and an interaction between EQP and distance to the closest military base; **Model 6:** the first stage includes time variation in the cumulative absolute variation in boots on the ground interacted with distance to the closest DLA storage facility and distance to the closest military base; **Model 7:** the first stage includes time variation in cumulative real military expenditure, similar to Bove and Gavrilova (2017), interacted with distance to the closest DLA storage facility and distance to the closest military base; **Model 8:** the first stage includes time variation in cumulative items in the nation, as in Harris et al. (2017), interacted with distance to the closest DLA storage facility and distance to the closest military base. The regression controls for population size, labor force, unemployment, per capita personal income, the police department's total budget, the number of full time sworn officers, the total number of employees, year fixed effects, and agency fixed effects. Standard errors are clustered at the commuting zone by year. Table results are in the Appendix Table C.2.

Table 4.5: The effect of military items on the probability of under-reporting.

	Total	Weapons	Vehicles	Optics	Clothes	Others	Rifles	Ammunition
Items	0.005*** (0.000)	0.094*** (0.000)	6.040*** (0.117)	0.239*** (0.000)	0.029*** (0.000)	0.005** (0.000)	0.263*** (0.000)	0.148*** (0.000)
Obs	235069	235069	235069	235069	235069	235069	235069	235069
F-statistic	40.27	46.38	6.61	44.76	42.21	34.83	46.24	51.66
Eqp*Storage	-0.455	-0.032	-0.000	-0.011	-0.108	-0.303	-0.011	-0.021
Eqp*Mil Base	0.545	-0.025	-0.001	-0.018	0.001	0.587	-0.013	-0.005

Notes: The table shows the effect of one hundred additional items on the probability of a positive difference between homicides by police reported in Fatal Encounters and in UCR. The F-statistic is the Kleibergen-Paap rk Wald F statistic. The rows display the first-stage coefficient of normalized boots on the ground interacted with the distance to DLA storage, and the interaction with distance to a military base. The regression controls for population size, labor force, unemployment, per capita personal income, the police department's total budget, the number of full-time sworn officers, the total number of employees, year fixed effects, and agency fixed effects. Clustered standard errors are computed at the commuting zone and year interaction.

In addition, we explore sociodemographic heterogeneities in the likelihood of disclosing civilian killings by police. Figure 4.9 illustrates the coefficients from 20 regressions. In each regression, we define under-reporting using the number of civilians recorded as killed by police across specific races or sex in each dataset. We estimate the program’s effects on the under-reporting of homicides across three racial categories (white, black, and non-white civilians) and across sex. The first five estimates in Figure 4.9 are computed using the stock of all military items transferred to police departments. The second set of estimates is limited to weapons and ammunition, the third on optics, and the last one groups together tactical items such as weapons, ammunition, and optics (Masera, 2021a).

We find little evidence of heterogeneities when aggregating tactical and non-tactical items, as coefficients are very small. One hundred additional weapons acquired by a police department increase the probability of not reporting the homicide of white and black civilians. The treatment effect is larger for the under-reporting of white civilians but there are no statistical differences across groups, as the confidence intervals overlap. The pattern remains in the acquisition of optics and tactical gear. The evidence contrasts with the unconditional aggregates in the killings of civilians, where the gap between Fatal Encounters and UCR is most predominant for non-white civilians. Our results appear to indicate that, while under-reported homicides are more frequent for racial minorities, the effect of the program has no racial heterogeneities.

Table 4.6 displays results in a sample restricted to periods when a police officer killed a civilian according to the Fatal Encounters data. We first explore the impact of military gear on the probability that the agency responds to UCR during that period. One hundred additional items decrease the likelihood of participating in UCR by 0.4 percentage points less. The probability of responding is even smaller for LEAs that acquire weapons and ammunition (7 percentage points) and optics (28 percentage points). The bottom panel in Table 4.6 shows the statistical association between 1033 transfers and under-reporting, for those departments that respond to UCR in

periods when an agent killed a civilian. In this instance, we cannot identify any effect of military items. The evidence is inconclusive as sample restrictions decrease the validity of the instrument. However, these findings suggests, when police officers kill a civilian, agencies abstain from disclosing their statistics to UCR, instead of misrepresenting the tally.

4.4.4 Leading indicators of 1033 take-up

Figure 4.10 explores the behavior of police departments prior to acquiring military hardware through the 1033 program. The vertical axis represents whether police departments are at p th quantile in the stock of items per sworn officers in 2000. The horizontal axis represents the quantile at which police departments transition by the end of the sample.⁴⁰ Each quadrant displays the probability, weighted by population, that a police department under-reported homicides by police between 2000 and 2010, before the major expansion of the 1033 program that followed the withdrawal of troops from Afghanistan and Iraq. For example, police departments in the top left quadrant in Figure 4.10 have a 5.3 percent probability of under-reporting. These LEAs were at the bottom quantile of military items acquired in 2000 and remained at the bottom of the distribution in 2020. The bottom right corner, instead, shows the probability of under-reporting for police departments that were at the top of the distribution and remained in the fifth quantile in 2020. These police departments have a probability of 0.3 percent of not disclosing civilian killings.

The transitional matrix in Figure 4.10 illustrates the main evidence that LEAs had a pattern of under-reporting civilian killings prior to participating in the program. LEAs that were low-takers and remain low-takers exhibit a lower probability of under-reporting than those that were low-takers but acquired a substantial quantity of items, transitioning to the top of the distribution of items acquired. Instead, LEAs that will

⁴⁰The distribution of military items acquired by police departments is highly skewed at the right tail. Over 50 percent of police departments never acquire items. Therefore, we define quantiles using only positive values larger than zero. The distribution remains skewed with 0.4 items per sworn officer in the first quantile and 5 items per sworn officer in the top quantile.

become heavy 1033 participants in the future exhibit a larger probability of under-reporting. For example, the probability of under-reporting is 7.7 percent for a police department that was at the bottom of the distribution but transitions to the second quintile. While the probability is 11 percent for those departments at the bottom that acquire enough items to be at the top of the distribution in 2020.

We observe a similar pattern using an alternative measure of adverse police behavior, computing average assets seized in counties transitioning from the bottom quintiles of 1033 items to the top quintiles. Asset forfeiture is a controversial policy that allows LEAs to keep a fraction of assets seized in drug arrests with lower legal restrictions than criminal conviction. We use asset forfeitures data at the county level from [Baicker and Jacobson \(2007\)](#). In Appendix [C.0.4](#), we present evidence that departments with more prominent asset seizure practices tend to transition between the bottom to top quintiles in the distribution of items per sworn officer. Seized assets are an alternative measure of negative police culture, as financial incentives have behavioral responses by police departments increasing arrests to fund their activities ([Baicker and Jacobson, 2007](#)). Asset forfeiture policies aim at increasing anti-drug policing, such that seizures may be larger in HIDTA areas favored by the 1033 program. Therefore, these findings cannot be causally interpreted but is indicative of negative behavioral patterns in police departments that participate in the 1033 program.

We present two sets of parametric results to back up our findings. First, Appendix [C.0.5](#) includes a linear estimator documenting the correlation between lagged under-reports and the acquisition of military gear through the 1033 program. The results point that police departments with a history of under-reports are more likely to acquire additional units of vehicles and optics. However, we cannot reject the hypothesis that police departments that under-reported in UCR are those located in areas with growing violence or other unobservable characteristics.⁴¹

⁴¹Appendix Table [C.9](#) reports the regression coefficients on a sample with all police department. Table [C.10](#) reports the statistical association between lagged under-reporting and the stock of military items, conditional on the agencies having previously responded to UCR.

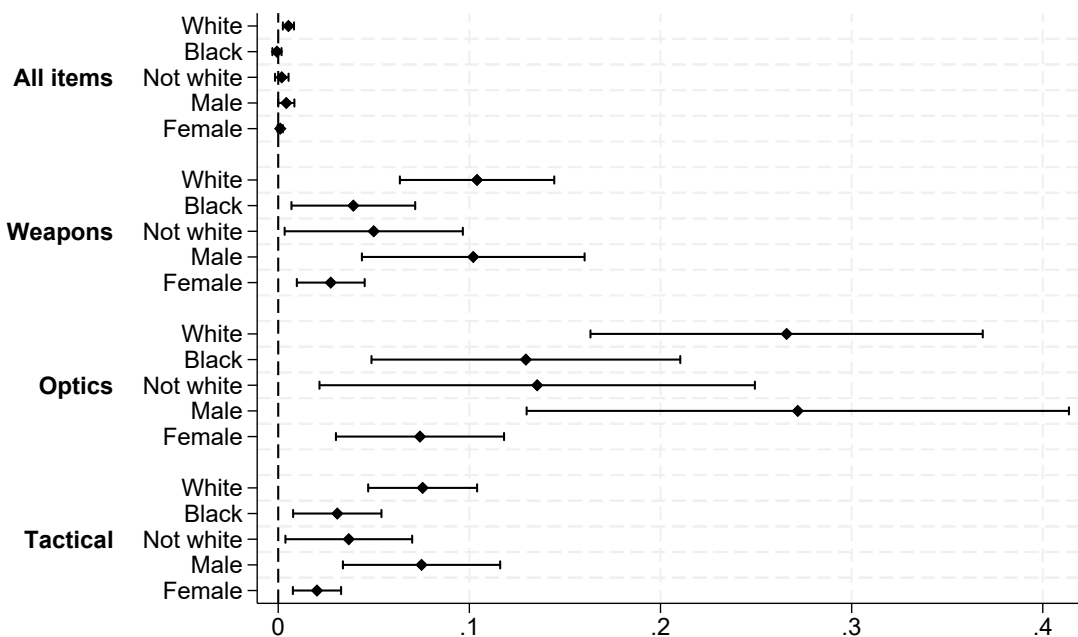


Figure 4.9: Racial and sex heterogeneities in the effect of under-reporting.

Notes: The figure shows the effect of one hundred additional items on the probability of a positive difference between homicides by police reported in Fatal Encounters and in UCR. Estimates are computed by instrumenting the stock of items with time variation in normalized boots in the ground (EQP) interacted with distance to a DLA storage facility, and interacted with the distance to the closest military base. Each coefficient is estimated with an individual regression on the difference in killings along sociodemographic characteristics (race and sex) in Fatal Encounters and UCR. All regressions controls for population size, labor force, unemployment, per capita personal income, the police department's total budget, the number of full-time sworn officers, the total number of employees, year fixed effects, and agency fixed effects. Standard errors are clustered at the commuting zone by year.

Table 4.6: Military items and under-reporting conditional on a civilian being killed.

	Total	Weapons	Vehicles	Optics	Clothes	Others
Probability of reporting to UCR						
Items	-0.004*** (0.001)	-0.070*** (0.022)	-2.010 (3.384)	-0.275** (0.117)	-0.017*** (0.006)	-0.005*** (0.002)
Obs	8399	8399	8399	8399	8399	8399
F-statistic	15.13	19.70	1.99	4.76	14.65	13.84
Eqp*Storage	-6.127	-0.358	0.000	-0.088	-1.395	-4.286
Eqp*Mil Base	5.855	0.133	-0.006	-0.023	1.077	4.674
Probability of under-reporting among current UCR respondents						
Items	0.003 (0.002)	0.082 (0.065)	-3.407 (3.255)	-0.257 (0.274)	0.013 (0.009)	0.004 (0.003)
Obs	5548	5548	5548	5548	5548	5548
F-statistic	5.96	3.41	2.81	1.83	7.72	5.56
Eqp*Storage	-6.239	-0.252	0.004	0.040	-1.646	-4.383
Eqp*Mil Base	8.131	0.131	-0.010	-0.132	1.652	6.491

Notes: The table shows the coefficients from two-step least squares of one hundred additional items on the probability of responding to UCR, and the probability of under-reporting conditional on being a current UCR respondent. The sample is limited to years in which the police department killed a civilian according to Fatal Encounters. The independent variable is the number of items acquired by the department, instrumented with time variation in normalized boots on the ground, the distance to a storage facility, and the distance to a military base. The F-statistic is the Kleibergen-Paap rk Wald F statistic. The rows display the first-stage coefficient of normalized boots on the ground interacted with the distance to DLA storage, and the interaction with distance to a military base. The regression controls for population size, labor force, unemployment, per capita personal income, the police department's total budget, the number of full-time sworn officers, the total number of employees, year fixed effects, and agency fixed effects. Clustered standard errors are computed at the commuting zone and year interaction.

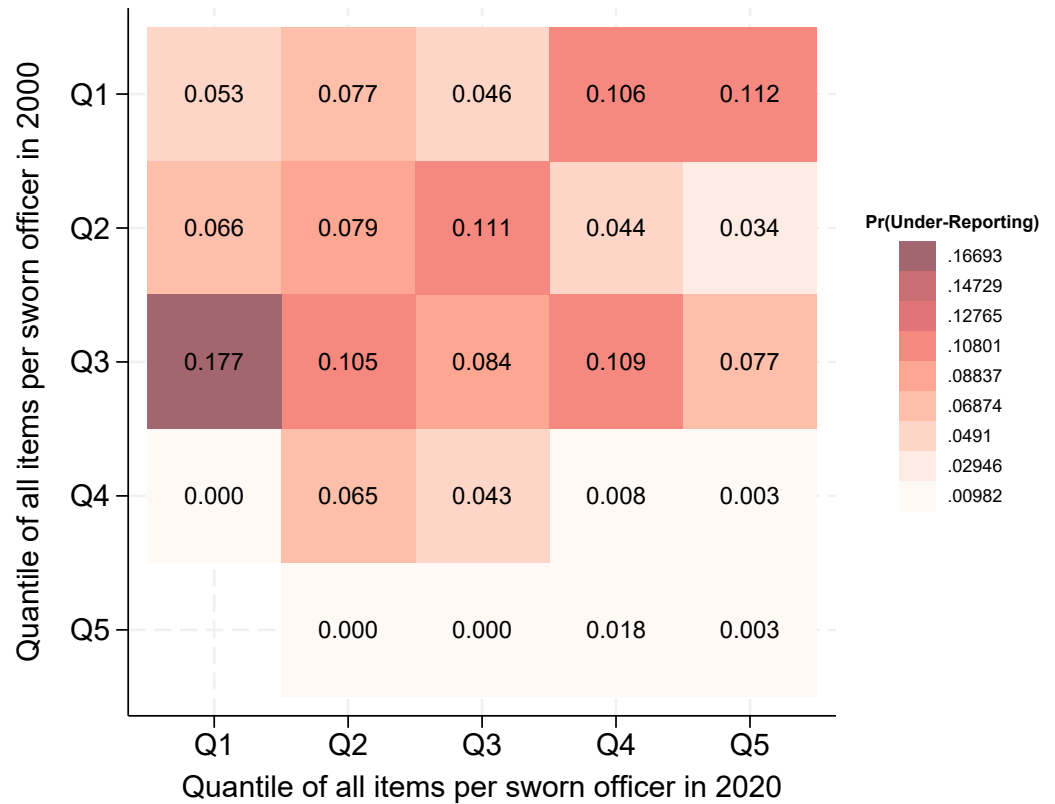


Figure 4.10: Probability of under-reports as police departments transition in the stock of items.

Notes: The figure plots the probability of under-reporting before 2010 as departments transition between the quintile of the distribution of 1033 items in an initial period (vertical axis) into a future period (horizontal axis). The probability of under-reporting is estimated with population weights.

In addition, twelve event studies are presented in Appendix C.0.6. Cohorts are defined according to the first shipment’s rollout (see Figure C.6) and, in six additional figures, according to the year of the largest 1033 shipment (see Appendix Figure C.0.6). Due to the heterogeneous timing in treatment, we compute estimates with a Two Way Fixed Effects (TWFE) model and the weighted dynamic effects estimator in Sun and Abraham (2021). In all cases, there is consistent evidence of pre-trends in under-reporting across both estimators. Furthermore, the TWFE estimator presents additional evidence of moral hazard in the shipment of weapons, vehicles, and optics: there is an increasing pattern of under-reporting after police departments acquire military gear.⁴² The parallel trend assumption indicates that police departments do not differ systematically in the trend of under-reporting behavior prior to acquiring gear.

4.5 Discussion

This paper explores the relationship between police departments acquiring military gear and their failure to report civilian homicide incidents to federal government databases. We reconcile mixed results in the literature and show that the most important difference between the studies that do and do not find linkages between receipts of tactical hardware and civilian killings by police is the data source for the outcome variable (Masera, 2021a; Gunderson et al., 2021; Bove and Gavrilova, 2017). We use instrumental variables methods that are well established in this literature, and match self-reported statistics of felons killed by police in the UCR dataset with third-party records of killings in the Fatal Encounters dataset at the agency-year level. In this paper, we analyze the timing of acquiring military items and the non-disclosure of homicides by police officers. We also scrutinize whether police departments exhibit

⁴²Sun and Abraham (2021)’s method shows more conservative standard errors, with suggestive evidence that vehicles and ammunition increase the probability of under-reporting in the long run. Overall, however, the event studies show no evidence of systematic pre-treatment differences among 1033 participants.

adverse behavior prior to participating in the 1033 program using both non-parametric and parametric estimators.

This paper lends support to both moral hazard and adverse selection hypotheses. The analysis via instrumental variable regressions reveals that the likelihood of agencies failing to disclose homicides escalates with an increase in military gear acquisition by police departments. Notably, the impact is most pronounced for tactical equipment, with a 9-percentage point increase associated with weapons and a 24-pp. rise for optics. In contrast, non-tactical gear, such as clothing and other items, exhibit low effects below 3-pp. The evidence indicates that military hardware influences the reporting behavior of LEAs. While non-disclosure is more prominent for the homicide of racial minorities, there is no evidence of racial heterogeneities in the program's treatment effects. Additionally, we find that police departments with a history of under-reporting and Civilian Asset Forfeiture are more inclined to build up their inventory of military equipment. The evidence suggests that police departments participating in the 1033 program may have previously engaged in selective or strategic non-disclosure of misconduct.

Previous studies examined whether violence involving police officers, both by and against them, rises following the acquisition of military hardware, yielding mixed results based on the data source. Research using self-reported data suggests that the incidence of civilian homicides by police is not influenced by the procurement of new military gear ([Harris et al., 2017](#); [Bove and Gavrilova, 2017](#)). Conversely, analysis based on data from third-party sources indicates a significant rise in civilian fatalities post-militarization of police departments ([Masera, 2021a](#)). Our study adds to this debate by underlining the intricate relationship between the acquisition of military hardware and the reporting behaviors of police departments. Specifically, we illuminate the dynamics of law enforcement disclosure and how it is influenced by and contribute to the process of militarization. The results support the theory that militarization influences organizational and cultural aspects of police departments

([Lawson Jr, 2019](#); [Delehanty et al., 2017](#)). However, identifying cultural changes is challenging and our results only offer suggestive evidence of adverse selection.

These findings have strong policy implications concerning federal oversight of law enforcement agencies that have been equipped with lethal military hardware. First, the minimal cost of entry and the absence of stringent approval processes have potentially enabled agencies with a history of harmful practices to exacerbate such behaviors. Second, the distribution of military equipment is shown to have adverse effects on police accountability. Despite the FBI's recognition of the urgent need to enhance the national statistics on law enforcement misconduct, federal authorities have struggled to motivate participation among a substantial portion of law enforcement agencies. This highlights a critical gap in efforts to monitor and improve law enforcement practices at the national level.

References

- Baicker, K. and Jacobson, M. (2007). Finders keepers: Forfeiture laws, policing incentives, and local budgets. *Journal of Public Economics*, 91(11-12):2113–2136.
- Banks, D., Couzens, L., Blanton, C., and Cribb, D. (2015). Arrest-related deaths program assessment. Technical report, Department of Justice, Office of Justice Programs, Bureau of Justice Statistics.
- Bove, V. and Gavrilova, E. (2017). Police Officer on the Frontline or a Soldier? The Effect of Police Militarization on Crime. *American Economic Journal: Economic Policy*, 9(3):1–18.
- Bruce, D. J., Carruthers, C. K., Harris, M. C., Murray, M. N., and Park, J. (2019). Do in-kind grants stick? The department of defense 1033 program and local government spending. *Journal of Urban Economics*, 112:111–121.
- Carriere, K. R. and Encinosa, W. (2017). The Risks of Operational Militarization: Increased Conflict Against Militarized Police. *Peace Economics, Peace Science and Public Policy*, 23(3):20170016.
- Connealy, N. T., Hall, A. R., and Fitch, C. H. (2023). 1033 Program data for studying Policing and militarization: an open-source database in response to the recommendations of Koslicki (2023). *Police Practice and Research*, pages 1–7.
- Delehanty, C., Mewhirter, J., Welch, R., and Wilks, J. (2017). Militarization and police violence: The case of the 1033 program. *Research & politics*, 4(2):2053168017712885.

- DLA (N/A). Law Enforcement Support Office (LESO). U.S. Defense Logistics Agency <https://www.dla.mil/Disposition-Services/Offers/Law-Enforcement/>. [Online; accessed 2023-12-14].
- Finch, B. K., Beck, A., Burghard, D. B., Johnson, R., Klinger, D., and Thomas, K. (2019). Using Crowd-Sourced Data to Explore Police-Related-Deaths in the United States (2000–2017): The Case of Fatal Encounters. *Open Health Data*, 6(1):1.
- Gunderson, A., Cohen, E., Schiff, K. J., Clark, T. S., Glynn, A. N., and Owens, M. L. (2021). Counterevidence of crime-reduction effects from federal grants of military equipment to local police. *Nature Human Behaviour*, 5(2):194–204.
- Harris, M. C., Park, J., Bruce, D. J., and Murray, M. N. (2017). Peacekeeping Force: Effects of Providing Tactical Equipment to Local Law Enforcement. *American Economic Journal: Economic Policy*, 9(3):291–313.
- Kaplan, J. (2021). Uniform Crime Reporting (UCR) Program Data: A Practitioner’s Guide. <https://ucrbook.com/>.
- Koslicki, W. (2022). Crucial considerations when using 1033 Program data as a measure of police militarization. *Police Practice and Research*, pages 1–7.
- Lawson Jr, E. (2019). Trends: Police militarization and the use of lethal force. *Political Research Quarterly*, 72(1):177–189.
- Masera, F. (2021a). Police safety, killings by the police, and the militarization of US law enforcement. *Journal of Urban Economics*, 124:103365.
- Masera, F. (2021b). Violent Crime and the Overmilitarization of US Policing. *The Journal of Law, Economics, and Organization*, 37(3):479–511.
- Mummolo, J. (2018). Militarization fails to enhance police safety or reduce crime but may harm police reputation. *Proceedings of the National Academy of Sciences*, 115(37):9181–9186.

- Ozkan, T., Worrall, J. L., and Zettler, H. (2018). Validating media-driven and crowdsourced police shooting data: a research note. *Journal of Crime and Justice*, 41(3):334–345.
- Ramey, D. M. and Steidley, T. (2018). Policing through subsidized firepower: An assessment of rational choice and minority threat explanations of police participation in the 1033 program. *Criminology*, 56(4):812–856.
- Steidley, T. and Ramey, D. M. (2019). Police militarization in the United States. *Sociology Compass*, 13(4):e12674.
- Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199.
- United States Bureau of Justice Statistics (2004). Law Enforcement Agency Identifiers Crosswalk [United States], 2000. <https://doi.org/10.3886/ICPSR04082.v1>.
- United States Bureau of Justice Statistics (2007). Law Enforcement Agency Identifiers Crosswalk [United States], 2005. <https://doi.org/10.3886/ICPSR04634.v1>.
- United States Bureau of Justice Statistics (2009). Bureau of Justice Statistics. Census of State and Local Law Enforcement Agencies (CSLLEA), 2000. <https://doi.org/10.3886/ICPSR03484.v4>.
- United States Bureau of Justice Statistics (2011a). Bureau of Justice Statistics. Census of State and Local Law Enforcement Agencies (CSLLEA), 2004. <https://doi.org/10.3886/ICPSR28001.v1>.
- United States Bureau of Justice Statistics (2011b). Bureau of Justice Statistics. Census of State and Local Law Enforcement Agencies (CSLLEA), 2008. <https://doi.org/10.3886/ICPSR27681.v1>.

United States Bureau of Justice Statistics (2018). Law Enforcement Agency Identifiers Crosswalk [United States], 2012. <https://doi.org/10.3886/ICPSR35158.v2>.

United States Bureau of Justice Statistics (2023). Bureau of Justice Statistics. Census of State and Local Law Enforcement Agencies (CSLLEA), 2018. <https://doi.org/10.3886/ICPSR38771.v1>.

Appendix A

A.1 Income and tax credits

Figure A.1 portrays the tax schedule of 2018 for single and married taxpayers. The horizontal axis is the adjusted gross income declared by a taxpayer and the vertical axis corresponds to the tax credit transfers the household is entitled to at each level of income. The schedule depends on the family characteristics. Solid lines are the earnings-to-transfers correspondences for single taxpayers, and dashed lines for married taxpayers filing jointly. The plot includes the tax schedule for taxpayers with no children, with one child, two children, and three children. The three regions are easily observed in the figure. First, wage subsidies in the phase-in of each schedule begin with the first dollar of earnings and ends in the income threshold where the maximum credit starts. The plateau region is where tax credits are a horizontal line that represents the maximum tax credit available for families. The last region acts as wage tax, the phase-out decreases the transfers linearly until the maximum eligible income.

Figure A.2 presents the distribution of cumulative EITC measurements during all childhood by cohort. This plot illustrates that children in later cohorts receive larger transfers. First, because of program expansions illustrated in Figure 2.1. However, it may also be attributed to shifting family characteristics. For instance, a child born during a later cohort is likely to have older parents or more siblings during their childhood. The measures for EITC exposure (left panel) and imputed EITC (right panel) reflect variations in family composition as much as variations in the tax credit schedules.¹ In the right panel, we observe that imputed transfers are skewed, likely because agents optimize in the right tail of the tax credit schedule. The imputed EITC distribution also shows that a proportion of the sample is not eligible to the

¹EITC exposure in the left panel in Figure A.2, closely resembles the PSID distribution in previous research Bastian and Micheltore (2018), although the earlier cohorts are a smaller share of the NLSY sample.

policy, receiving zero transfers when income is above the AGI threshold. As the EITC expanded, the proportion of noneligibles decreased across cohorts.

Figure A.3 displays the distribution of adult income by intensity of tax credit transfers. I divide the sample by whether the child was in the bottom quintile the cumulative EITC distribution (red), receiving the least amount of EITC transfers when they were between 0 and 18 years old, and in the top quintile when children were eligible to the largest cumulative tax credits during their childhood (blue). The left graph shows the adult income of children between ages 22 and 40, conditional on their exposure to the EITC. In general, the distribution overlays but there is a visible tail for children with the least exposure in their childhood. The distribution of income in the right panel is determined by the percentile of imputed EITC. The lowest quintile of the tax credit distribution are more likely to have higher levels of adult income. The figure illustrates in blue that the large fraction of the sample in the top quintile of the EITC distribution have mean adult income positively skewed near zero. Conversely, those in the bottom quintile of EITC transfers were less likely to experience poverty during their childhood. As a result, we observe a negative relationship in the distribution of income as tax credits increase.

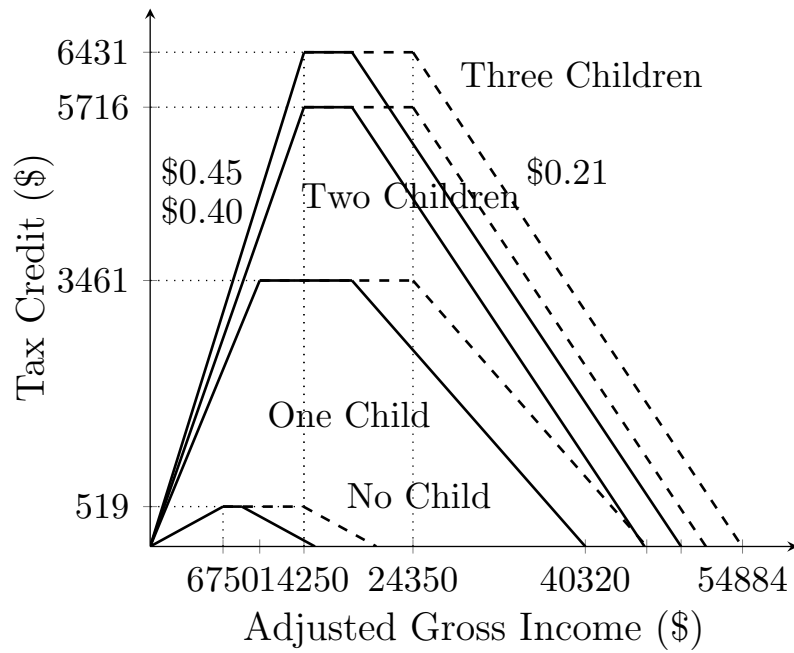


Figure A.1: EITC Schedule for Single and Married Households in 2018.

Notes: Solid lines represent the tax credit for a single head of household. Dashed lines represent the maximum tax credit for a married couple filing jointly. The phase-in and phaseout slopes for the Three-Children case are shown on the left and right, respectively.

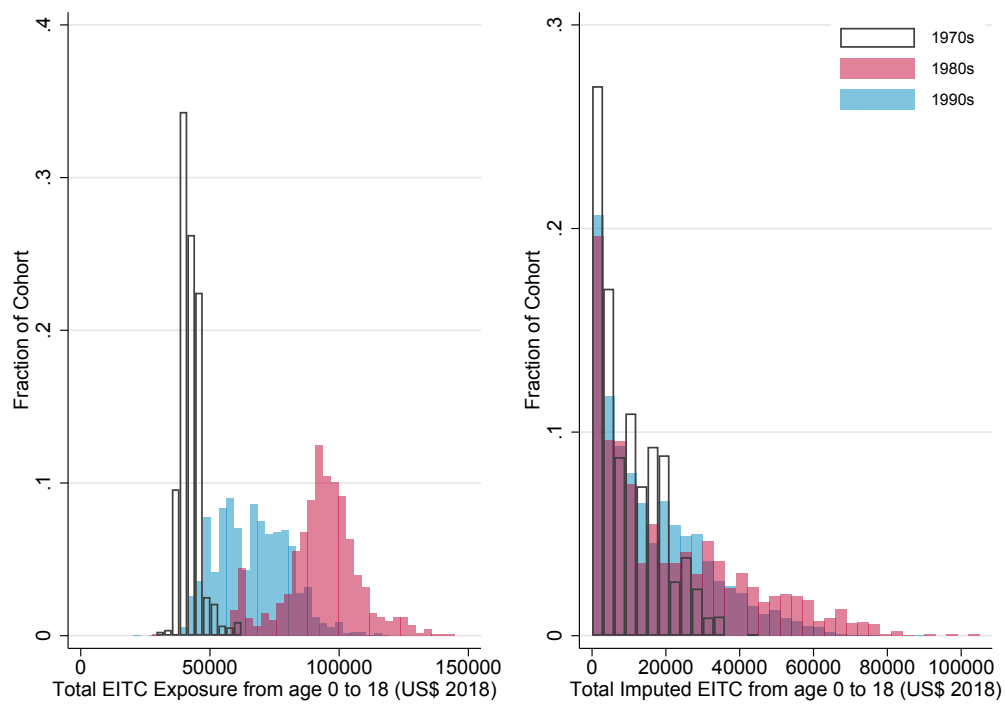


Figure A.2: Distribution of EITC exposure and imputed EITC, by cohort.

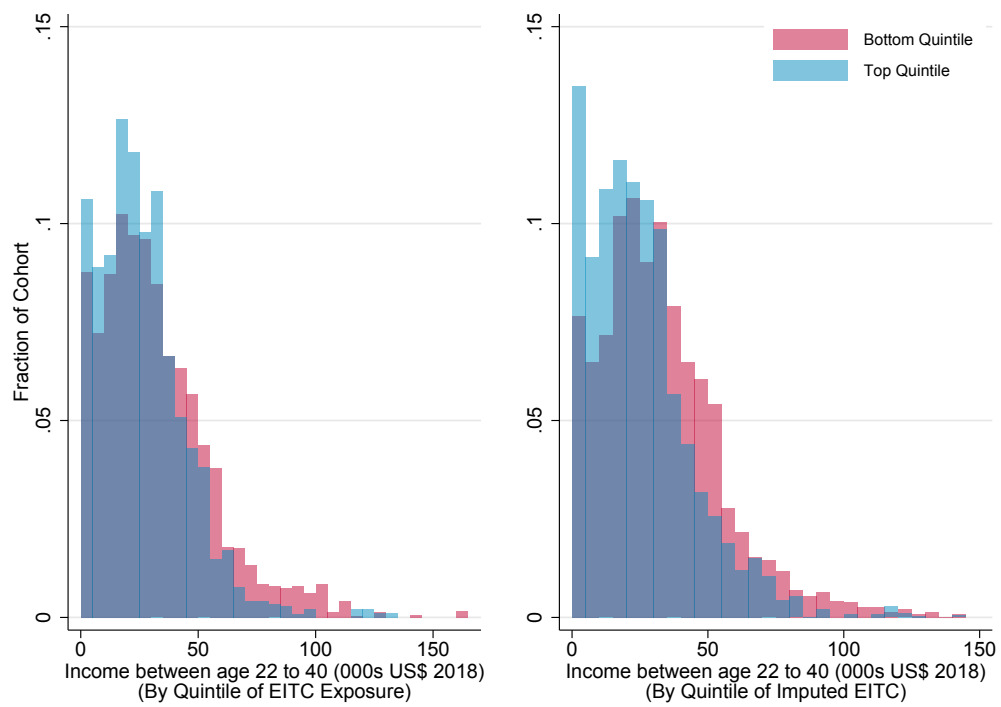


Figure A.3: Distribution of income by percentile EITC exposure and by percentile of imputed EITC.

A.2 Policy Variation

Table A.1: Moments of the distribution of tax credit variables.

	(1) Mean	(2) SD	(3) Skewness	(4) Kurtosis
EITC Exposure	2617.11	2216.62	0.47	1.84
Imputed EITC	445.76	1186.59	3.20	13.22
Policy Variation	73.00	199.34	4.07	23.83
Observations	370490			

Notes: The Table presents the moments of the distribution of three tax credit measurements in the full panel of mothers in the NLSY from 1979 to 2018. All values are calculated in real dollars in 2018. The first column shows the first moment, the second has the standard deviation, the third is the skewness and, lastly, the kurtosis of the distribution.

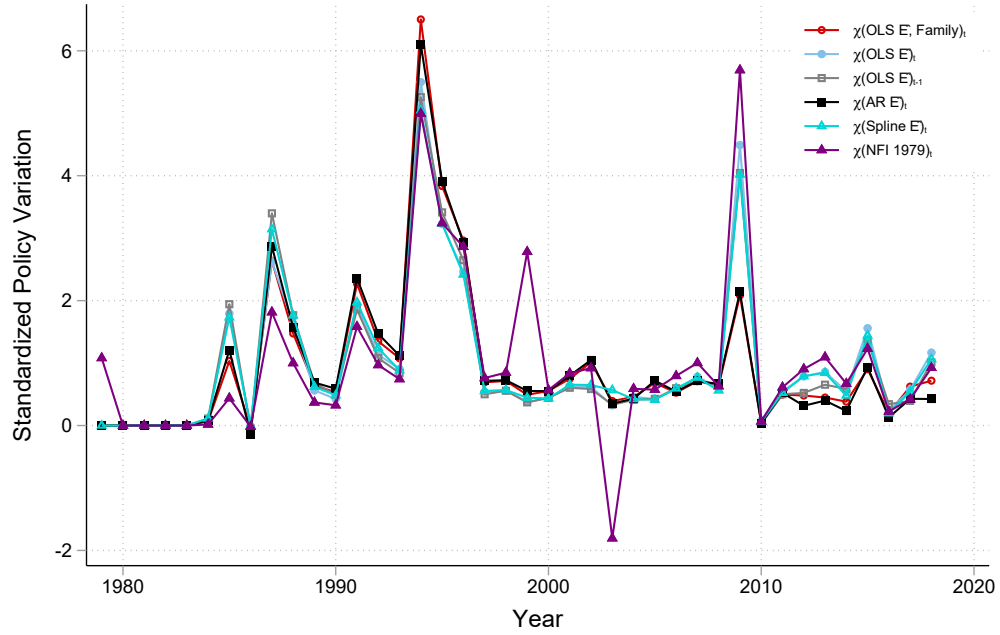


Figure A.4: Standardized policy variation across alternative constructions of the instrument.

Notes: The figure plots normalized measurements of policy variation across time. Each line represents a different input in calculating the simulated policy variation. First: the instrument is calculated with linearly predicted family income for the current period and family characteristics; Second: the instrument is calculated with linearly predicted family income for the current period and all mothers are assumed as single with 3 dependents; Third: the instrument is calculated with linearly predicted family income for the previous period and all mothers are assumed as single with 3 dependents; Fourth: the instrument is calculated with an autoregressive prediction of family income and all mothers are assumed as single with 3 dependents; Fifth: the instrument is calculated with fifth-order spline of lagged family income and all mothers are assumed as single with 3 dependents; Sixth: the instrument is calculated with net family income in 1979 adjusted for inflation, and all mothers are assumed as single with 3 dependents. Predicted family income is regressed with lagged wages, year fixed effects, and mother fixed effects.

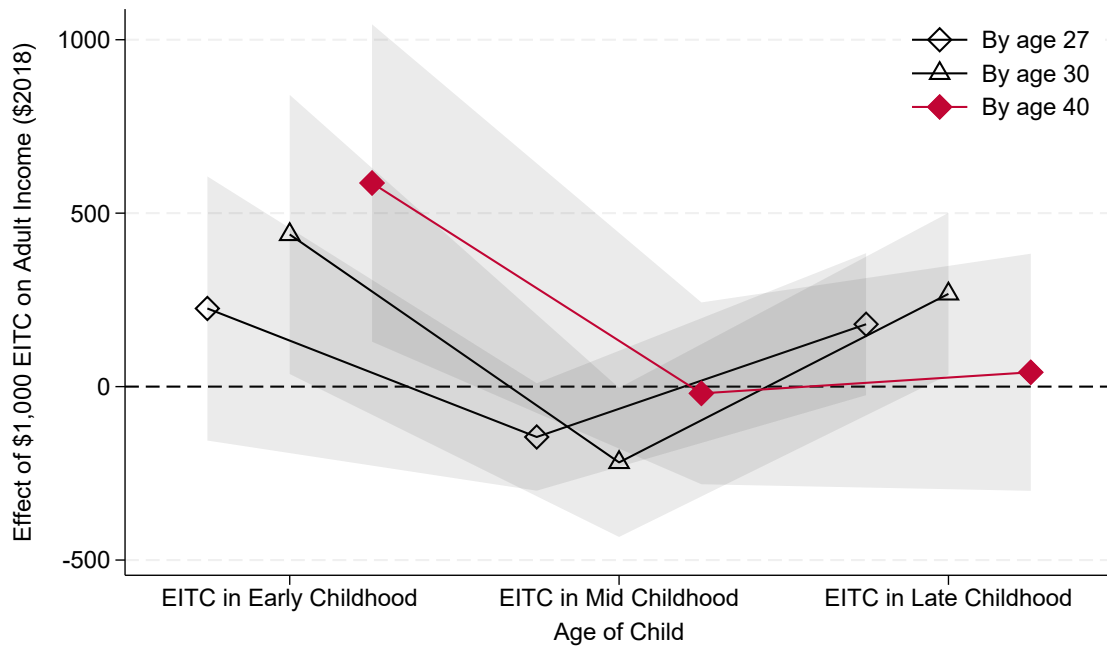


Figure A.5: 2SLS Coefficients of imputed earned income tax credit on the adult income by childhood stage.

Notes: The figure plots the 2SLS coefficients of \$1,000 in imputed tax credit transfers on adult income, instrumented with policy variation. The corresponding childhood stage in each coefficient is shown in the horizontal axis. The black diamonds are the coefficients on the adult income between age 22 and 27, the triangle represents adult income between age 22 and 30, and the red diamond is adult income up to age 40. 95% confidence interval clustered at the state level in grey. The regressions control for mother fixed effects, cohort fixed effects, quadratic and cubic age in 2018, indicators for black, Hispanic, female, whether the mother has been married by age 18, an indicator for whether the mother graduated high school, an indicator for whether the mother has some college education, and fixed effects for the number of siblings at age 18, the three-way interactions between race, ethnicity, and sex by number of siblings, birth year and state. The specifications accounts for state policy variables measured at age 18 (GDP per capita, unemployment rate, the top marginal tax rate, the minimum wage, tax revenues, higher education spending, and the maximum welfare benefits), state fixed effects, and year of birth fixed effects.

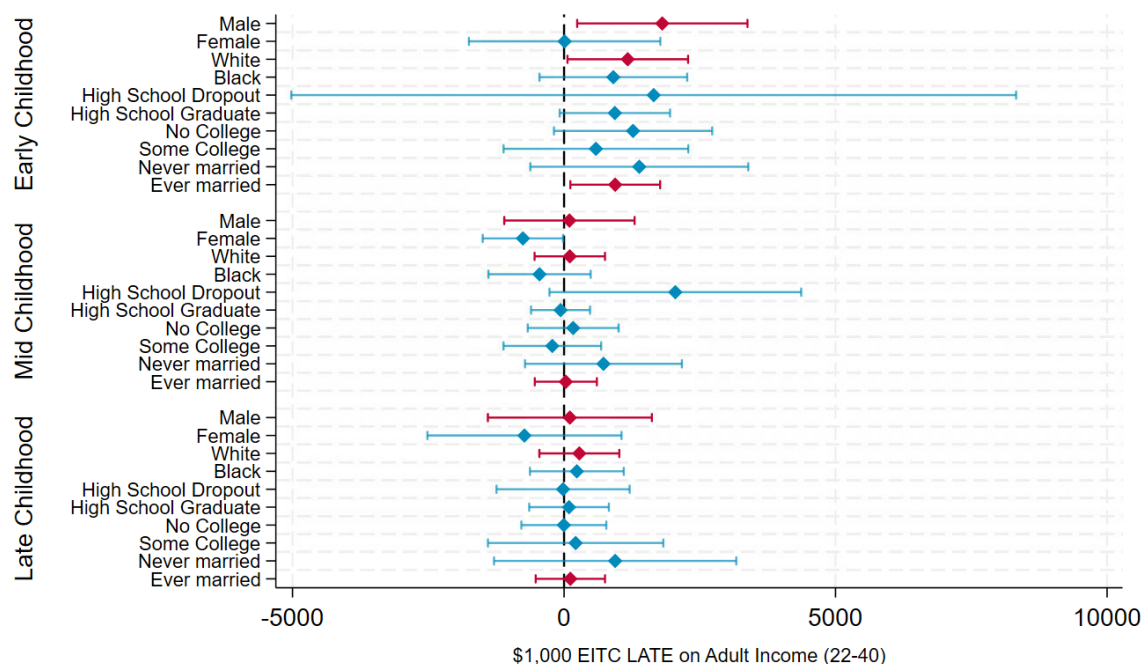


Figure A.6: 2SLS subgroup analyses of tax credits on adult income between ages 22 and 40.

Notes: The figure plots the 2SLS coefficients of \$1,000 in imputed tax credit transfers on adult income across subgroups. Imputed tax credits are instrumented with policy variation. The vertical axis is the size of coefficients and the 95 percent confidence interval clustered at the state level. The horizontal axis shows the sample restriction of relevance. Coefficients for early childhood tax credits are shown at the top, mid-childhood tax credits in the middle, and late childhood at the bottom. The regressions control for mother fixed effects, cohort fixed effects, quadratic and cubic age in 2018, indicators for black, Hispanic, female, whether the mother has been married by age 18, an indicator for whether the mother graduated high school, an indicator for whether the mother has some college education, and fixed effects for the number of siblings at age 18, the three-way interactions between race, ethnicity, and sex by number of siblings, birth year and state. The specifications accounts for state policy variables measured at age 18 (GDP per capita, unemployment rate, the top marginal tax rate, the minimum wage, tax revenues, higher education spending, and the maximum welfare benefits), state fixed effects, and year of birth fixed effects.

A.3 EITC Exposure in the CNLSY

Table A.2: OLS and 2SLS coefficients of earned income tax credit exposure on adult income between age 22 and 40.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Exposure in Early Childhood	-362.83 (223.33)	9859.02 (6853.05)	8358.59 (26770.78)	12708.83 (33848.69)	24913.40 (106140.29)	45320.91 (409317.71)	-7013.43 (9216.26)
Exposure in Mid Childhood	-45.54 (184.86)	2522.15 (2331.96)	-10849.29 (34688.22)	-3042.10 (14489.43)	-11153.12 (59936.59)	-39834.99 (397174.41)	204.80 (1529.01)
Exposure in Late Childhood	-366.21** (180.47)	-9498.54 (6085.30)	-27511.75 (68682.18)	-21194.71 (53146.21)	-34159.06 (146058.26)	-73554.32 (688447.37)	-3053.78 (2365.22)
Obs.	5253	5253	5253	5253	5253	5253	5253
R2	0.59	-1.98	-11.41	-7.41	-20.70	-96.81	-0.29
Estimator	OLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
KP Wald F		1.31	0.05	0.05	0.02	0.00	0.74
SW p-value (1)		0.05	0.65	0.66	0.80	0.90	0.06
SW p-value (2)		0.01	0.69	0.69	0.81	0.90	0.00
SW p-value (3)		0.05	0.67	0.69	0.80	0.90	0.00
Pred. Income Period		Current	Current	Lagged	Current	Current	1979
Pred. Income Function		Linear	Linear	Linear	Autoregressive	Spline	Fixed
IV Demographics		Yes	No	No	No	No	No

Notes: The Table presents coefficients of \$1,000 EITC exposure on mean adult income between ages 22 and 40. 2SLS coefficients are instrumented with cumulative policy variation in each childhood stage. The rows show the number of observations, adjusted R-squared, the Kleibergen-Paap F statistic (KP Wald F), and the Sanderson-Windmeijer (SW) multivariate F test corresponding to each first stage equation ([Sanderson and Windmeijer, 2016](#)). Column 1 is estimated with OLS, and Columns 2 to 7 are estimated with 2SLS and different policy variation instruments. Policy variation is calculated holding predicted family income constant for the current period of tax credits, the previous year (lagged), or with net family income of mothers in 1979. The linear predicted income function is regressed with lagged wages, year fixed effects, and mother fixed effects. The autoregressive function uses one income lag. The spline uses a fifth-order spline of lagged income. Fixed predicted income is the grandfather's income in 1979 adjusted for inflation. The instrument is computed with demographics (number of children, marital status, and state of residency), or assuming that all mothers are single with three children residing in the state where the child was born. Regressions control for mother fixed effects, cohort fixed effects, quadratic and cubic age in 2018, indicators for black, Hispanic, female, whether the mother has been married by age 18, an indicator for whether the mother graduated high school, an indicator for whether the mother has some college education, and fixed effects for the number of siblings at age 18, the three-way interactions between race, ethnicity, and sex by number of siblings, birth year and state. The specifications also accounts for state policy variables measured at age 18 (GDP per capita, unemployment rate, the top marginal tax rate, the minimum wage, tax revenues, higher education spending, and the maximum welfare benefits), state fixed effects, and year of birth fixed effects. Robust standard errors clustered at the state-level in parentheses. *** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$.

A.4 First Stage

The figure shows remarkably sensitivity of all first stage estimates to the inclusion of covariates. The figure also shows the dynamic correlation of transfers across stages of childhood. EITC transfers during early childhood are impacted only by policy variation during this stage, with no statistical association related to later variations as the cash transfers do not have expectations of future policy changes. Transfers during mid childhood do show a positive association both with early and mid childhood policy variation, reflecting the dynamics of policy changes. Policy variation during late childhood does appear positively associated with mid childhood EITC benefits. Lastly, transfers during later years are mostly affected by policy variation in this stage, with no association related to early childhood policy variation and little correlation with the previous stage.

I also show that there is no systematic correlation between policy variation and a set of covariates.

The results regarding state-policy variables presented in Table [A.4](#) support the general findings. The majority of the variables exhibit no significant relationship with policy variation in the EITC. Notably, the generosity of welfare policy, as measured by the maximum TANF benefit, shows no correlation with the EITC. This is consistent with recent research on the political economy of tax policy, suggesting that local and state tax changes are not driven by economic needs, state politics, institutional rules, neighbor competition, or demographics ([Robinson and Tazhitdinova, 2023](#)). The presence of a small but significant correlation between the EITC policy variation and minimum wages suggests that states with higher tax credits are more likely to have more generous minimum wage policies. Consequently, the lack of a systematic pattern observed in Table [A.4](#) mitigates concerns of endogenous policies.

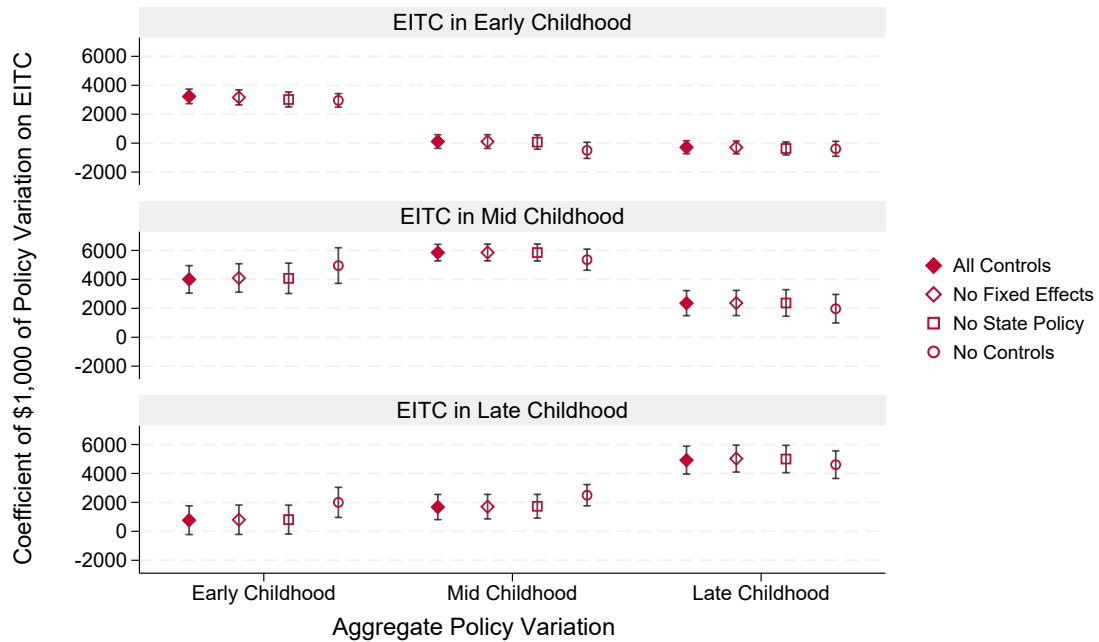


Figure A.7: Sensitivity of first stage estimates to covariates.

Notes: The figure plots the coefficients of \$1,000 in simulated policy variation on imputed tax credit transfers in each childhood stage. The vertical axis is the size of coefficients and the 95 percent confidence interval clustered at the state level. The horizontal axis is the childhood stage corresponding to policy variation. The regressions control for mother fixed effects, cohort fixed effects, quadratic and cubic age in 2018, indicators for black, Hispanic, female, whether the mother has been married by age 18, an indicator for whether the mother graduated high school, an indicator for whether the mother has some college education, and fixed effects for the number of siblings at age 18, the three-way interactions between race, ethnicity, and sex by number of siblings, birth year and state. The specifications accounts for state policy variables measured at age 18 (GDP per capita, unemployment rate, the top marginal tax rate, the minimum wage, tax revenues, higher education spending, and the maximum welfare benefits), state fixed effects, and year of birth fixed effects.

Table A.3: Correlation between policy variation and covariates across childhood stages.

	Policy Variation ($Z_{i(a)}$)			Obs.	R2 Adj.	F
	Early Childhood	Mid Childhood	Late Childhood			
Family Income (000s \$)	1.248 (1.123)	2.191 (1.475)	-1.722 (1.513)	5253	0.79	6.16
Female	0.012 (0.027)	0.016 (0.028)	0.014 (0.028)	5253	-0.02	0.12
Mother's Age in 2018	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	5253	1.00	.
GDP per capita	1.350*** (0.400)	0.477* (0.277)	-0.726** (0.283)	5253	0.89	35.33
State Min. Wage	0.459*** (0.123)	-0.111 (0.170)	-0.095 (0.106)	5253	0.81	5.04
Total Tax Revenue	1.315 (0.793)	-0.170 (0.475)	-1.046 (0.686)	5253	0.94	2.18
High Ed. Spending	6.490** (3.091)	9.696*** (2.345)	0.385 (2.274)	5253	0.79	8.61
Max. Welfare Benefit	-11.759 (18.949)	17.295 (14.538)	3.222 (10.640)	5253	0.80	10.95
Unemployment Rate	1.182*** (0.189)	-0.030 (0.144)	0.140 (0.132)	5253	0.37	29.89
Top Marginal Tax Rate	0.032 (0.043)	0.078*** (0.026)	0.133** (0.050)	5253	0.95	5.34

Notes: The Table shows coefficients for ten regressions of tax credit transfers on covariates. Dependent variables are in the rows. Each regression is computed with three independent variables: policy variation in early childhood, policy variation in mid childhood, and policy variation in late childhood. Regressions do not include controls. Robust standard errors in parentheses. *** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$.

Table A.4: Correlation between policy variation and covariates.

	(1) Age (2018)	(2) Female	(3) Mother's (YOB)	(4) GDP	(5) High Ed. Spending	(6) Min. Wage	(7) Tax Revenue	(8) Welfare Benefit	(9) Marginal Tax Rate
Policy EITC age 0-18	0.269*** (0.034)	-0.000 (0.003)	0.068** (0.032)	-0.147 (0.099)	-0.792 (0.900)	-0.079*** (0.015)	-0.296 (0.214)	-0.546 (2.194)	-0.027 (0.064)
Obs.	5393	5393	5393	5393	5393	5393	5393	5393	5393
R2 Adj.	0.01	-0.00	0.00	0.00	0.00	0.02	0.00	-0.00	0.00
F	62.41	0.02	4.52	2.20	0.77	28.64	1.92	0.06	0.18

Notes: The Table shows coefficients for nine regressions of total tax credit transfers on covariates. The dependent variables, in columns, are the age of children in 2018, their sex, the year of birth of mothers, the state's GDP per capita, spending in higher education, minimum wage, tax revenue, maximum welfare benefits for a family of three, and the top marginal tax rate. The rows are the coefficient for total tax credits across all childhood, the number of observations, adjusted R-squared, and the F-test value. Regressions do not include controls. Robust standard errors in parentheses. *** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$.

A.5 Sensitivity to covariates

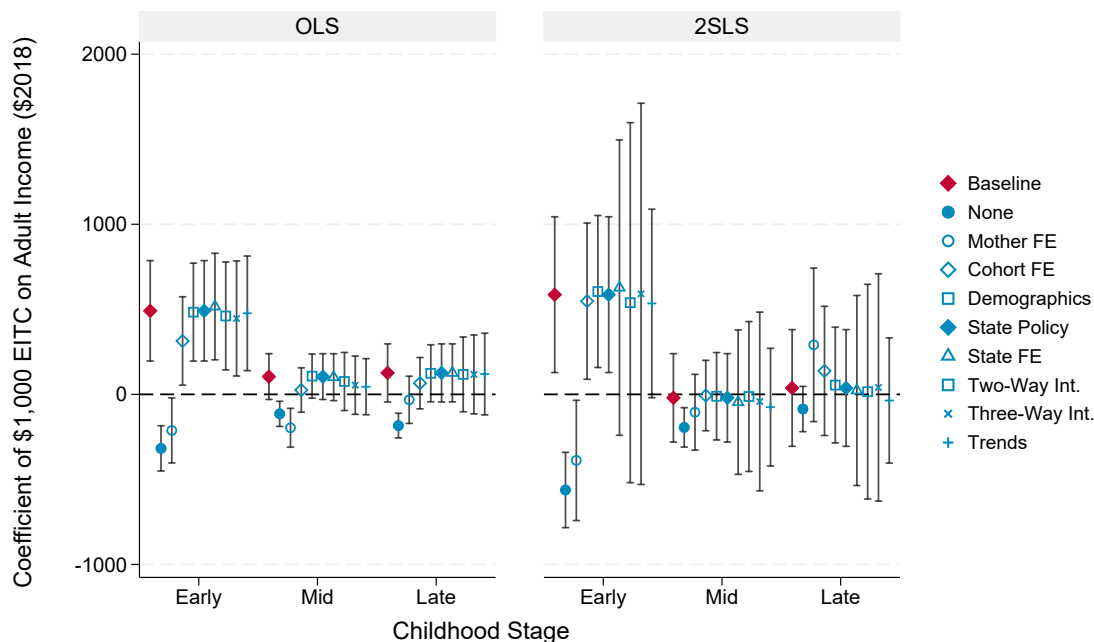


Figure A.8: Sensitivity to covariates on adult income.

Notes: The Figure plots coefficients for OLS (left) and 2SLS (right) regressions of \$1,000 in tax credits on adult income between ages 22 and 40. Estimates for all three childhood stages are computed across ten specifications. **Baseline:** the main specification; **None:** no covariates included; **Mother FE:** controls for mother fixed effects; **Cohort FE:** controls for cohort fixed effects; **Demographics:** controls for quadratic and cubic age in 2018, black, Hispanic, female, whether the mother has been married by age 18, an indicator for whether the mother graduated high school, an indicator for whether the mother has some college education, and fixed effects for the number of siblings at age 18; **State Policy:** controls for variables measured at age 18 (GDP per capita, unemployment rate, the top marginal tax rate, the minimum wage, tax revenues, higher education spending, and the maximum welfare benefits); **State FE:** controls for state fixed effects; **Two-Way Int.:** controls for two-way interactions between race, ethnicity, and sex by number of siblings, birth year and state; **Three-Way Int.:** controls for three-way interactions between race, ethnicity, and sex by number of siblings, birth year and state; **Trends:** controls for state by year time trends. The baseline and three-way specifications are equivalent. The interval is the 95% confidence interval estimated with standard errors clustered at the state-level.

A.6 Placebo Test

I conduct additional tests to examine the potential statistical association between policy variation in the EITC and a range of covariates. If the instrument exhibits a systematic statistical relationship with observable characteristics, it raises concerns regarding the independence of unobservable heterogeneities. The estimation excludes covariates from all regressions to focus specifically on the relationship between policy variation in the EITC and selected controls, including age, sex, the birth cohort of mothers, and state policy indicators. By examining the statistical associations between the EITC instrument and these covariates, we gain insights into the independence and validity of the instrumental variable. A lack of significant associations increases the confidence in the conditional independence of the instrument and enhances the reliability of the estimated causal effects of the EITC on the outcomes of interest.

The results are presented in Table [A.4](#). Overall, the table indicates a small but statistically significant association between policy variation in the EITC and the age of children. This is related to exposure to larger cash transfers in younger cohorts, likely attributable to the Welfare Reform. Additionally, there is a negative association between the year of birth of mothers and the EITC policy variation. These associations are expected and underscore the significance of including year of birth fixed effects and state-policy variables in the main regression models. These controls help mitigate the influence of omitted variables that may affect each cohort differently, particularly after the implementation of policy reforms. By accounting for these factors, the main regression models provide more robust estimates of the causal effects of the EITC on the outcomes of interest.

To test the robustness of the findings, the validity of the instrument may be tested by comparing it with a placebo. The placebo instrument is constructed with random family characteristics assigned to each child in 1985. Income follows a log normal distribution that simulates the moments observed in the data. Family income is then

adjusted for inflation. Marriage and fertility are assigned uniformly to all families. The intuition is that this test does not feature the advantages of capturing policy variation localized to a family. After running 100 simulations, regression results are shown in Figure [A.9](#). All coefficients are null with a large proportion of zeroes, and the instrument is weak according to the SW F statistic in every iteration. This robustness test reassures that the policy variation instrument is superior a random non-localized variation in the program. Furthermore, the test validates the strength of the instrument.

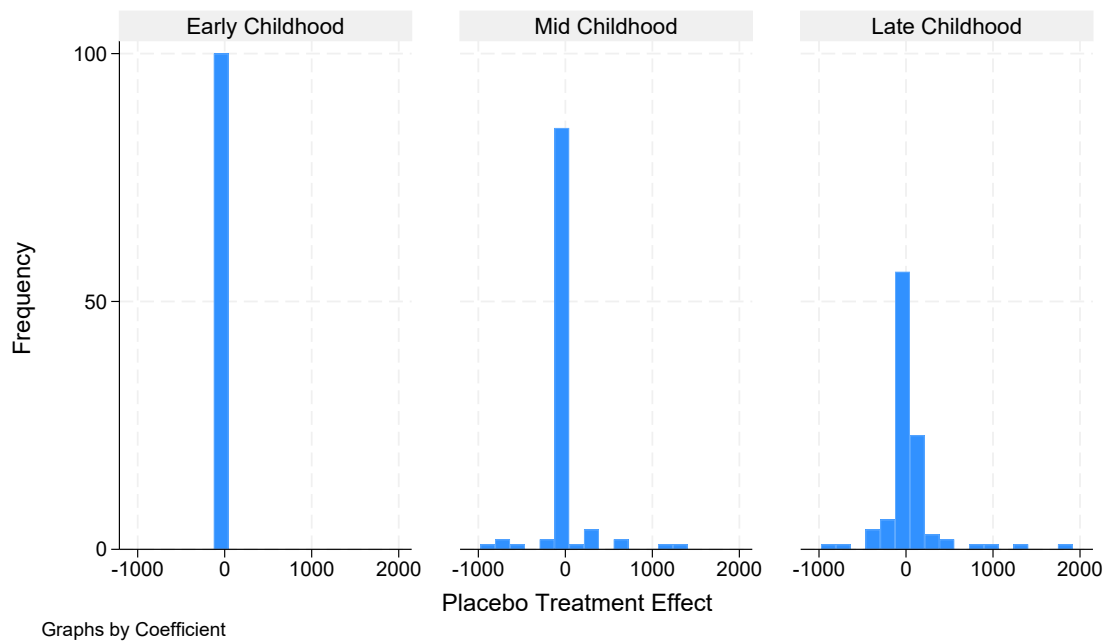


Figure A.9: Distribution of 2SLS coefficients with placebo instruments.

Notes: Each distribution shows an estimate of β_i in Equation 2.1. Imputed tax credits are instrumented with placebo instruments. Placebo instruments were estimated using 100 Monte Carlo Simulations of income and family characteristics. Family income is computed with a log-normal distribution that approximately matches the moments of observed income distribution in the NLSY. Family size and marital status are assigned with an uniform distribution.

A.7 Maternal Labor Supply

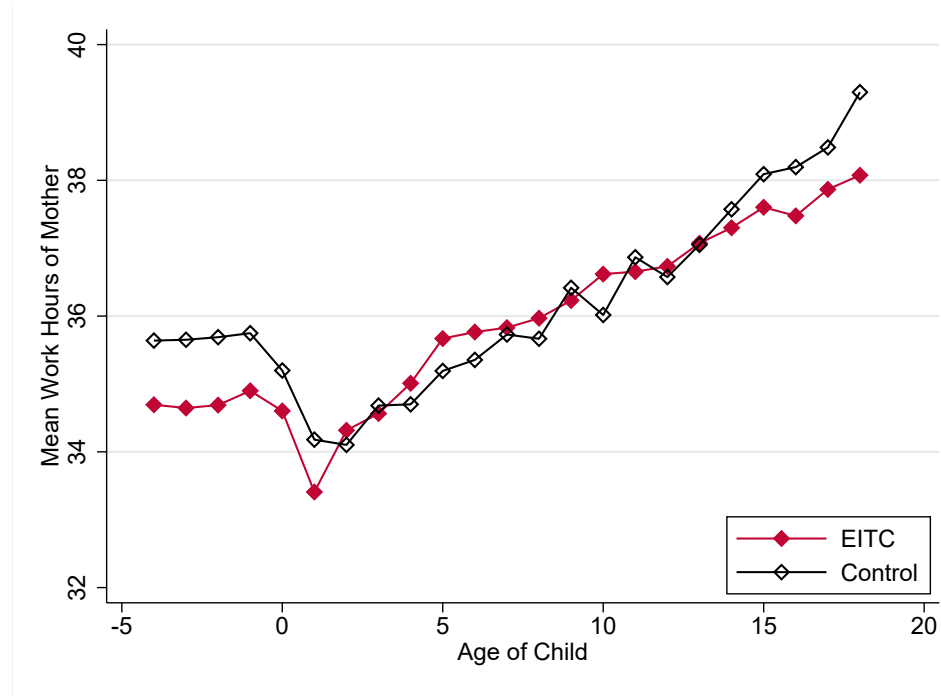


Figure A.10: Maternal Labor Supply after birth by EITC status.

Notes: The figure plots the unconditional weekly work hours of mothers and the age of their child. The first line includes the sample of households that are EITC eligible (red) while the second is the sample of non EITC eligible households (black).

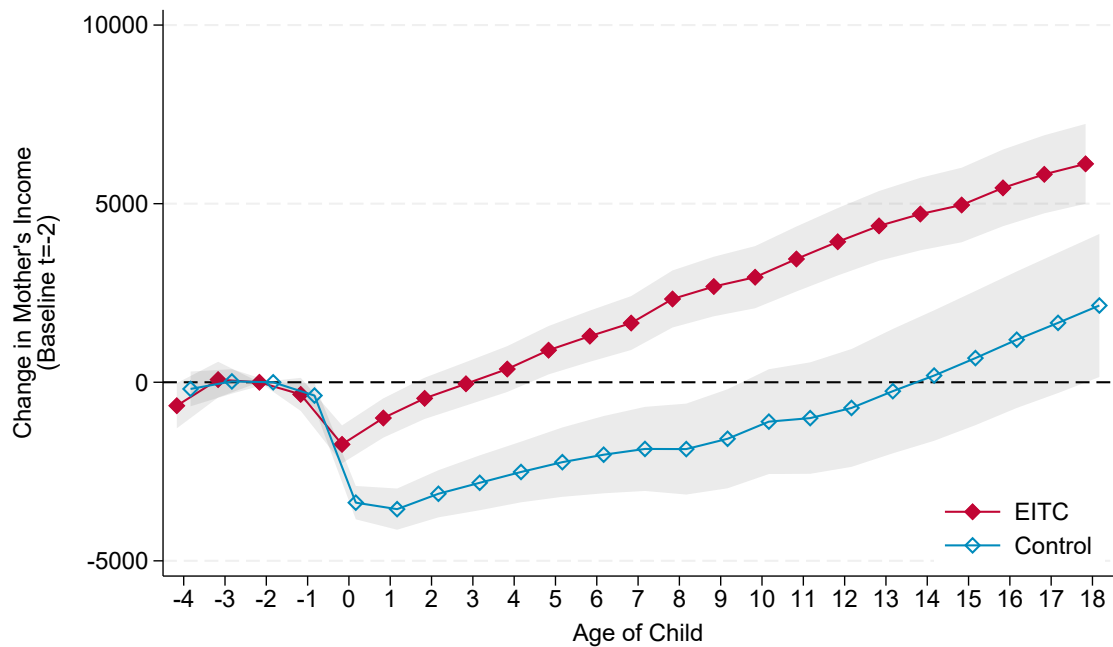


Figure A.11: Maternal income after birth by EITC status.

Notes: The figure plots the coefficients resulting from two two-way fixed effects regression. The coefficients capture the correlation between maternal income and the age of children. The first regression includes a sample of households that are EITC eligible (red) while the second is the sample of non eligible households (blue). The 95% confidence interval is constructed with robust standard errors clustered at the state-level.

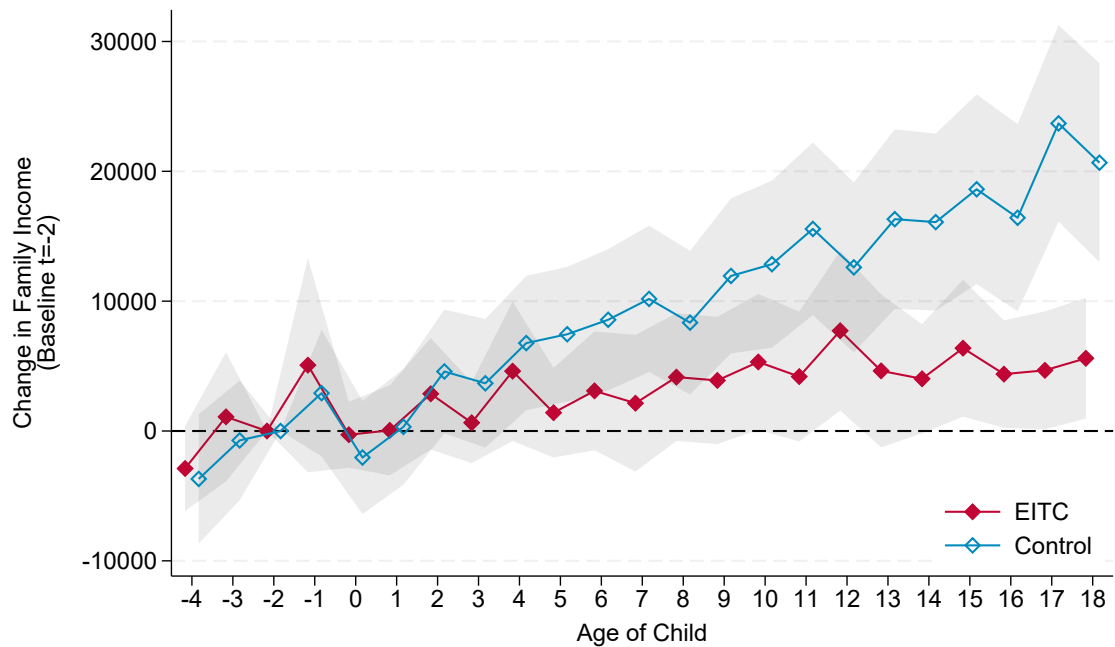


Figure A.12: Family income after birth by EITC status.

Notes: The figure plots the coefficients resulting from two two-way fixed effects regression. The coefficients capture the correlation between family income and the age of children. The first regression includes a sample of households that are EITC eligible (red) while the second is the sample of non eligible households (blue). The 95% confidence interval is constructed with robust standard errors clustered at the state-level.

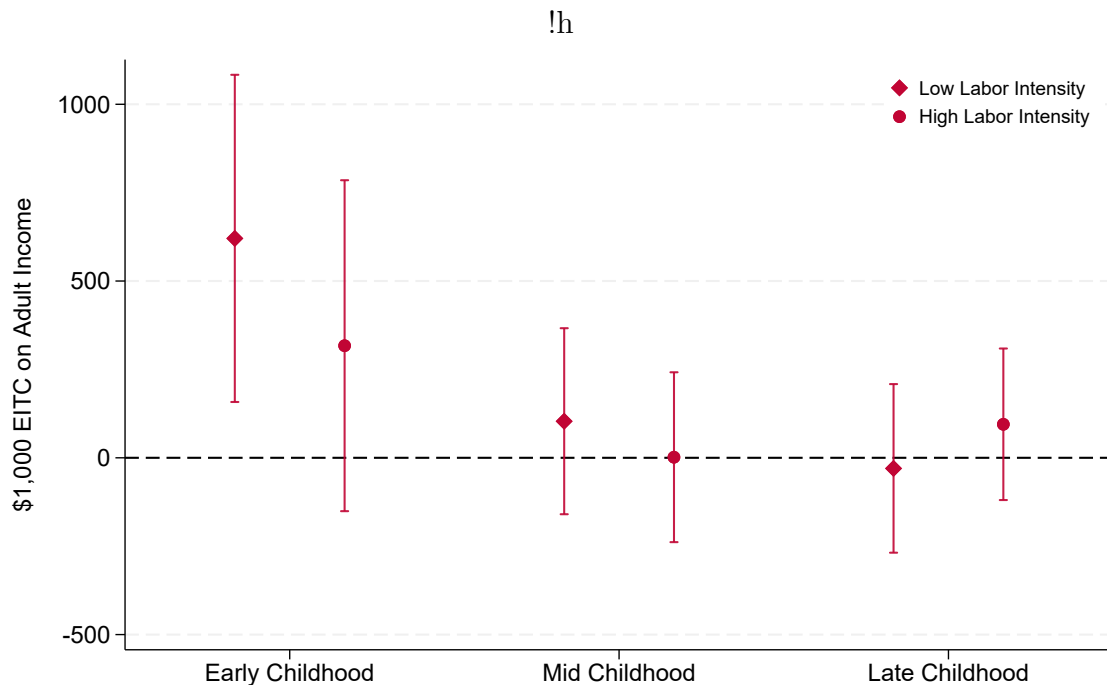


Figure A.13: Intergenerational effects of the EITC on adult income conditional on maternal labor intensity prior to childbearing.

Notes: The figure plots the OLS coefficients of \$1,000 in imputed tax credit transfers on adult income conditional on maternal work intensity. The vertical axis is the size of coefficients and the 95 percent confidence interval clustered at the state level. The horizontal axis is the childhood stage in which tax credits are received. Low labor intensity (red diamond) are children whose mothers are below the median in the distribution of predicted work hours prior to childbearing. High labor intensity (red circle) are children whose mothers are above the median in the distribution of predicted work hours prior to childbearing. The distribution of maternal labor intensity is computed before the child's birth year. The regressions control for mother fixed effects, cohort fixed effects, quadratic and cubic age in 2018, indicators for black, Hispanic, female, whether the mother has been married by age 18, an indicator for whether the mother graduated high school, an indicator for whether the mother has some college education, and fixed effects for the number of siblings at age 18, the three-way interactions between race, ethnicity, and sex by number of siblings, birth year and state. The specifications accounts for state policy variables measured at age 18 (GDP per capita, unemployment rate, the top marginal tax rate, the minimum wage, tax revenues, higher education spending, and the maximum welfare benefits), state fixed effects, and year of birth fixed effects.

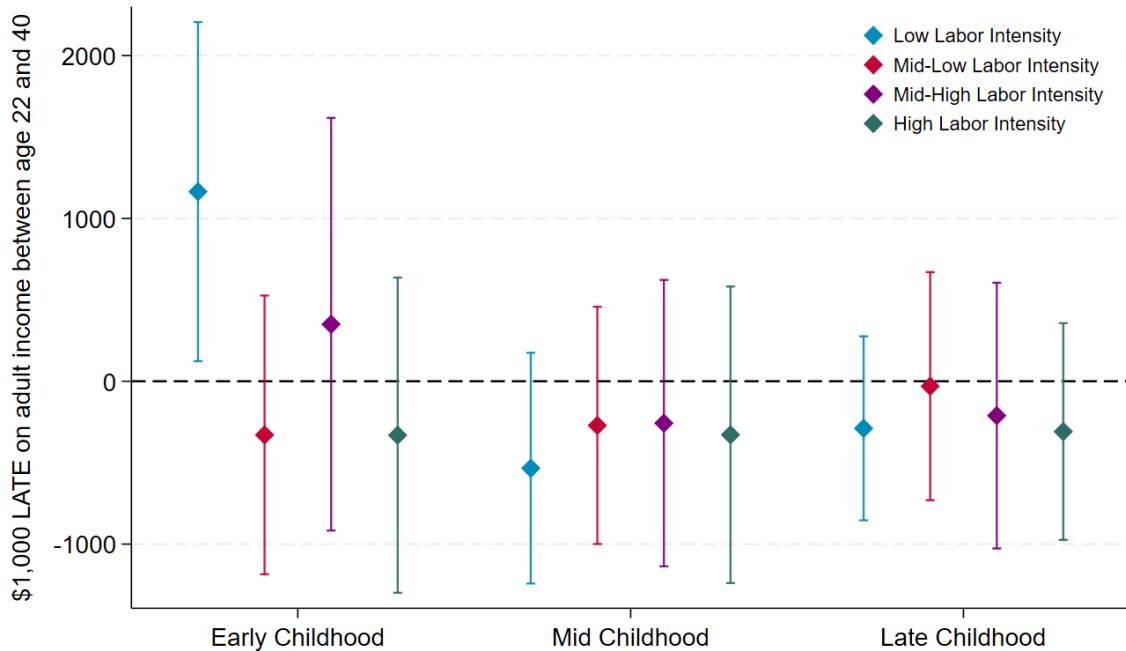


Figure A.14: LATE of the EITC on adult income conditional on maternal labor intensity during early childhood.

Notes: The figure plots the 2SLS coefficients of \$1,000 in imputed tax credits on adult income conditional on maternal work intensity. Imputed tax credits are instrumented with policy variation. The vertical axis is the size of coefficients and the 95 percent confidence interval clustered at the state level. The horizontal axis is the childhood stage in which tax credits are received. Maternal labor supply intensity is estimated when the child is between 0 and 5 years old. Maternal labor intensity is constructed at the quintiles of work hours such that there is low intensity (25th percentile), mid-low intensity (below 50th percentile), mid-high intensity (below the 75th percentile), and high intensity (above the 75th percentile). The regressions control for mother fixed effects, cohort fixed effects, quadratic and cubic age in 2018, indicators for black, Hispanic, female, whether the mother has been married by age 18, an indicator for whether the mother graduated high school, an indicator for whether the mother has some college education, and fixed effects for the number of siblings at age 18, the three-way interactions between race, ethnicity, and sex by number of siblings, birth year and state. The specifications accounts for state policy variables measured at age 18 (GDP per capita, unemployment rate, the top marginal tax rate, the minimum wage, tax revenues, higher education spending, and the maximum welfare benefits), state fixed effects, and year of birth fixed effects.

Appendix B

B.1 Industry Estimates

Table B.1: Average impact of the EITC on industry outcomes in logs.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of Establishments								
EITC Enactment	-0.032 (0.024)	-0.024 (0.021)	-0.041 (0.025)	-0.016 (0.027)	0.011 (0.029)	0.023 (0.030)	0.018 (0.031)	0.045 (0.053)
Employment								
EITC Enactment	-0.055*** (0.020)	-0.052** (0.020)	-0.063** (0.024)	-0.064** (0.030)	-0.024 (0.035)	-0.028 (0.036)	-0.041 (0.036)	-0.001 (0.058)
Real Weekly Wages								
EITC Enactment	-0.005 (0.008)	-0.010 (0.007)	-0.010 (0.008)	-0.014** (0.006)	-0.017** (0.007)	-0.012* (0.007)	-0.014** (0.007)	-0.005 (0.009)
Obs.	317367	317367	293155	252753	317367	317367	293155	252753
R2	0.90	0.90	0.90	0.90	0.90	0.91	0.91	0.91
Covariates	N	Y	Y	Y	N	Y	Y	Y
Sample	All	All	Refundable	Non-Refundable	All	All	Refundable	Non-Refundable
Exposure	N	Y	N	N	Y	Y	Y	Y

Notes: The table reports difference-in-difference and triple difference estimates of the EITC's impact on the log of the number of establishments, employment, and real weekly wages across industries. The treatment variable captures whether the state passed the EITC. Estimators with exposure use the interaction between the enactment of the tax credit and an indicator for EITC exposure above the 75th percentile of exposure. Estimates based on the enactment of a refundable (or non-refundable) EITC supplement exclude observations from states that ever enacted a non-refundable (refundable) program. All models include state fixed effects, year fixed effects, and industry fixed effects. Models with covariates control for the unemployment rate, state per capita GDP, minimum wage, state marginal tax rate, the percentage of female workers, average race and ethnicity, educational achievement, number of children, median family income, and the age of workers in the industry. Standard errors are clustered at the state level. $p < 0.01$ $p < 0.05$ $p < 0.1$.

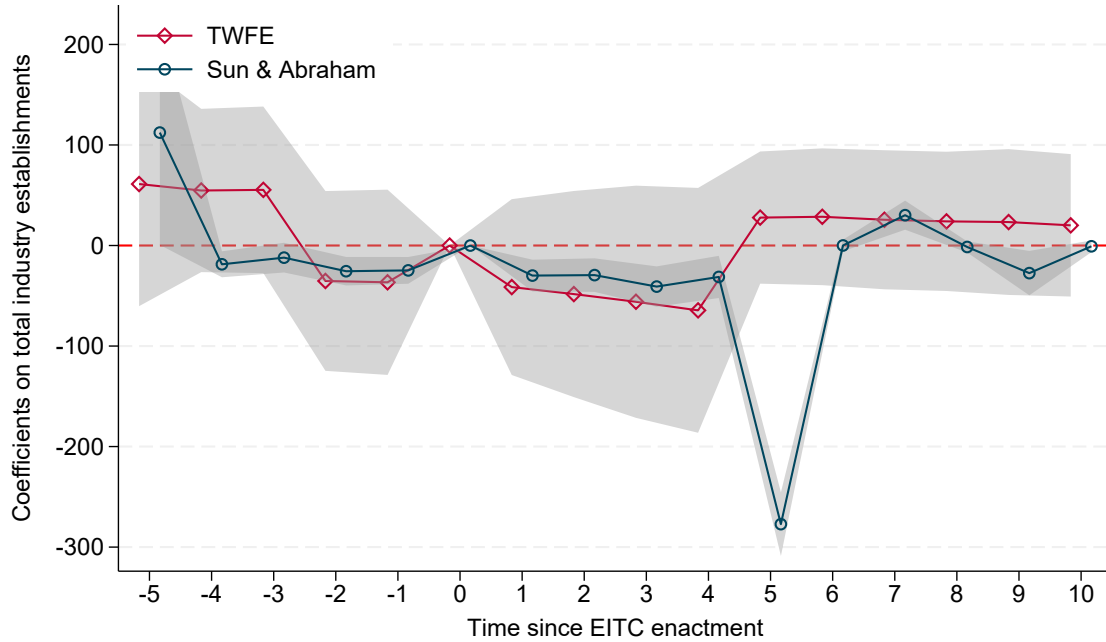


Figure B.1: Event study of EITC roll-out and establishments.

Notes: The figure presents standard event studies using the Two Way Fixed Effects (TWFE) estimator and the [Sun and Abraham \(2021\)](#) estimator. The outcome is the number of establishments in an industry. The treatment variable captures the interaction between the enactment of the EITC and an indicator for EITC exposure above the 75th percentile of exposure. Models control for occupation fixed effects, state fixed effects, year fixed effects, the unemployment rate, state per capita GDP, minimum wage, state marginal tax rate, the percent of female workers, average race and ethnicity, educational achievement, number of children, median family income, and the age of workers in the industry. The 95 percent confidence intervals are estimated with cluster-robust standard errors at the state-level.

B.2 Occupational Estimates

Table B.2: Estimates of the EITC on the log of occupational outcomes.

	(1) Employment	(2) Mean Wages	(3) Median Wages	(4) 10th Pct	(5) 25th Pct	(6) 75th Pct	(7) 90th Pct
EITC Enactment	-0.099*** (0.005)	-0.011*** (0.001)	-0.016*** (0.001)	-0.020*** (0.001)	-0.020*** (0.001)	-0.009*** (0.001)	-0.002 (0.001)
Obs.	353100	358546	358106	358544	358517	356264	348839
R2	0.88	0.95	0.93	0.91	0.93	0.93	0.93
Covariates	Y	Y	Y	Y	Y	Y	Y
Sample	All	All	All	All	All	All	All
Exposure	Y	Y	Y	Y	Y	Y	Y
Outcome baseline	7.07	2.82	2.76	2.34	2.53	2.98	3.16

Notes: The table reports estimates of the EITC's impact on the log of employment and the log of real hourly wages across occupations. Column 1 shows employment in levels. Columns 2 to 7 are the real hourly wages across the percentiles of the distribution paid by occupations. The treatment variable captures the interaction between the enactment of the EITC and an indicator for EITC exposure above the 75th percentile of exposure. Models control for the unemployment rate, state per capita GDP, minimum wage, state marginal tax rate, the percent of female workers, average race and ethnicity, educational achievement, number of children, median family income, and the age of workers in the industry. Standard errors are clustered at the state level. *** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$.

Table B.3: Estimates of the EITC on occupational outcomes without occupational exposure.

	(1) Employment	(2) Mean Wages	(3) Median Wages	(4) 10th Pct	(5) 25th Pct	(6) 75th Pct	(7) 90th Pct
EITC Enactment	-103.701 (89.035)	0.059* (0.031)	0.057* (0.030)	0.129*** (0.020)	0.106*** (0.024)	0.017 (0.037)	-0.072 (0.045)
Obs.	353100	358546	358106	358544	358517	356264	348839
R2	0.50	0.88	0.87	0.84	0.86	0.88	0.87
Covariates	Y	Y	Y	Y	Y	Y	Y
Sample	All	All	All	All	All	All	All
Exposure	N	N	N	N	N	N	N
Outcome baseline	5095.72	17.25	16.18	10.51	12.87	20.29	24.42

Notes: The table reports difference-in-difference and triple difference estimates of the EITC's impact on the employment and real weekly wages across occupations. Column 1 shows employment in levels. Columns 2 to 7 are the real hourly wages across the percentiles of the distribution paid by occupations. The treatment variable captures the interaction between the enactment of the EITC and an indicator for EITC exposure above the 75th percentile of exposure. Models control for the unemployment rate, state per capita GDP, minimum wage, state marginal tax rate, the percent of female workers, average race and ethnicity, educational achievement, number of children, median family income, and the age of workers in the industry. Standard errors are clustered at the state level. *** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$.

Table B.4: Impact of the EITC on occupational outcomes conditional on refundability.

	(1) Employment	(2) Mean Wages	(3) Median Wages	(4) 10th Pct	(5) 25th Pct	(6) 75th Pct	(7) 90th Pct
Refundable Sample							
EITC Enactment	-157.539* (90.799)	-1.710*** (0.031)	-1.648*** (0.029)	-0.931*** (0.020)	-1.284*** (0.024)	-1.941*** (0.036)	-2.083*** (0.043)
Obs.	322729	327401	327020	327399	327375	325325	318624
R2	0.50	0.89	0.88	0.85	0.87	0.88	0.87
Covariates	Y	Y	Y	Y	Y	Y	Y
Sample	Refundable	Refundable	Refundable	Refundable	Refundable	Refundable	Refundable
Exposure	Y	Y	Y	Y	Y	Y	Y
Outcome baseline	4787.65	18.74	17.59	11.35	13.96	22.04	26.39
Non-Refundable Sample							
EITC Enactment	-143.850 (119.053)	-1.592*** (0.038)	-1.575*** (0.036)	-0.924*** (0.024)	-1.236*** (0.029)	-1.788*** (0.045)	-1.784*** (0.054)
Obs.	260645	264934	264628	264932	264910	263283	258073
R2	0.49	0.88	0.87	0.84	0.86	0.87	0.87
Covariates	Y	Y	Y	Y	Y	Y	Y
Sample	Non-Refundable	Non-Refundable	Non-Refundable	Non-Refundable	Non-Refundable	Non-Refundable	Non-Refundable
Exposure	Y	Y	Y	Y	Y	Y	Y
Outcome baseline	4893.04	17.65	16.56	10.75	13.17	20.76	24.89

Notes: The table reports difference-in-difference and triple difference estimates of the EITC's impact on the employment and real weekly wages across occupations. The sample is restricted to exclude states with nonrefundability in the upper panel, and excluding states with refundability in the bottom panel. Column 1 shows employment in levels. Columns 2 to 7 are the real hourly wages across the percentiles of the distribution paid by occupations. The treatment variable captures the interaction between the enactment of the EITC and an indicator for EITC exposure above the 75th percentile of exposure. Models control for the unemployment rate, state per capita GDP, minimum wage, state marginal tax rate, the percent of female workers, average race and ethnicity, educational achievement, number of children, median family income, and the age of workers in the industry. *** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$.

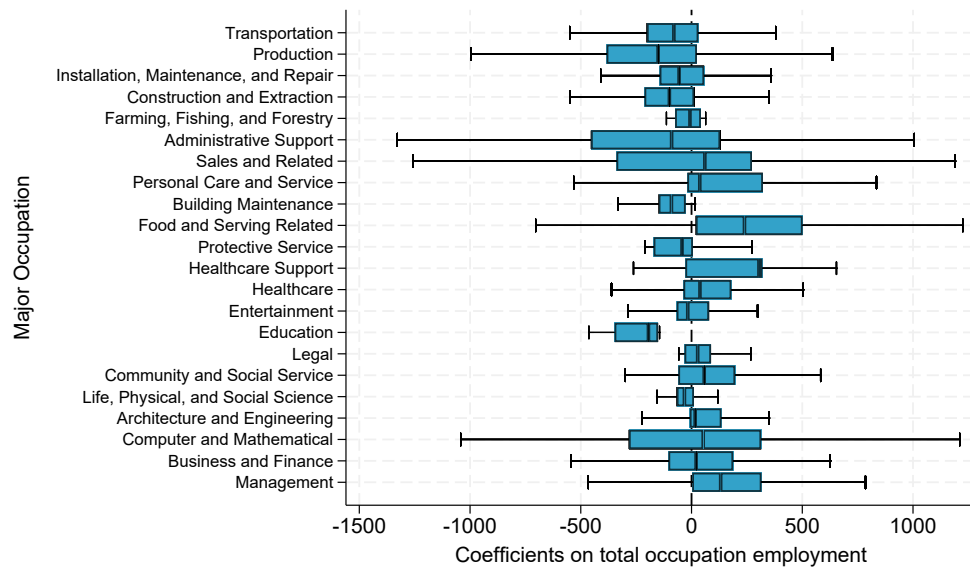


Figure B.2: Heterogeneous effects of the EITC across major occupations: Employment.

Notes: The figure presents box plots for the distribution of estimates of the impact of EITC roll-out grouped by major occupations. Models control for state and year fixed effects, the unemployment rate, state per capita GDP, minimum wage, state marginal tax rate, the percent of female workers, average race and ethnicity, educational achievement, number of children, median family income, and the age of workers in the industry. Standard errors are clustered at the state level.

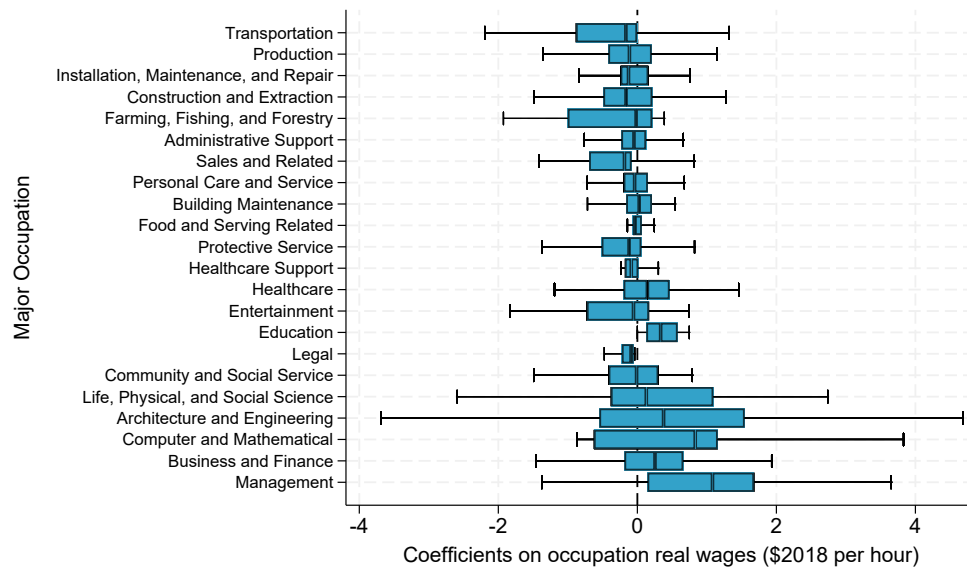


Figure B.3: Heterogeneous effects of the EITC across major occupations: Hourly wages.

Notes: The figure presents box plots for the distribution of estimates of the impact of EITC roll-out grouped by major occupations. Models control for state and year fixed effects, the unemployment rate, state per capita GDP, minimum wage, state marginal tax rate, the percent of female workers, average race and ethnicity, educational achievement, number of children, median family income, and the age of workers in the industry. Standard errors are clustered at the state level.

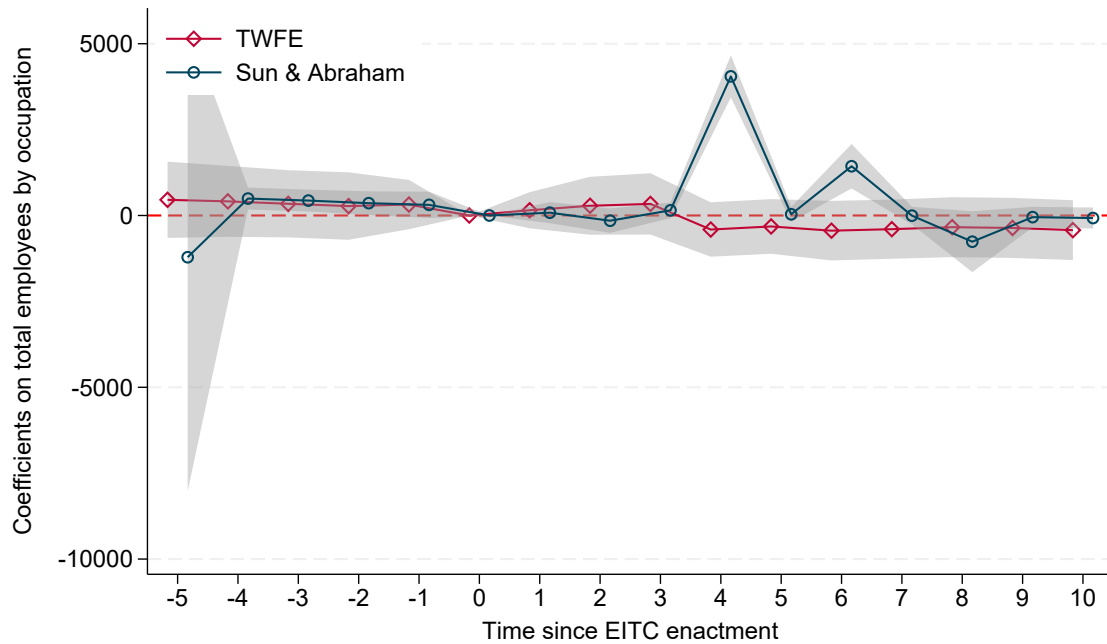


Figure B.4: Event study of EITC roll-out and employment across occupations.

Notes: The figure presents standard event studies using the Two Way Fixed Effects (TWFE) estimator and the [Sun and Abraham \(2021\)](#) estimator. The outcome is the number of employed workers in each occupation. The treatment variable captures the interaction between the enactment of the EITC and an indicator for EITC exposure above the 75th percentile of exposure. Models control for occupation fixed effects, state fixed effects, year fixed effects, the unemployment rate, state per capita GDP, minimum wage, state marginal tax rate, the percent of female workers, average race and ethnicity, educational achievement, number of children, median family income, and the age of workers in the industry. The 95 percent confidence intervals are estimated with cluster-robust standard errors at the state-level.

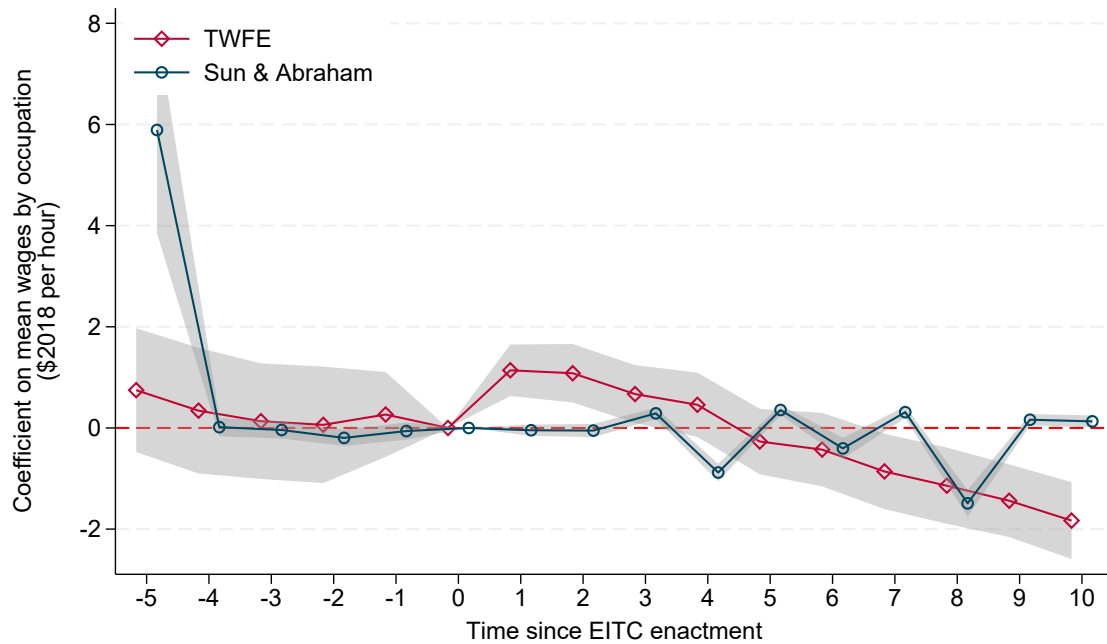


Figure B.5: Event study of EITC roll-out and wages across occupations.

Notes: The figure presents standard event studies using the Two Way Fixed Effects (TWFE) estimator and the [Sun and Abraham \(2021\)](#) estimator. The outcome is the real wages across occupations. The treatment variable captures the interaction between the enactment of the EITC and an indicator for EITC exposure above the 75th percentile of exposure. Models control for occupation fixed effects, state fixed effects, year fixed effects, the unemployment rate, state per capita GDP, minimum wage, state marginal tax rate, the percent of female workers, average race and ethnicity, educational achievement, number of children, median family income, and the age of workers in the industry. The 95 percent confidence intervals are estimated with cluster-robust standard errors at the state-level.

B.3 Additional robustness tests

Table B.5: Population weighted coefficients of the EITC on industry outcomes.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of Establishments								
EITC Enactment	-95.454 (51.664)	-56.362* (22.688)	-62.517* (26.356)	-50.328* (22.241)	-110.433 (58.903)	-108.066 (56.282)	-99.996 (57.797)	-75.636 (65.275)
Employment								
EITC Enactment	-1179.284* (555.061)	-801.807** (238.574)	-845.448** (278.123)	-1038.104*** (294.599)	-2967.940* (1225.977)	-2978.008* (1215.051)	-2919.862* (1224.526)	-2455.220 (1251.536)
Real Weekly Wages								
EITC Enactment	1.357 (15.835)	14.722 (8.412)	21.490* (8.993)	5.325 (10.410)	-80.437** (27.596)	-67.290* (26.151)	-77.312** (27.351)	-81.636* (36.048)
Obs.	2.02e+09	2.02e+09	1.90e+09	1.66e+09	2.02e+09	2.02e+09	1.90e+09	1.66e+09
R2	0.49	0.50	0.50	0.53	0.50	0.51	0.50	0.53
Covariates	N	Y	Y	Y	N	Y	Y	Y
EITC	All	All	Refundable	Non-Refundable	All	All	Refundable	Non-Refundable
Exposure	N	Y	N	N	Y	Y	Y	Y

Notes: The table reports difference-in-difference and triple difference estimates of the EITC's impact on the number of establishments, employment, and real weekly wages across industries. The treatment variable captures whether the state passed the EITC. Estimators with exposure use the interaction between the enactment of the tax credit and an indicator for EITC exposure above the 75th percentile of exposure. Estimates based on the enactment of a refundable (or non-refundable) EITC supplement exclude observations from states that ever enacted a non-refundable (refundable) program. All models include state fixed effects, year fixed effects, and industry fixed effects. Models with covariates control for the unemployment rate, state per capita GDP, minimum wage, state marginal tax rate, the percentage of female workers, average race and ethnicity, educational achievement, number of children, median family income, and the age of workers in the industry. Standard errors are clustered at the state level. All results are weighted by the state population. $p < 0.01$ $p < 0.05$ $p < 0.1$.

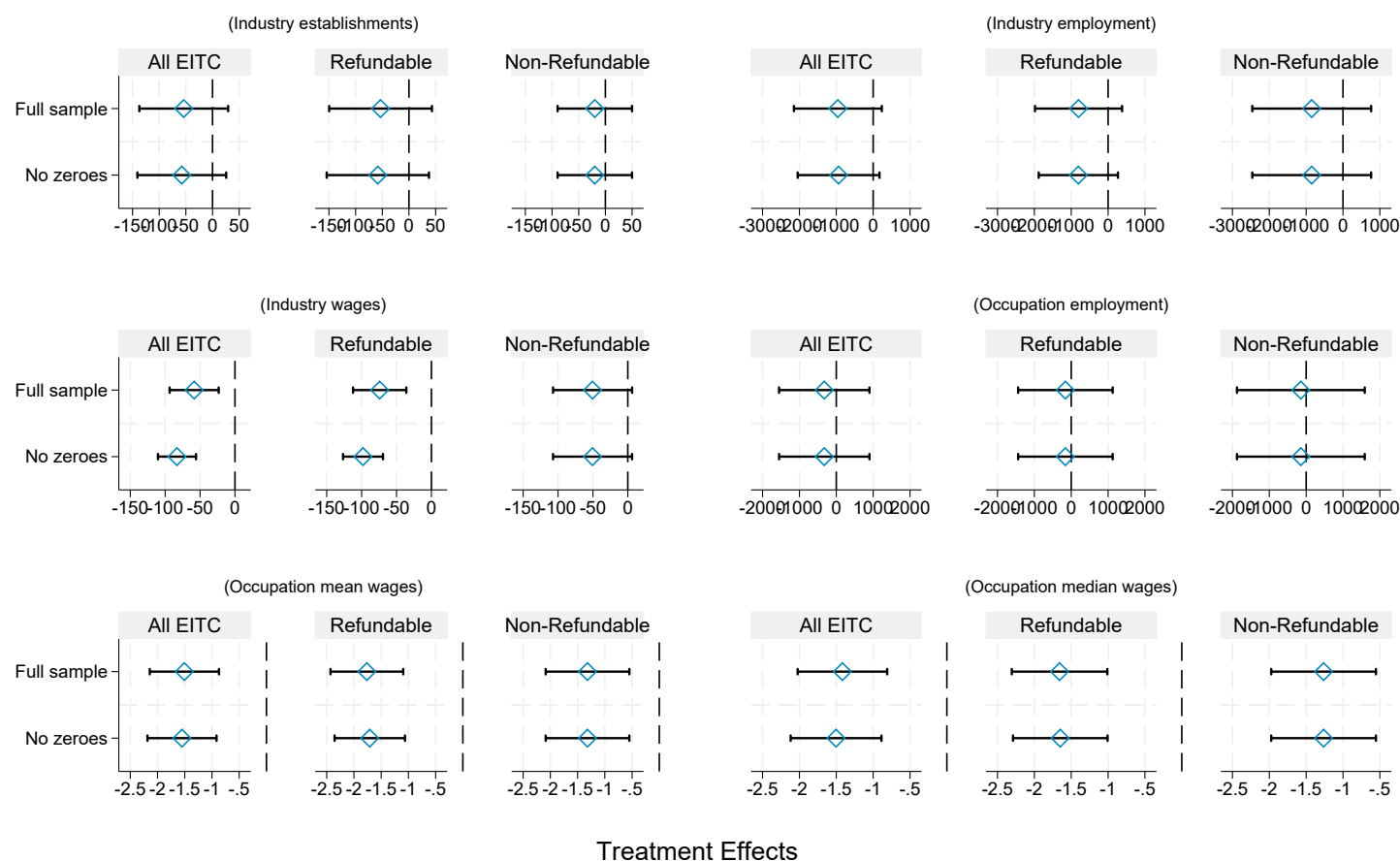


Figure B.6: Sensitivity of coefficients to the exclusion of zeroes in the sample.

Notes: The figure plots the triple difference coefficients from regressions on industry and occupational outcomes. The treatment variable interacts state roll-out and whether an industry or occupation is above the 75th percentile of exposure. “Full sample” models include all observations. “No zeroes” models exclude observations where the outcome is zero. Models control for occupation fixed effects, state fixed effects, year fixed effects, the unemployment rate, state per capita GDP, minimum wage, state marginal tax rate, the percent of female workers, average race and ethnicity, educational achievement, number of children, median family income, and the age of workers in the industry. The 95 percent confidence intervals are estimated with cluster-robust standard errors at the state-level.

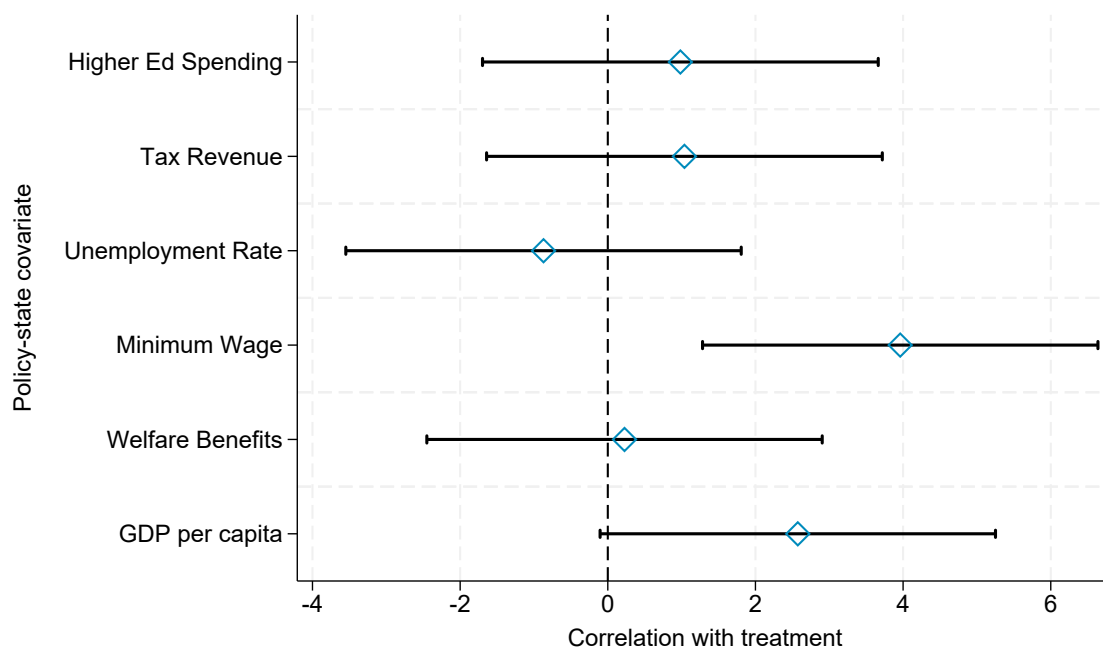


Figure B.7: Correlation between state EITC roll-out and policy state covariates.

Notes: The figure presents the correlation between policy covariates and the EITC roll-out across states. Each diamond represents a different regression where the outcome is state roll-out of the EITC supplementary program. Coefficients are re-scaled for cosmetic reasons, dividing the coefficient by the standard error. Models include no controls.

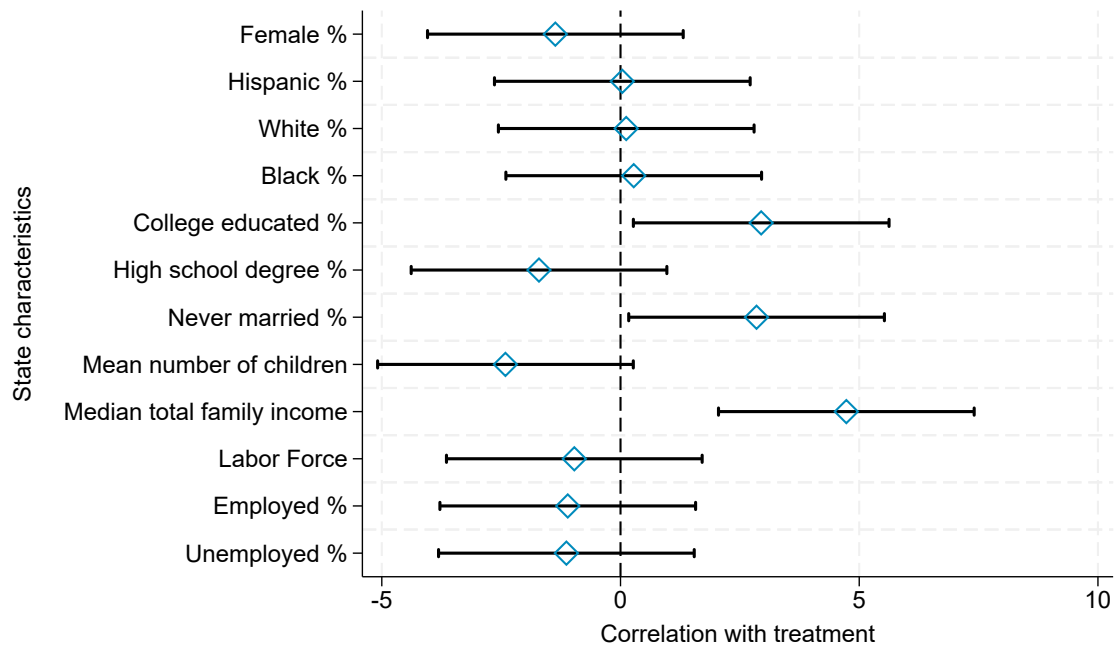


Figure B.8: Correlation between state EITC roll-out and state sociodemographic characteristics.

Notes: The figure presents the correlation between state's sociodemographic characteristics and the EITC roll-out across states. Each diamond represents a different regression where the outcome is state roll-out of the EITC supplementary program. Coefficients are re-scaled for cosmetic reasons, dividing the coefficient by the standard error. Models include no controls.

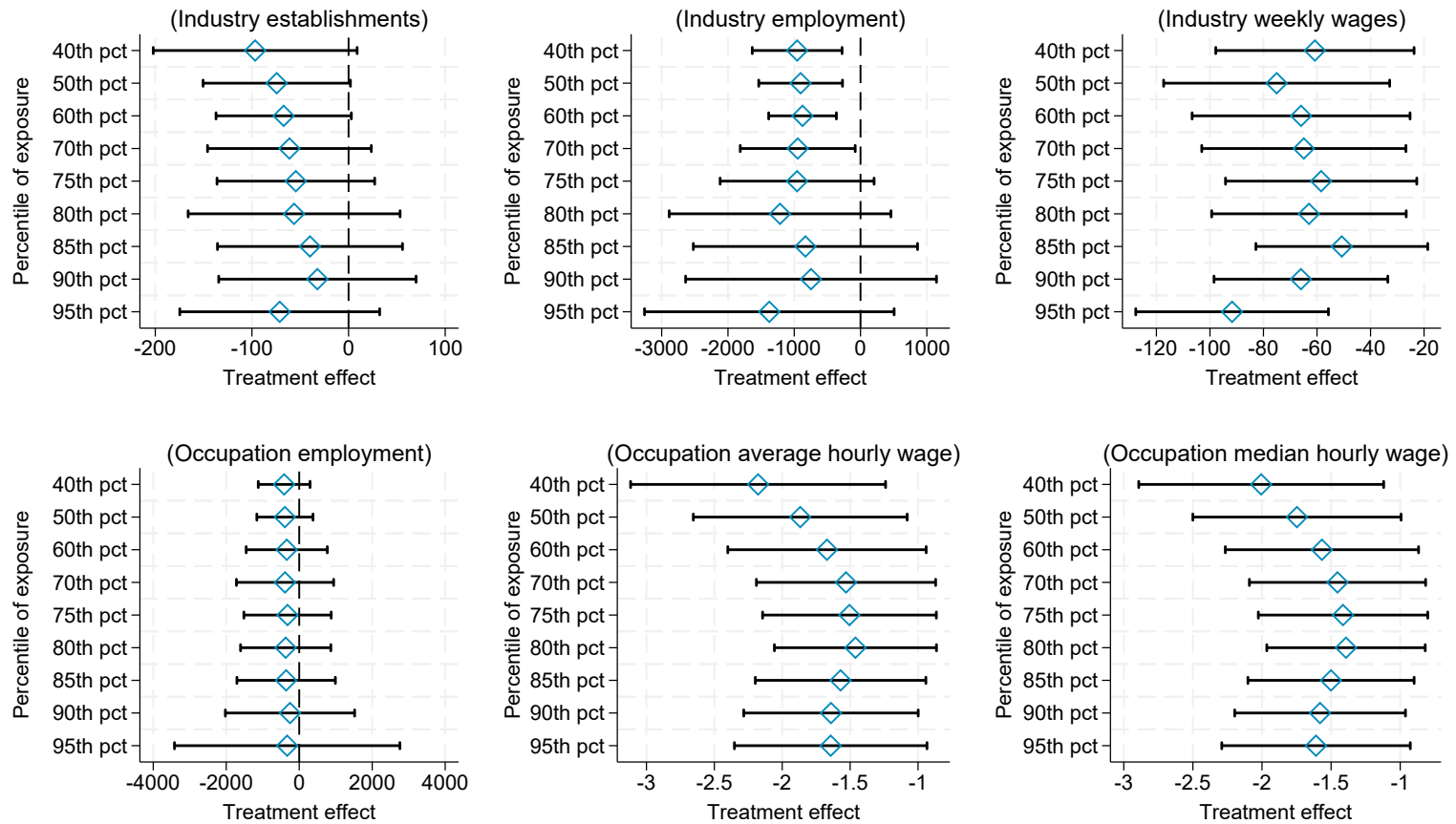


Figure B.9: Sensitivity of treatment effects to various thresholds of exposure.

Notes: The figure plots the estimated triple difference coefficients from 54 regressions. Each diamond represents a different regression on the industry and occupation outcomes. The treatment variable varies across regressions as EITC exposure is defined above different percentiles of the distribution. Models control for occupation fixed effects, state fixed effects, year fixed effects, the unemployment rate, state per capita GDP, minimum wage, state marginal tax rate, the percent of female workers, average race and ethnicity, educational achievement, number of children, median family income, and the age of workers in the industry. The 95 percent confidence intervals are estimated with cluster-robust standard errors at the state-level.

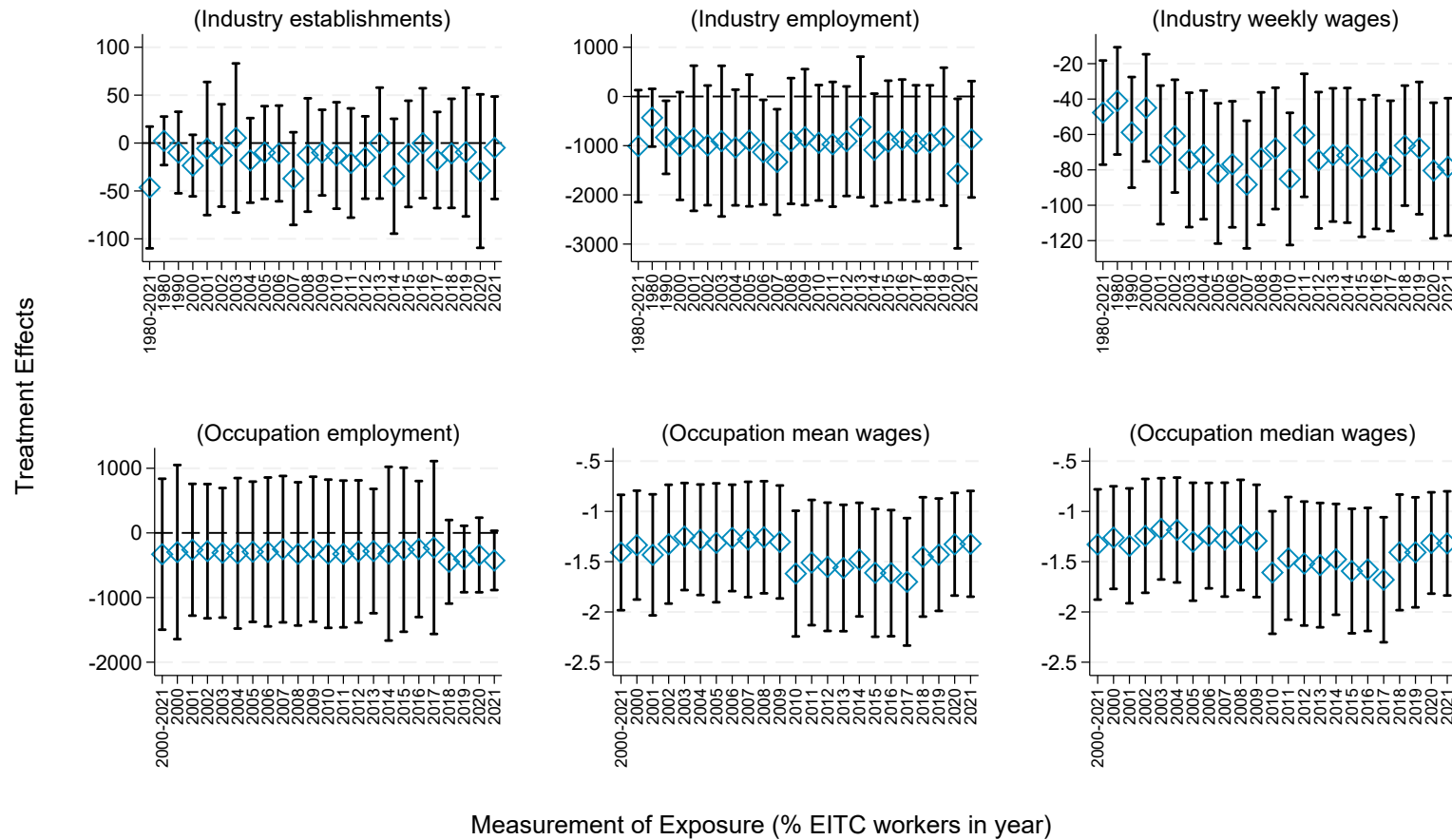


Figure B.10: Sensitivity of treatment effects to year of exposure.

Notes: The figure plots the triple difference coefficients from regressions on industry and occupational outcomes. The treatment variable varies across regressions as EITC exposure is defined using the 75th percentile distribution of eligible workers in various year in the ACS. The first estimate is computed with the percent of eligible workers in each industry across all years (1980 to 2021 in industries, and 2000 to 2021 in occupations). Models control for occupation fixed effects, state fixed effects, year fixed effects, the unemployment rate, state per capita GDP, minimum wage, state marginal tax rate, the percent of female workers, average race and ethnicity, educational achievement, number of children, median family income, and the age of workers in the industry. The 95 percent confidence intervals are estimated with cluster-robust standard errors at the state-level.

Appendix C

C.0.1 Descriptive Tables and Figures

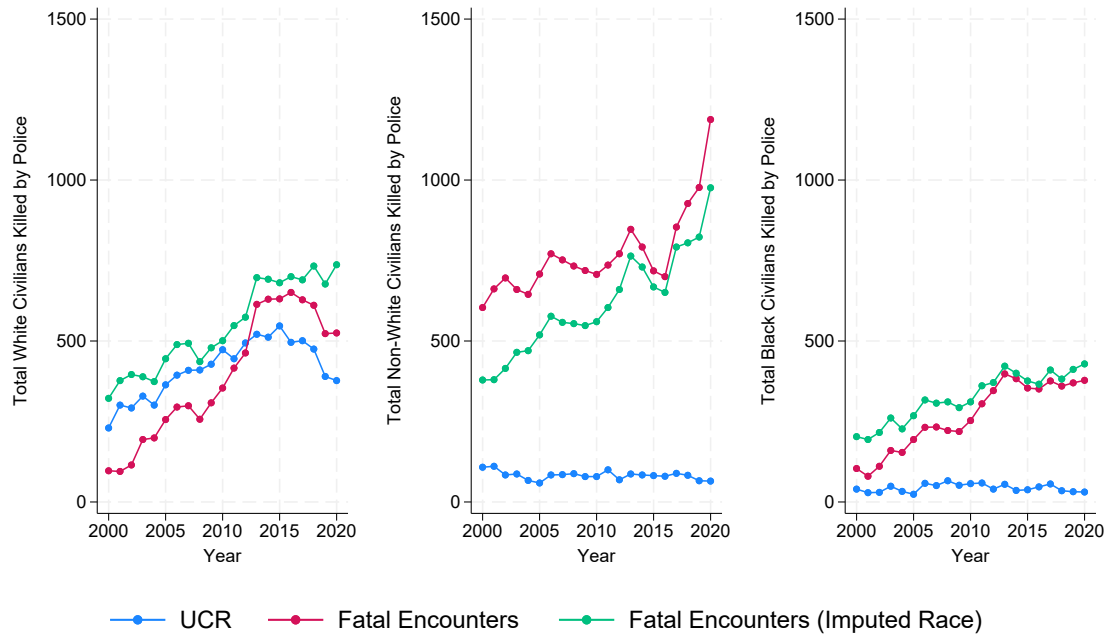


Figure C.1: Total civilians killed by police by race.

Table C.1: Total Original Acquisition Value of 1033 items shipped by category.

Period	Total	Weapons	Vehicles	Optics	Clothing	Others
1990-1994	1.68	0.40	1.02	0.26	-	0.00
1995-1999	36.59	5.00	27.39	3.95	-	0.26
2000-2004	44.77	7.87	32.31	4.46	-	0.13
2005-2009	150.47	23.98	106.12	17.81	0.30	2.26
2010-2014	1,566.65	44.98	1,025.17	141.80	14.49	340.21
2015-2020	1,596.48	72.46	802.11	115.25	29.14	577.52
Total	3,396.71	154.68	1,994.18	283.53	43.92	920.39

	Total	Tactical	Non-Tactical	Military Vehicles
1990-1994	1.68	0.66	0.00	1.02
1995-1999	36.59	8.95	0.26	27.39
2000-2004	44.77	12.33	0.13	32.31
2005-2009	150.47	41.79	15.65	93.03
2010-2014	1,566.65	186.77	374.49	1,005.38
2015-2020	1,596.48	187.71	659.03	749.74
Total	3,396.71	438.21	1,049.56	1,908.94

	Guns, Rifles	Magazines
1990-1994	0.40	-
1995-1999	4.86	-
2000-2004	7.68	-
2005-2009	23.45	0.02
2010-2014	19.92	1.58
2015-2020	2.85	1.17
Total	59.16	2.77

Notes: The original acquisition value is in millions of current dollars.

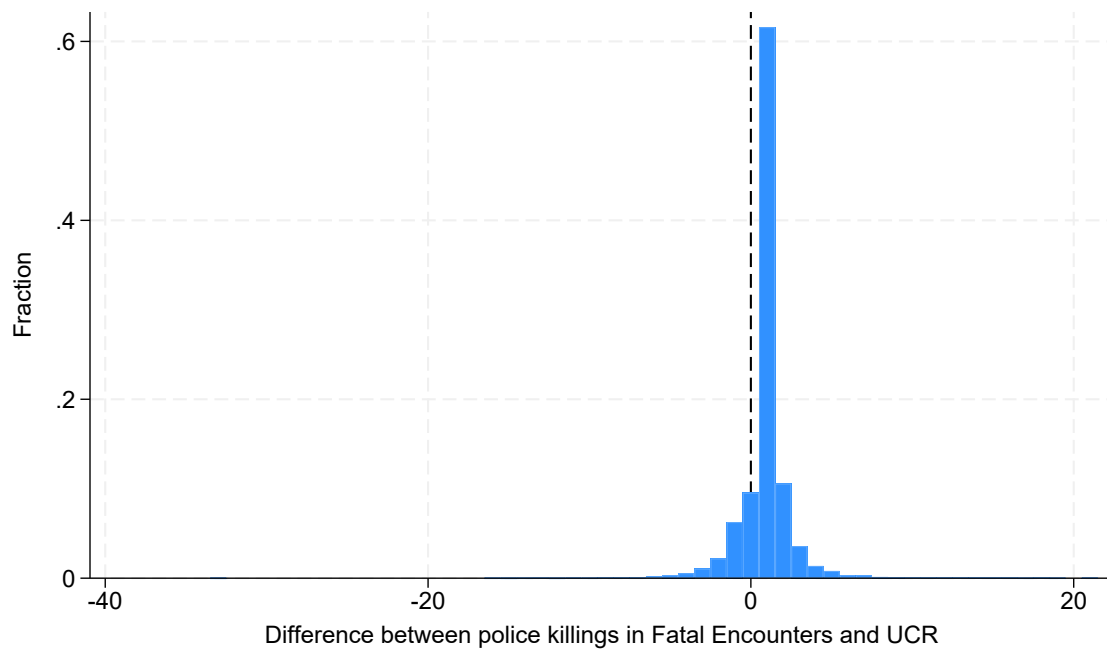


Figure C.2: Distribution of the gap in civilian killings between Fatal Encounters and UCR.

Notes: The Figure plots distribution of differences in the total count of civilian killed by police every quarter between Fatal Encounters and UCR.

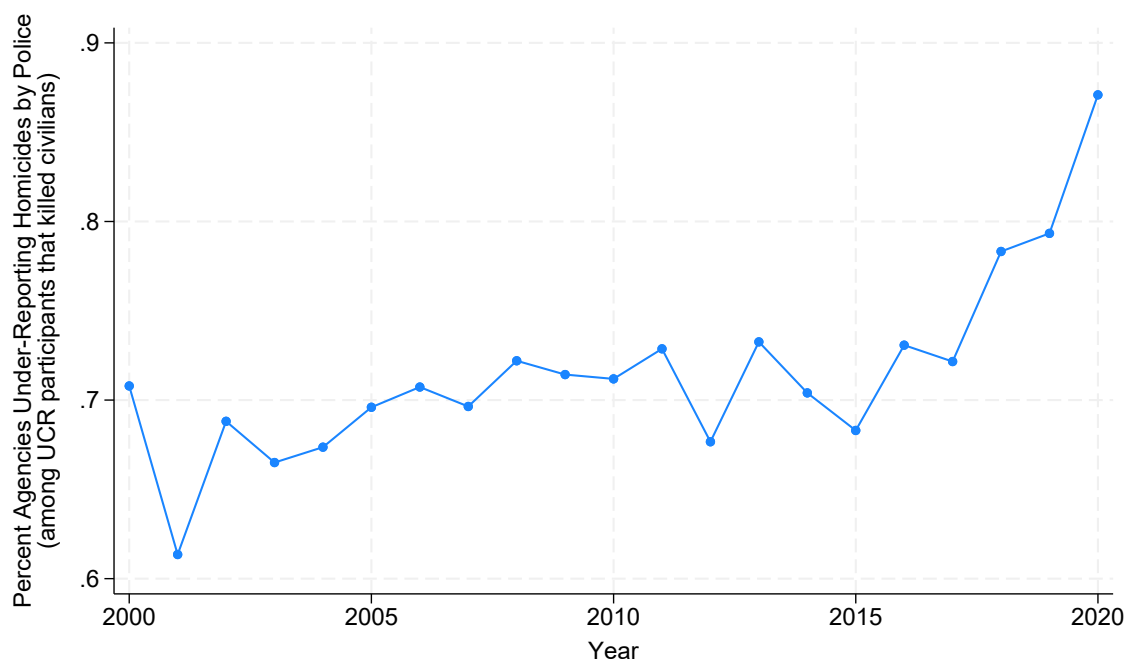


Figure C.3: Percent of UCR respondents that do not disclose a homicide by police.

Notes: The Figure plots the percent of UCR respondents that do not disclose a homicide by police conditional on a homicide recorded in Fatal Encounters.

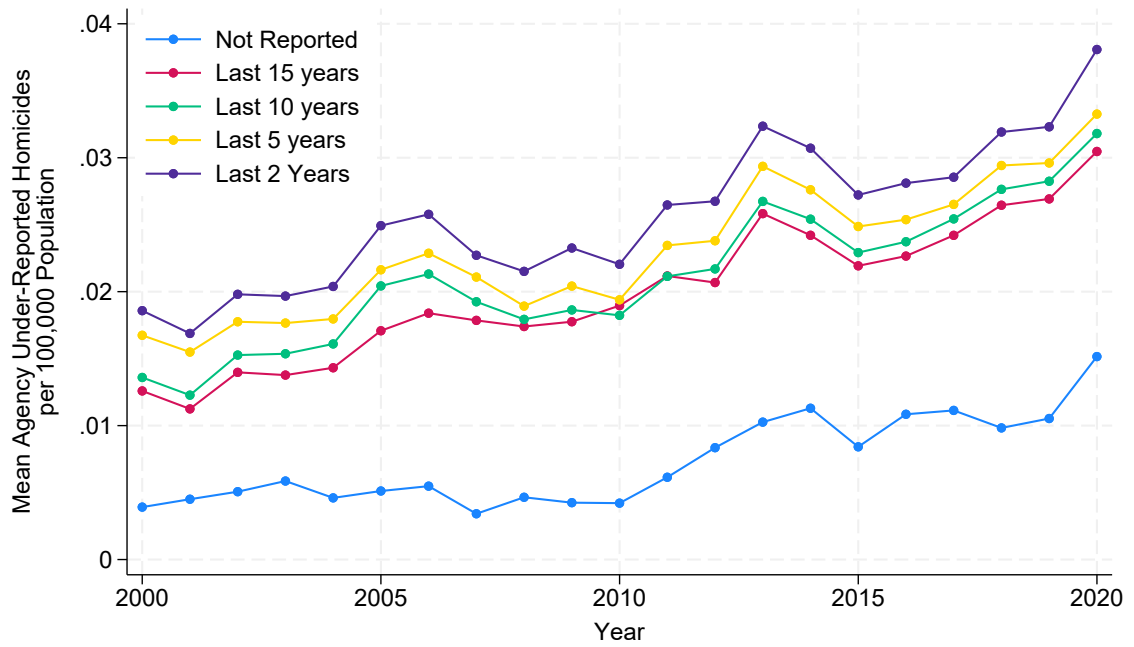


Figure C.4: Average gap between Fatal Encounters and UCR conditional on recent reporting behavior.

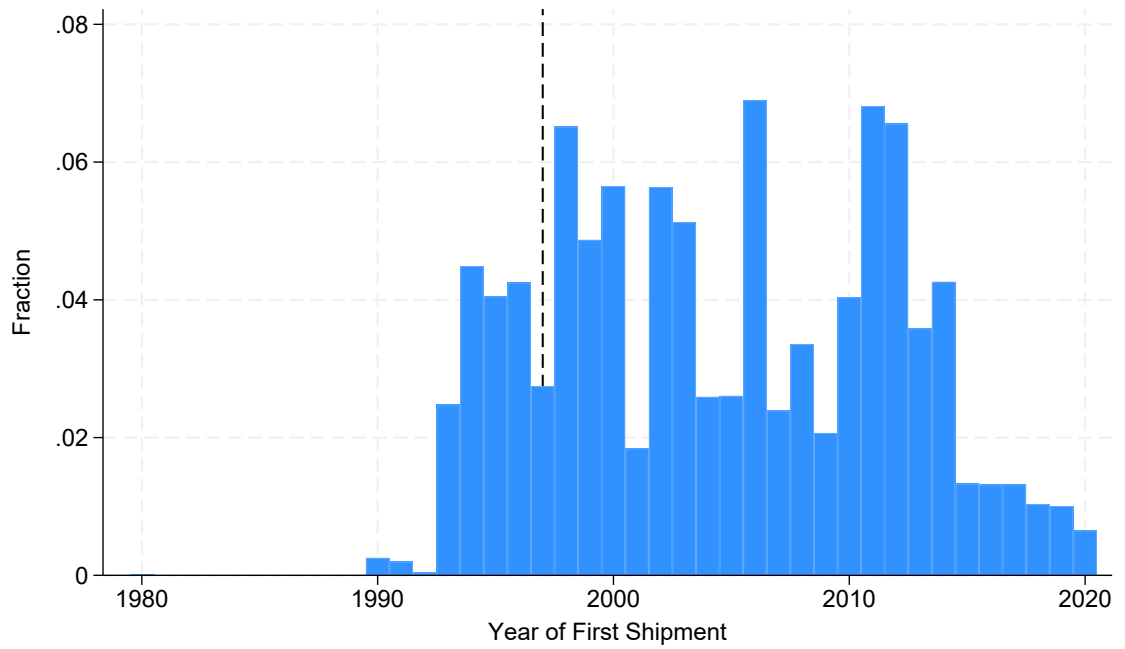


Figure C.5: The distribution of 1033 take-up year.

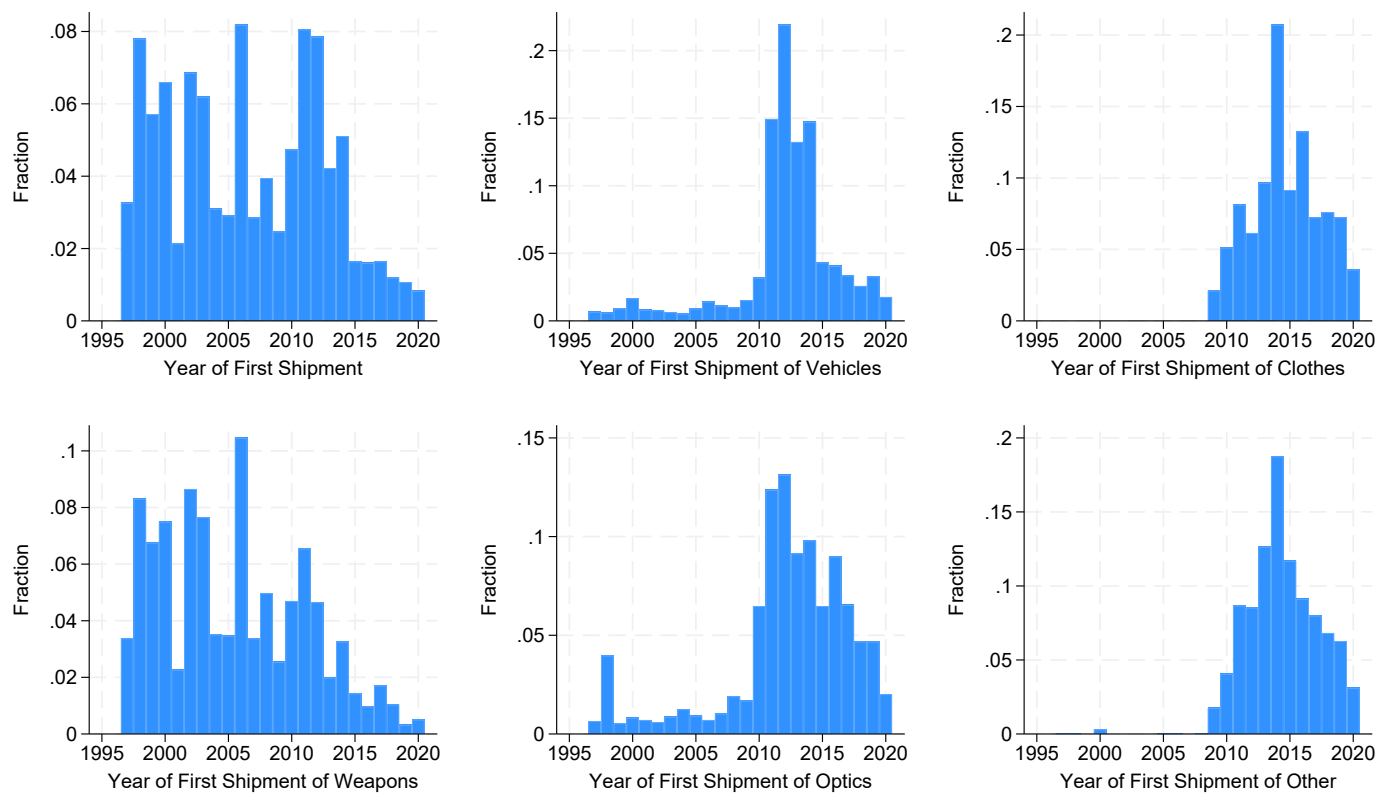


Figure C.6: The distribution of 1033 take-up year by shipment type.

C.0.2 Robustness tests

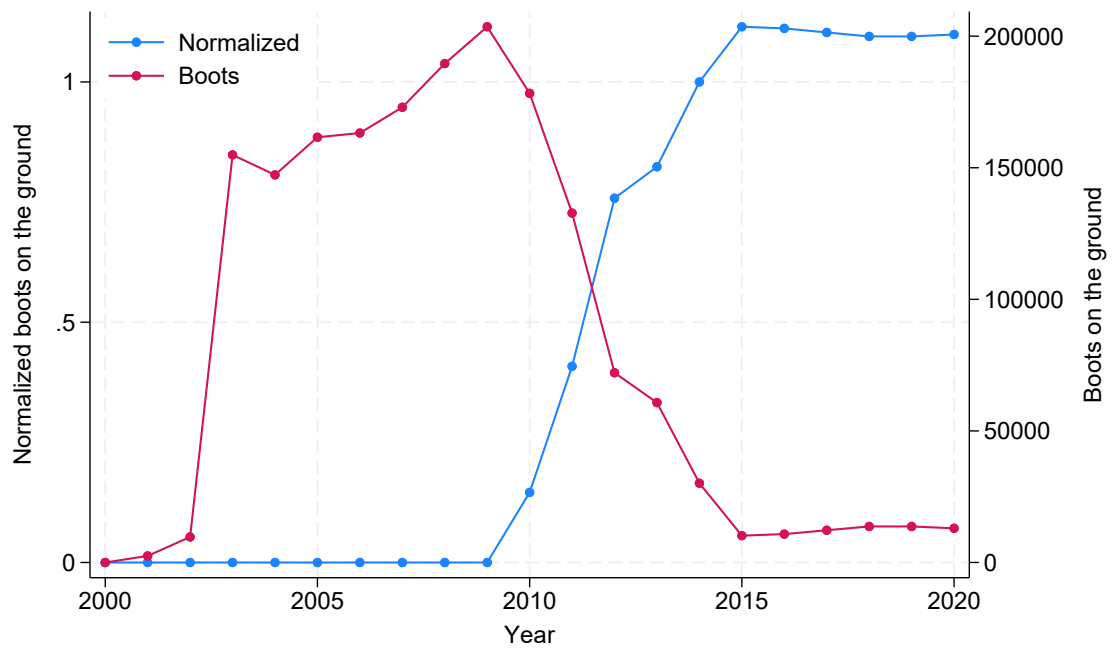


Figure C.7: Boots on the ground

Notes: The figure plots two measurements of boots on the ground in Afghanistan and Iraq. The count of boots on the ground, and the normalized series based on Masera (2021a,2021b).

Table C.2: The effect of military items on killings by police with alternative instruments.

	Main	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Dependent variable: killed by police in UCR								
Items	0.000 (0.002)	-0.001 (0.003)	-0.000 (0.003)	-0.000 (0.003)	0.001 (0.003)	0.002 (0.003)	0.001 (0.003)	-0.001 (0.001)
Dependent Variable: civilians killed by police in Fatal Encounters								
Items	0.009*** (0.003)	0.013*** (0.004)	0.013*** (0.004)	0.012*** (0.004)	0.009*** (0.003)	0.010*** (0.003)	0.008*** (0.003)	0.000 (0.001)
Obs	235069	235069	235069	235069	235069	235069	235069	235069
F-statistic	40.27	55.37	44.36	32.07	27.64	32.12	30.24	63.71
Close DLA	Yes	Yes	Quadratic	Cubic	Cubic	Yes	Yes	Yes
Close Mil	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Time	EQP	EQP	EQP	EQP	EQP	$\sum \Delta b $	Mil Exp	Total Items
Time*Close	-0.455	-0.405	-0.680	-0.243	-0.454	-0.000	-0.000	0.002
Time*Close Sq			0.001	-0.001	-0.001			
Time*Close Cu				0.000	0.000			
Time*Mil	0.545				0.533	0.000	0.000	-0.002

Notes: The table shows the effect of one hundred additional items on the homicides by police in UCR and Fatal Encounters. The F-statistic is the Kleibergen-Paap rk Wald F statistic. Each column uses a different specification in the first stage: **Main model:** the instrument uses time variation in normalized boots in the ground (EQP) interacted with distance to a DLA storage facility and interacted with the distance to the closest military base (Masera, 2021a); **Model 2:** the first stage includes time variation in normalized boots in the ground (EQP) interacted with distance to a DLA storage facility; **Model 3:** the first stage includes time variation in EQP interacted with distance to the closest DLA storage facility and a second interaction between EQP and quadratic distance to the closest DLA facility; **Model 4:** the first stage includes time variation in EQP interacted with distance to the closest DLA storage facility, an interaction between EQP and quadratic distance, and an interaction between EQP and cubic distance; **Model 5:** the first stage includes time variation in EQP interacted with distance to a DLA storage facility, quadratic distance, and cubic distance, and an interaction between EQP and distance to the closest military base; **Model 6:** the first stage includes time variation in the cumulative absolute variation in boots on the ground interacted with distance to the closest DLA storage facility and distance to the closest military base; **Model 7:** the first stage includes time variation in cumulative real military expenditure, similar to Bove and Gavrilova (2017), interacted with distance to the closest DLA storage facility and distance to the closest military base; **Model 8:** the first stage includes time variation in cumulative items in the nation, as in Harris et al. (2017), interacted with distance to the closest DLA storage facility and distance to the closest military base. The regression controls for population size, labor force, unemployment, per capita personal income, the police department's total budget, the number of full time sworn officers, the total number of employees, year fixed effects, and agency fixed effects. Standard errors are clustered at the commuting zone by year.

Table C.3: The effect of one additional standard deviation in military items on the probability of under-reporting.

	Total	Weapons	Vehicles	Optics	Clothes	Others	Rifles	Ammunition
Items	0.069*** (0.020)	0.097*** (0.019)	0.179*** (0.057)	0.108*** (0.022)	0.075*** (0.016)	0.051** (0.021)	0.141*** (0.028)	0.093*** (0.019)
Obs	235069	235069	235069	235069	235069	235069	235069	235069
F-statistic	40.27	46.38	6.61	44.76	42.21	34.83	46.24	51.66

Notes: The table shows the effect of one standard deviation in additional items on the probability of a positive difference between homicides by police reported in Fatal Encounters and in UCR. The F-statistic is the Kleibergen-Paap rk Wald F statistic. The regression controls for population size, labor force, unemployment, per capita personal income, the police department's total budget, the number of full time sworn officers, the total number of employees, year fixed effects, and agency fixed effects. Clustered standard errors are computed at the commuting zone and year interaction.

Table C.4: The effect of military items on total under-reports.

	Total	Weapons	Vehicles	Optics	Clothes	Others	Rifles	Ammunition
Items	0.006** (0.002)	0.135*** (0.034)	9.256*** (3.115)	0.357*** (0.082)	0.040*** (0.011)	0.005 (0.003)	0.386*** (0.095)	0.206*** (0.054)
Obs	235069	235069	235069	235069	235069	235069	235069	235069
F-statistic	40.27	46.38	6.61	44.76	42.21	34.83	46.24	51.66
Eqp*Storage	-0.455	-0.032	-0.000	-0.011	-0.108	-0.303	-0.011	-0.021
Eqp*Mil Base	0.545	-0.025	-0.001	-0.018	0.001	0.587	-0.013	-0.005

Notes: The table shows the effect of one hundred additional items on the positive difference between homicides by police reported in Fatal Encounters and in UCR. The F-statistic is the Kleibergen-Paap rk Wald F statistic. The rows show the first-stage coefficient of normalized boots on the ground interacted with the distance to DLA storage, and the interaction with distance to a military base. The regression controls for population size, labor force, unemployment, per capita personal income, the police department's total budget, the number of full time sworn officers, the total number of employees, year fixed effects, and agency fixed effects. Clustered standard errors are computed at the commuting zone and year interaction.

Table C.5: The effect of logarithmic military items on total under-reports.

	Total	Weapons	Vehicles	Optics	Clothes	Others	Rifles	Ammunition
Items	0.067*** (0.015)	0.148*** (0.036)	0.411*** (0.119)	0.126*** (0.028)	0.065*** (0.015)	0.067*** (0.015)	0.385*** (0.133)	0.114*** (0.026)
Obs	235069	235069	235069	235069	235069	235069	235069	235069
F-statistic	66.80	30.84	11.82	70.83	183.62	98.10	5.98	148.76
Eqp*Storage	-0.001	-0.000	-0.000	-0.000	-0.001	-0.001	-0.000	-0.000
Eqp*Mil Base	-0.001	-0.000	-0.000	-0.001	-0.000	-0.001	-0.000	-0.000

Notes: The table shows the effect of a 1 percent increase in items on the positive difference between homicides by police reported in Fatal Encounters and in UCR. The F-statistic is the Kleibergen-Paap rk Wald F statistic. The rows show the first-stage coefficient of normalized boots on the ground interacted with the distance to DLA storage, and the interaction with distance to a military base. The regression controls for population size, labor force, unemployment, per capita personal income, the police department's total budget, the number of full time sworn officers, the total number of employees, year fixed effects, and agency fixed effects. The dependent variable is constructed by taking the log of items plus one $\log(\text{items}_{it} + 1)$. Standard errors are clustered at the commuting zone by year.

Table C.6: The effect of military items (OAV) on the probability of under-reporting.

	Total	Weapons	Vehicles	Optics	Clothes	Others	Rifles	Ammunition
Items	0.002*** (0.000)	0.023*** (0.000)	0.005*** (0.000)	0.014*** (0.000)	0.053*** (0.000)	0.007*** (0.000)	0.099*** (0.000)	1.125*** (0.000)
Obs	235069	235069	235069	235069	235069	235069	235069	235069
Mean value	197216	12648	379453	82356	19211	244958	7511	2312
F-statistic	30.60	42.81	22.07	50.45	52.06	14.45	64.47	51.98
Eqp*Storage	-95.559	-13.472	-21.829	-12.089	-5.749	-42.419	-0.693	-0.281
Eqp*Mil Base	-195.362	-10.394	-125.577	-43.268	-4.098	-12.026	-6.310	-0.070

Notes: The table shows the effect of \$10,000 in additional items on the probability of a positive difference between homicides by police reported in Fatal Encounters and in UCR. The F-statistic is the Kleibergen-Paap rk Wald F statistic. The rows show the average value of items per category, the first-stage coefficient of normalized boots on the ground interacted with the distance to DLA storage, and the interaction with distance to a military base. The regression controls for population size, labor force, unemployment, per capita personal income, the police department's total budget, the number of full time sworn officers, the total number of employees, year fixed effects, and agency fixed effects. Standard errors are clustered at the commuting zone by year.

Table C.7: The effect of military items on the probability of under-reporting using categories in Masera (2021a).

	Tactical	Non-Tactical	Military Vehicles
Items	0.068*** (0.000)	0.005*** (0.000)	5.749*** (0.093)
Obs	235069	235069	235069
F-statistic	52.43	38.27	9.26
Eqp*Storage	-0.043	-0.411	-0.000
Eqp*Mil Base	-0.042	0.588	-0.001

Notes: The table shows the effect of one hundred additional items on the probability of a positive difference between homicides by police reported in Fatal Encounters and in UCR. The F-statistic is the Kleibergen-Paap rk Wald F statistic. The rows display the first-stage coefficient of normalized boots on the ground interacted with the distance to DLA storage, and the interaction with distance to a military base. The regression controls for population size, labor force, unemployment, per capita personal income, the police department's total budget, the number of full-time sworn officers, the total number of employees, year fixed effects, and agency fixed effects. Clustered standard errors are computed at the commuting zone and year interaction.

Table C.8: The effect of total military items on the probability of under-reporting with alternative instruments.

	Main Model	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Items	0.005*** (0.001)	0.008*** (0.002)	0.008*** (0.002)	0.007*** (0.002)	0.005*** (0.001)	0.006*** (0.002)	0.005*** (0.002)	0.000** (0.000)
Obs	235069	235069	235069	235069	235069	235069	235069	235069
F-statistic	40.27	55.37	44.36	32.07	27.64	32.12	30.24	63.71
Dist Storage	Yes	Yes	Quadratic	Cubic	Cubic	Yes	Yes	Yes
Dist Mil Base	Yes	No	No	No	Yes	Yes	Yes	Yes
Time	EQP	EQP	EQP	EQP	EQP	$\sum \Delta b $	Mil Exp	Total Items
Time*Dist Storage	-0.455	-0.405	-0.680	-0.243	-0.454	-0.000	-0.000	0.002
Time*Dist Storage Sq			0.001	-0.001	-0.001			
Time*Dist Storage Cu				0.000	0.000			
Time*Dist Mil Base	0.545				0.533	0.000	0.000	-0.002

Notes: The table shows the effect of one hundred additional items on the probability of a positive difference between homicides by police reported in Fatal Encounters and in UCR. The F-statistic is the Kleibergen-Paap rk Wald F statistic. Each column includes a different specification in the first stage: **Main model:** the instrument uses time variation in normalized boots in the ground (EQP) interacted with distance to a DLA storage facility and interacted with the distance to the closest military base (Masera, 2021a); **Model 2:** the first stage includes time variation in normalized boots in the ground (EQP) interacted with distance to a DLA storage facility; **Model 3:** the first stage includes time variation in EQP interacted with distance to the closest DLA storage facility and a second interaction between EQP and quadratic distance to the closest DLA facility; **Model 4:** the first stage includes time variation in EQP interacted with distance to the closest DLA storage facility, an interaction between EQP and quadratic distance, and an interaction between EQP and cubic distance; **Model 5:** the first stage includes time variation in EQP interacted with distance to a DLA storage facility, quadratic distance, and cubic distance, and an interaction between EQP and distance to the closest military base; **Model 6:** the first stage includes time variation in the cumulative absolute variation in boots on the ground interacted with distance to the closest DLA storage facility and distance to the closest military base; **Model 7:** the first stage includes time variation in cumulative real military expenditure, similar to Bove and Gavrilova (2017), interacted with distance to the closest DLA storage facility and distance to the closest military base; **Model 8:** the first stage includes time variation in cumulative items in the nation, as in Harris et al. (2017), interacted with distance to the closest DLA storage facility and distance to the closest military base. The regression controls for population size, labor force, unemployment, per capita personal income, the police department's total budget, the number of full time sworn officers, the total number of employees, year fixed effects, and agency fixed effects. Standard errors are clustered at the commuting zone by year.

C.0.3 Transitional Matrices

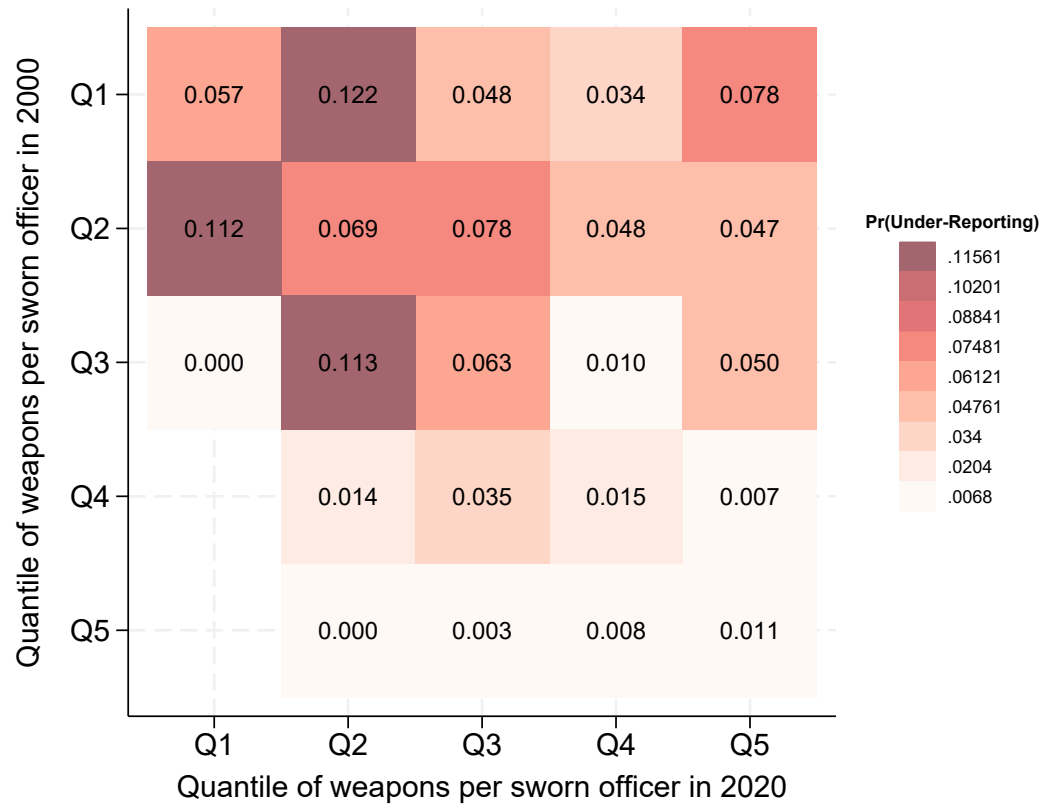


Figure C.8: Probability of under-reports as police departments transition in the stock of weapons.

Notes: The figure plots the probability of under-reporting during the last 5 years as departments transition between the quintile of the distribution of 1033 weapons in an initial period (vertical axis) into a future period (horizontal axis). The probability of under-reporting is estimated with population weights.

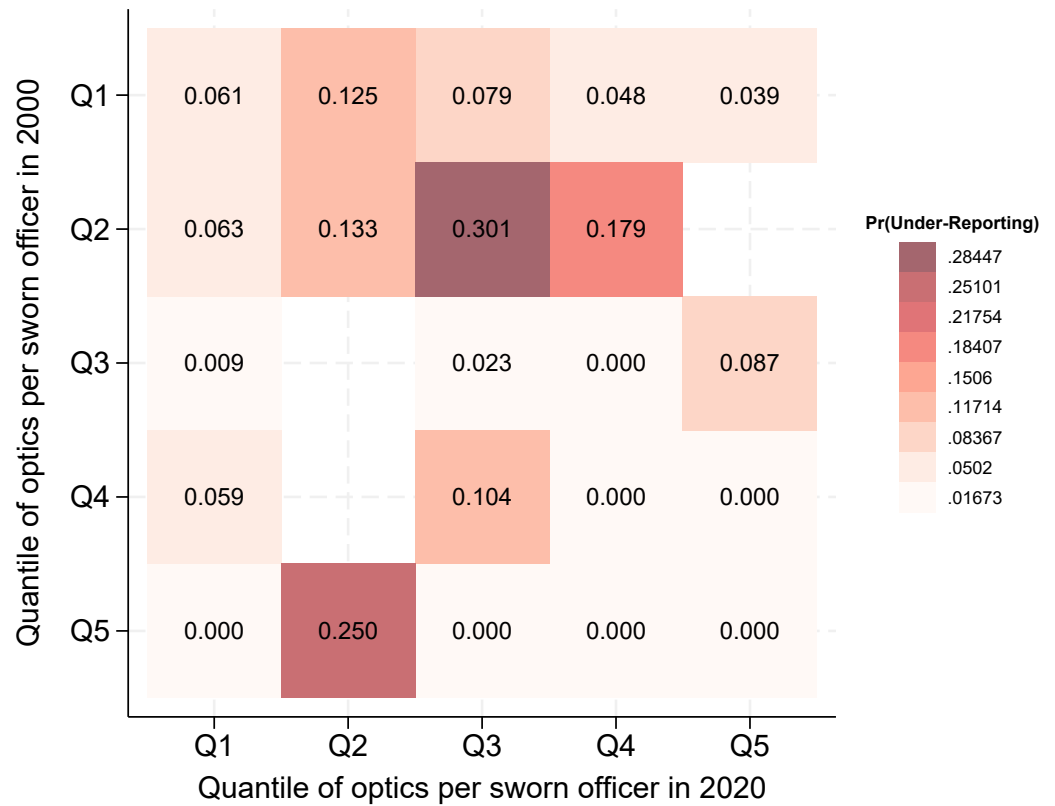


Figure C.9: Probability of under-reports as police departments transition in the stock of optics.

Notes: The figure plots the probability of under-reporting during the last 5 years as departments transition between the quintile of the distribution of 1033 optics in an initial period (vertical axis) into a future period (horizontal axis). The probability of under-reporting is estimated with population weights.

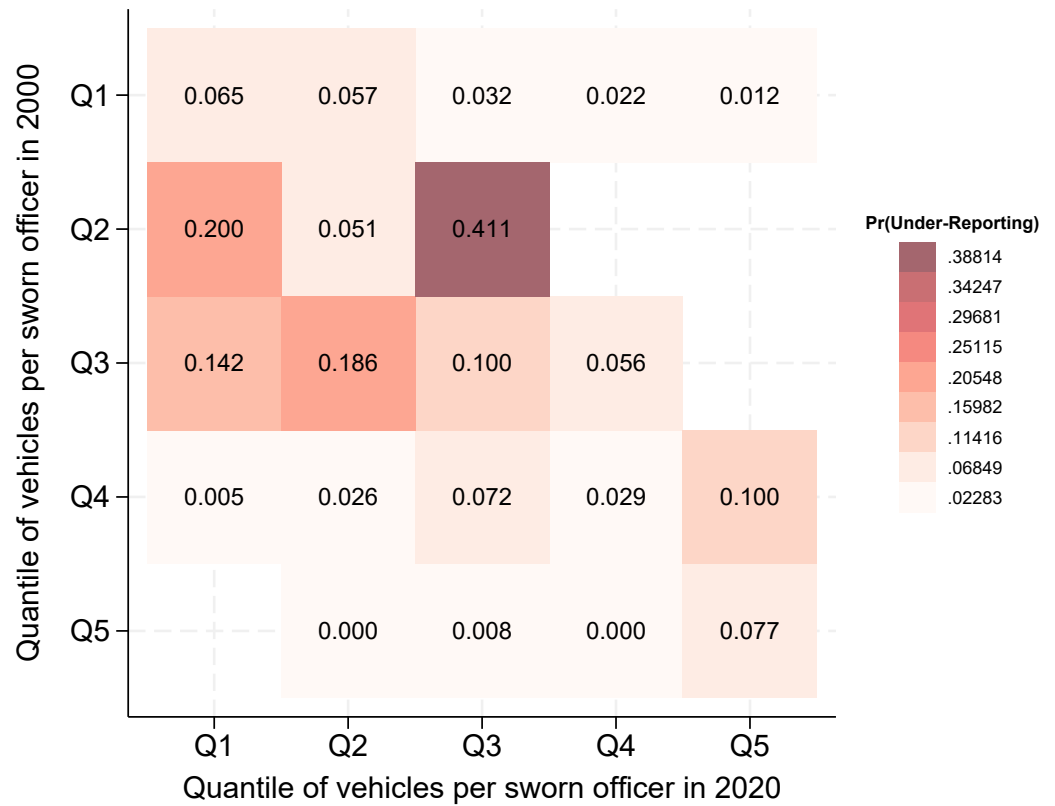


Figure C.10: Probability of under-reports as police departments transition in the stock of vehicles.

Notes: The figure plots the probability of under-reporting during the last 5 years as departments transition between the quintile of the distribution of 1033 vehicles in an initial period (vertical axis) into a future period (horizontal axis). The probability of under-reporting is estimated with population weights.

C.0.4 Assets Forfeiture

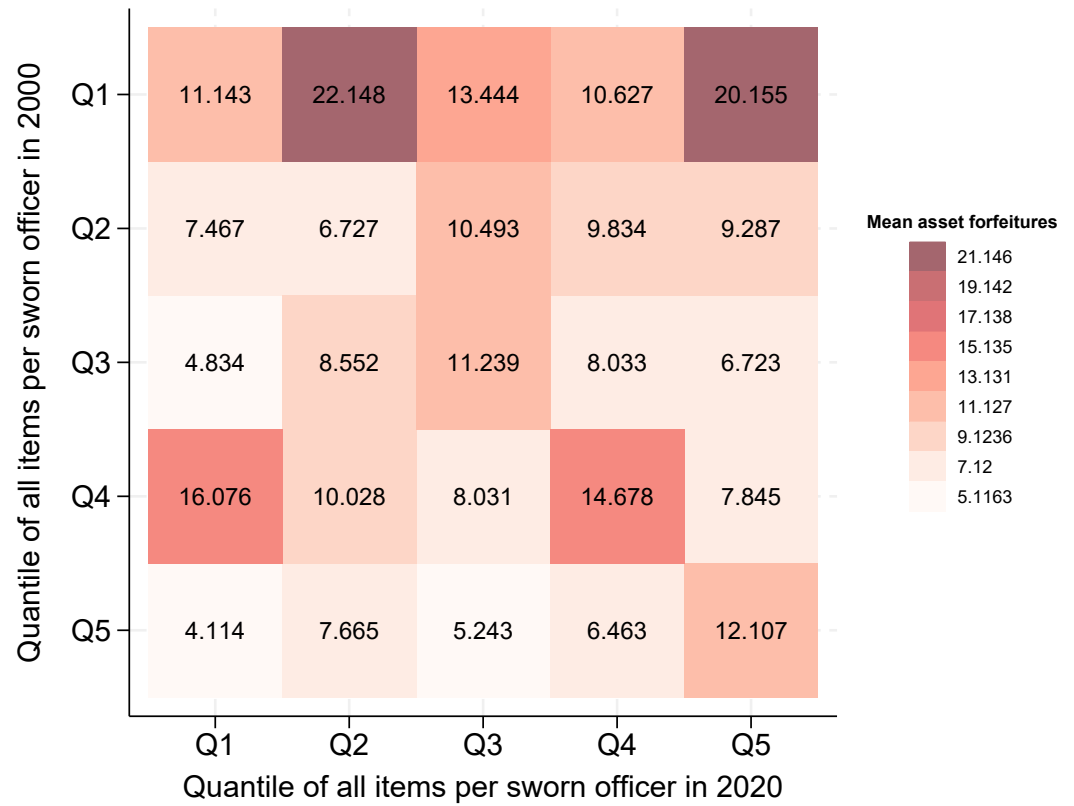


Figure C.11: Average real dollars per capita in assets seized as police departments transition in the stock of 1033 items.

Notes: The figure plots the mean assets forfeiture between 1977 and 1997 as departments transition between the quintile of the distribution of 1033 items in an initial period (vertical axis) into a future period (horizontal axis).

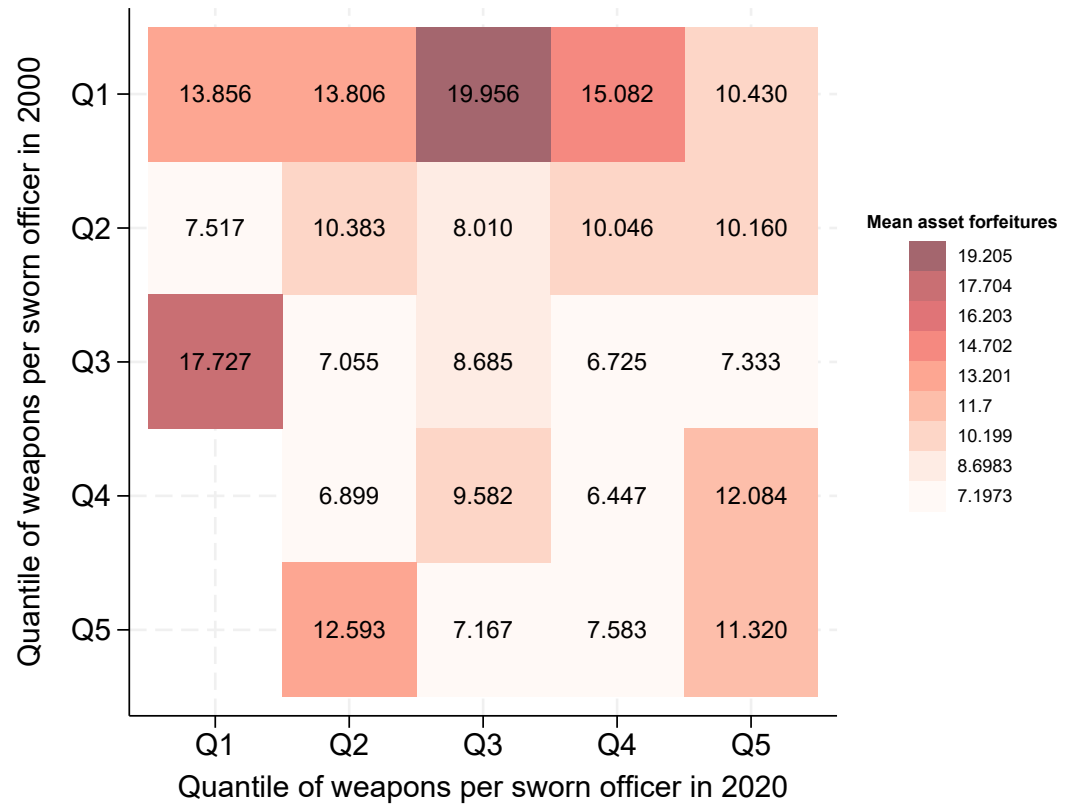


Figure C.12: Average real dollars per capita in assets seized as police departments transition in the stock of weapons.

Notes: The figure plots the mean assets forfeiture between 1977 and 1997 as departments transition between the quintile of the distribution of 1033 weapons in an initial period (vertical axis) into a future period (horizontal axis).

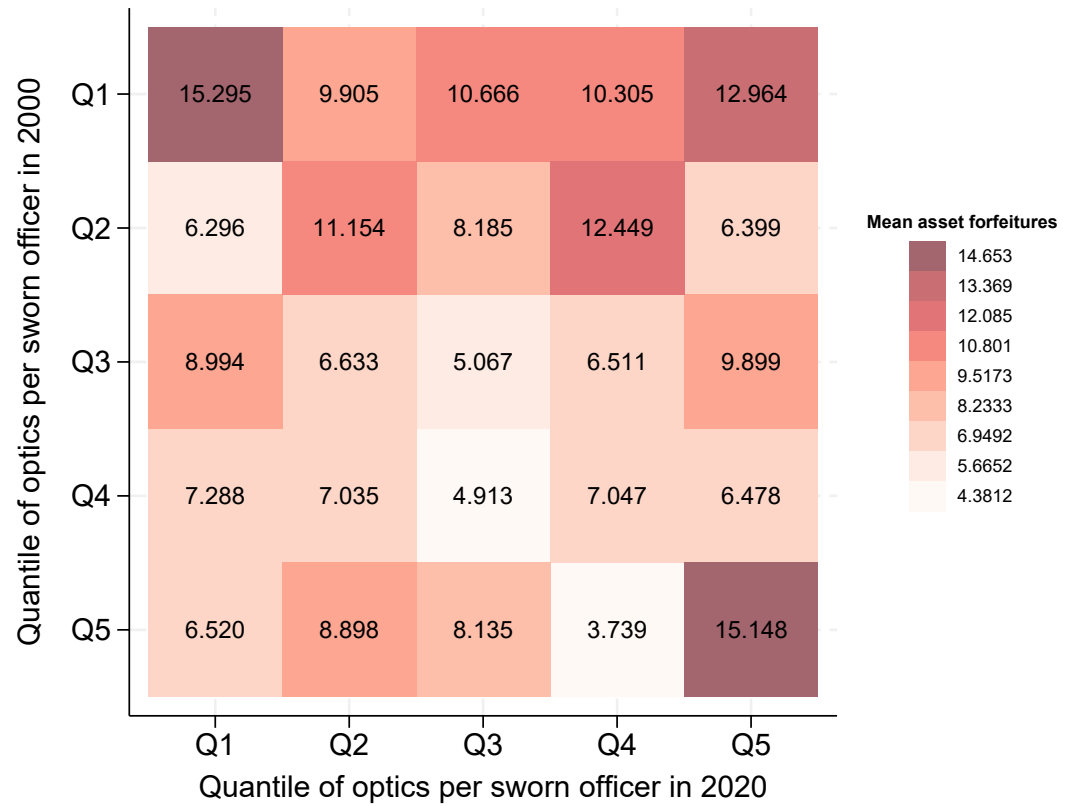


Figure C.13: Average real dollars per capita in assets seized as police departments transition in the stock of optics.

Notes: The figure plots the mean assets forfeiture between 1977 and 1997 as departments transition between the quintile of the distribution of 1033 optics in an initial period (vertical axis) into a future period (horizontal axis).

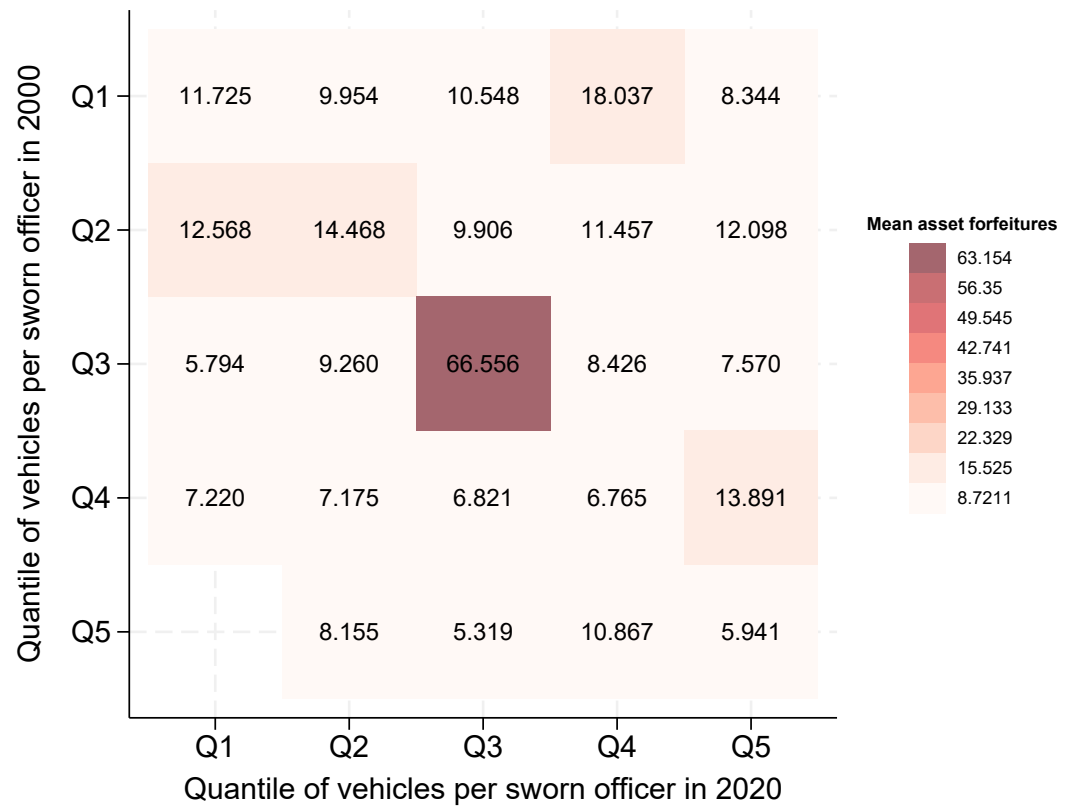


Figure C.14: Average real dollars per capita in assets seized as police departments transition in the stock of vehicles.

Notes: The figure plots the mean assets forfeiture between 1977 and 1997 as departments transition between the quintile of the distribution of 1033 vehicles in an initial period (vertical axis) into a future period (horizontal axis).

C.0.5 Lagged under-reports

Descriptive evidence in Section 4.4.4 is complemented with two statistical approximations to document whether under-reports is correlated with the probability of take-up. First, we explore a lagged conditional regression that captures whether the stock of items in current period is correlated with under-reports in previous periods. The outcome is the number of items acquired by a police department, as we aim to explain whether the underlying reporting culture in police departments leads to a larger stock of items. The process is a function of lagged under-reports:

$$\text{items}_{it} = \sum_{l=1}^L \eta_l \Delta 1(\text{report}_{il} > 0) + \pi X_{it} + \omega_i + \omega_t + \varepsilon_{it} \quad (\text{C.1})$$

The parameter η_l measures the statistical association between under-reporting in lagged year l and the accumulation of items in time t . We acknowledge that this approach has important limitations. Equation C.1 examines whether police departments accumulate more military gear if they have pre-existing instances of misreporting. However, as emphasized by Equation 4.2, there is a simultaneity bias between both behavioral choices.

Table C.9: Previous under-reports and the stock of 1033 items conditional on former UCR respondents.

	Total	Weapons	Vehicles	Optics	Clothes	Others
L.Under-Report	25.243 (28.735)	1.692 (3.121)	0.240*** (0.064)	2.763* (1.445)	10.419 (6.504)	10.130 (20.340)
L2.Under-Report	19.102 (29.282)	2.155 (3.008)	0.154** (0.065)	3.496** (1.433)	7.226 (6.545)	6.070 (21.116)
L3.Under-Report	21.652 (30.438)	0.897 (3.190)	0.166** (0.069)	3.862*** (1.476)	5.364 (6.697)	11.365 (21.891)
L4.Under-Report	23.047 (31.404)	0.545 (3.262)	0.110 (0.071)	5.186*** (1.384)	2.831 (6.963)	14.375 (22.376)
L5.Under-Report	26.082 (33.651)	-0.144 (3.547)	0.112 (0.072)	5.445*** (1.561)	2.201 (7.456)	18.468 (23.947)
Obs	109102	109102	109102	109102	109102	109102
R2	0.40	0.59	0.44	0.57	0.37	0.39
F-statistic	11.41	12.36	9.17	13.80	11.77	9.41
F $\eta = 0$	2.07	0.32	24.34	29.83	2.77	1.11
p-value $\eta = 0$	0.15	0.57	0.00	0.00	0.10	0.29

Notes: The table shows the OLS coefficients with items acquired by departments as a dependent variable. The independent variable are lagged indicators on whether the police department under-reported statistics. The sample is restricted to police departments that have responded to UCR in some instance after 2000. Standard F-statistics are computed and the joint significance test that all coefficients of lagged variables are equal to zero. The regression controls for population size, labor force, unemployment, per capita personal income, the police department's total budget, the number of full time sworn officers, the total number of employees, year fixed effects, and agency fixed effects. The regression also controls for normalized boots on the ground interacted with the distance to a DLA storage facility, and the interaction with distance to a military base. Clustered standard errors are computed at the commuting zone and year interaction.

Table C.10: Previous under-reports and the stock of 1033 items.

	Total	Weapons	Vehicles	Optics	Clothes	Others
L.Under-Report	30.648 (24.773)	2.693 (2.691)	0.287*** (0.059)	2.855** (1.241)	11.293** (5.644)	13.521 (17.502)
L2.Under-Report	25.289 (25.456)	3.069 (2.593)	0.205*** (0.060)	3.477*** (1.240)	8.754 (5.712)	9.783 (18.407)
L3.Under-Report	26.241 (25.926)	1.738 (2.741)	0.193*** (0.063)	3.792*** (1.272)	6.337 (5.725)	14.179 (18.624)
L4.Under-Report	26.793 (26.793)	1.656 (2.799)	0.152** (0.064)	4.863*** (1.195)	3.993 (5.965)	16.129 (19.064)
L5.Under-Report	28.999 (28.956)	0.943 (3.045)	0.158** (0.065)	5.072*** (1.343)	3.397 (6.412)	19.429 (20.612)
Obs	179094	179094	179094	179094	179094	179094
R2	0.40	0.59	0.44	0.57	0.37	0.39
F-statistic	12.03	14.76	12.12	16.00	13.88	9.45
F $\eta = 0$	3.90	1.67	45.45	37.16	5.24	2.11
p-value $\eta = 0$	0.05	0.20	0.00	0.00	0.02	0.15

Notes: The table shows the statistical association between items acquired and lagged indicators on whether the police department under-reported homicides in UCR. Standard F-statistics are computed, and a joint significance test whether all η coefficients of lagged variables are jointly equal to zero. The regression controls for population size, labor force, unemployment, per capita personal income, the police department's total budget, the number of full-time sworn officers, the total number of employees, year fixed effects, and agency fixed effects. The regression also controls for normalized boots on the ground interacted with the distance to a DLA storage facility, and the interaction with distance to a military base. Clustered standard errors are computed at the commuting zone and year interaction.

C.0.6 Event studies

Lastly, as a last source of evidence, we propose an event-study specification as follows:

$$1(\Delta \text{reports}_{it} > 0) = \alpha_i + \alpha_t + \sum_{\tau=-q}^{-2} \theta_{\tau} 1(\text{items}_{it} \geq 0) + \sum_{\tau=0}^p \vartheta_{\tau} 1(\text{items}_{it} \geq 0) + e_{it} \quad (\text{C.2})$$

Treatment is defined as a binary indicator that the items' stock is above zero $1(\text{items}_{it} \geq 0)$ indicating the before and after first shipment. Equation C.2 allows for testing adverse selection as a violation in the pre-trends assumption. Failure to hold the parallel trends assumption, observed in the lagged coefficient θ_{τ} , suggests that trends in reporting behavior of police departments systematically differ before 1033 take-up.

One feature of shipment choice in Equation C.2 is that variation in the variation in the timing of shipments. As a result, the control group in a two-way fixed effects specification may display biases related to inadequate comparisons. For example, the control group may include future-takers among police departments. We estimate results from a Two Way Fixed Effects (TWFE) estimator and the weighted dynamic effects estimator in Sun and Abraham (2021), robust to heterogeneous timing.

This empirical approach continues to suffer from simultaneity bias. Evidence of adverse selection is therefore suggestive and only serves to provide statistical back up to non-parametric evidence that departments selecting not to respond to UCR tend to acquire more weapons and vehicles from 1033 while 1033 participants may also be larger police departments with a larger likelihood to engage in violence against civilians. In that sense, police departments participating in section 1033 may have underlying differences in police culture and procedures before the transfer of equipment.

Event studies after first shipment

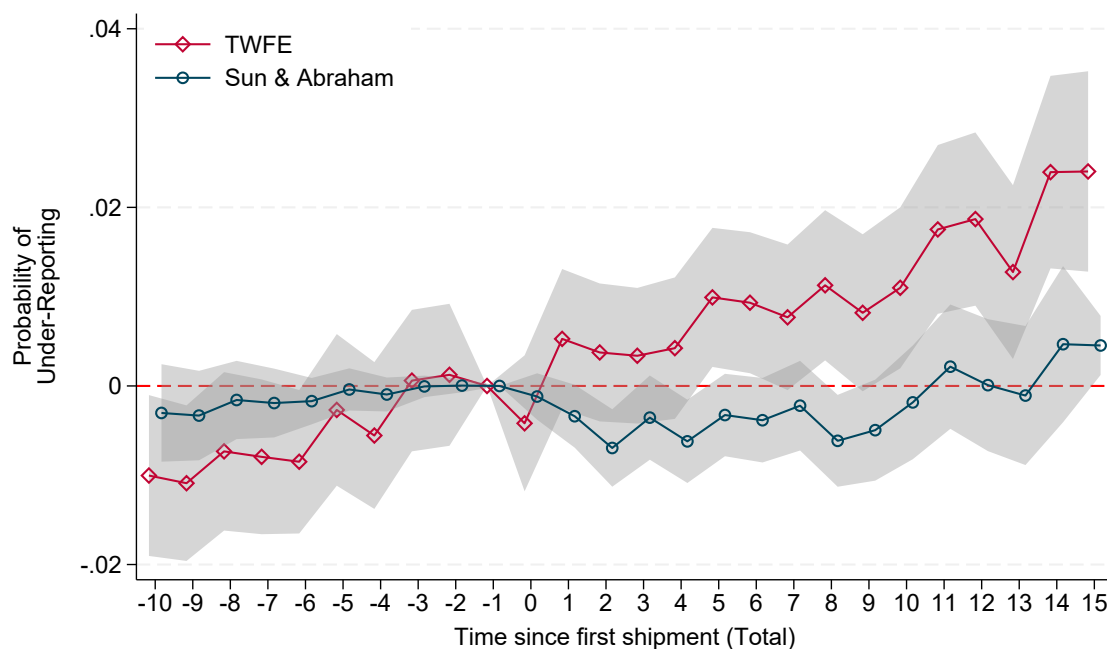


Figure C.15: The probability of under-reporting after acquiring 1033 items (All items).

Notes: The figure represents an event study on the time since the first shipment and the probability of under-reporting homicides by police. The panel presents coefficients after the first shipment of any 1033 item. In all instances, the control group excludes police departments that participate in the 1033 program. The red coefficients are computed with a two-way fixed effect estimator, and the Sun & Abraham estimator is in blue (Sun and Abraham, 2021). The regression controls for population size, labor force, unemployment, per capita personal income, the police department's total budget, the number of full-time sworn officers, the total number of employees, year fixed effects, and agency fixed effects. Clustered standard errors are computed at the commuting zone and year interaction. The graph plots the 95 percent confidence interval.

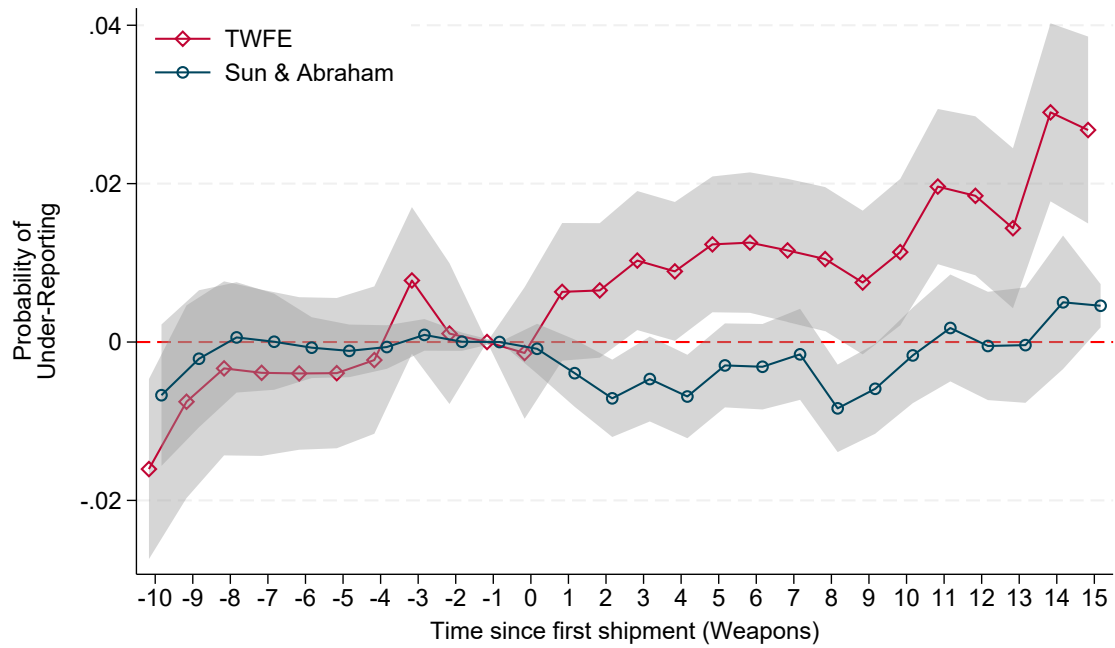


Figure C.16: The probability of under-reporting after acquiring 1033 weapons.

Notes: The figure represents an event study on the time since the first shipment and the probability of under-reporting homicides by police. The panel presents coefficients after the first shipment of weapons. In all instances, the control group excludes police departments that participate in the 1033 program. The red coefficients are computed with a two-way fixed effect estimator, and the Sun & Abraham estimator is in blue (Sun and Abraham, 2021). The regression controls for population size, labor force, unemployment, per capita personal income, the police department's total budget, the number of full-time sworn officers, the total number of employees, year fixed effects, and agency fixed effects. Clustered standard errors are computed at the commuting zone and year interaction. The graph plots the 95 percent confidence interval.

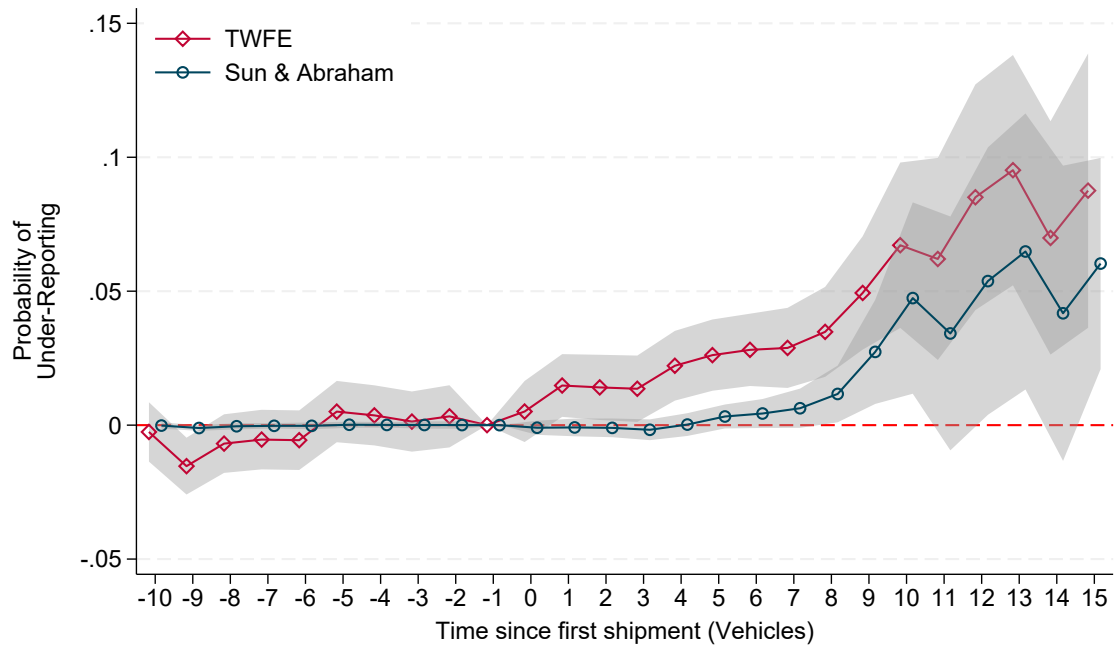


Figure C.17: The probability of under-reporting after acquiring 1033 vehicles.

Notes: The figure represents an event study on the time since the first shipment and the probability of under-reporting homicides by police. The panel presents coefficients after the first shipment of vehicles. In all instances, the control group excludes police departments that participate in the 1033 program. The red coefficients are computed with a two-way fixed effect estimator, and the Sun & Abraham estimator is in blue (Sun and Abraham, 2021). The regression controls for population size, labor force, unemployment, per capita personal income, the police department's total budget, the number of full-time sworn officers, the total number of employees, year fixed effects, and agency fixed effects. Clustered standard errors are computed at the commuting zone and year interaction. The graph plots the 95 percent confidence interval.

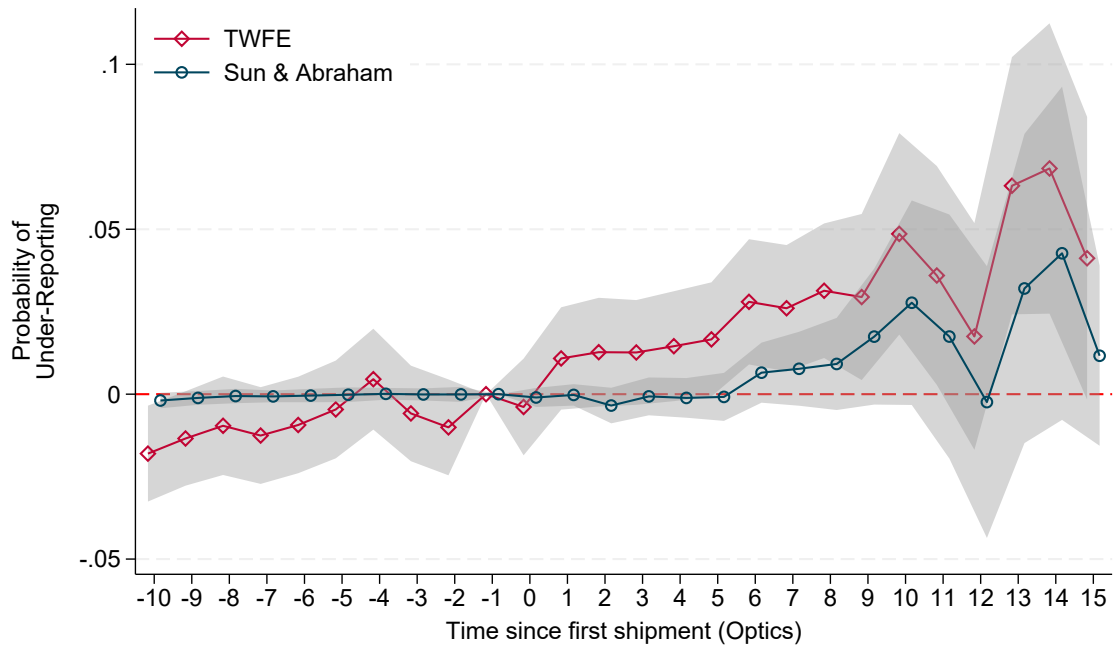


Figure C.18: The probability of under-reporting after acquiring 1033 optics.

Notes: The figure represents an event study on the time since the first shipment and the probability of under-reporting homicides by police. The panel presents coefficients after the first shipment of optics. In all instances, the control group excludes police departments that participate in the 1033 program. The red coefficients are computed with a two-way fixed effect estimator, and the Sun & Abraham estimator is in blue (Sun and Abraham, 2021). The regression controls for population size, labor force, unemployment, per capita personal income, the police department's total budget, the number of full-time sworn officers, the total number of employees, year fixed effects, and agency fixed effects. Clustered standard errors are computed at the commuting zone and year interaction. The graph plots the 95 percent confidence interval.

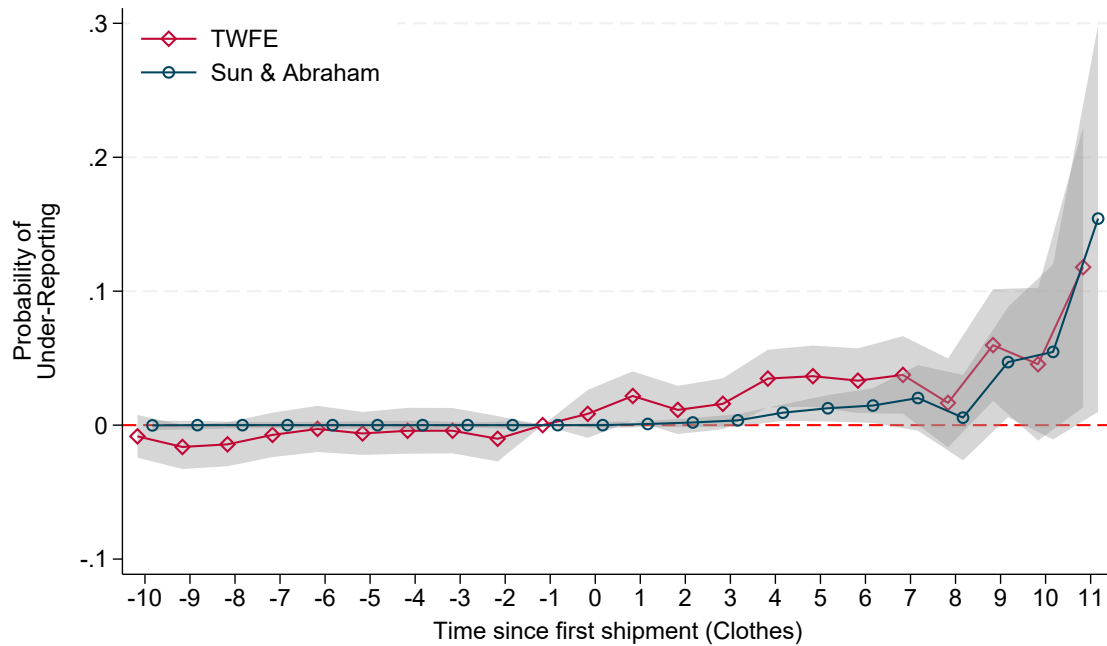


Figure C.19: The probability of under-reporting after acquiring 1033 clothes.

Notes: The figure represents an event study on the time since the first shipment and the probability of under-reporting homicides by police. The panel presents coefficients after the first shipment of clothes. In all instances, the control group excludes police departments that participate in the 1033 program. The red coefficients are computed with a two-way fixed effect estimator, and the Sun & Abraham estimator is in blue (Sun and Abraham, 2021). The regression controls for population size, labor force, unemployment, per capita personal income, the police department's total budget, the number of full-time sworn officers, the total number of employees, year fixed effects, and agency fixed effects. Clustered standard errors are computed at the commuting zone and year interaction. The graph plots the 95 percent confidence interval.

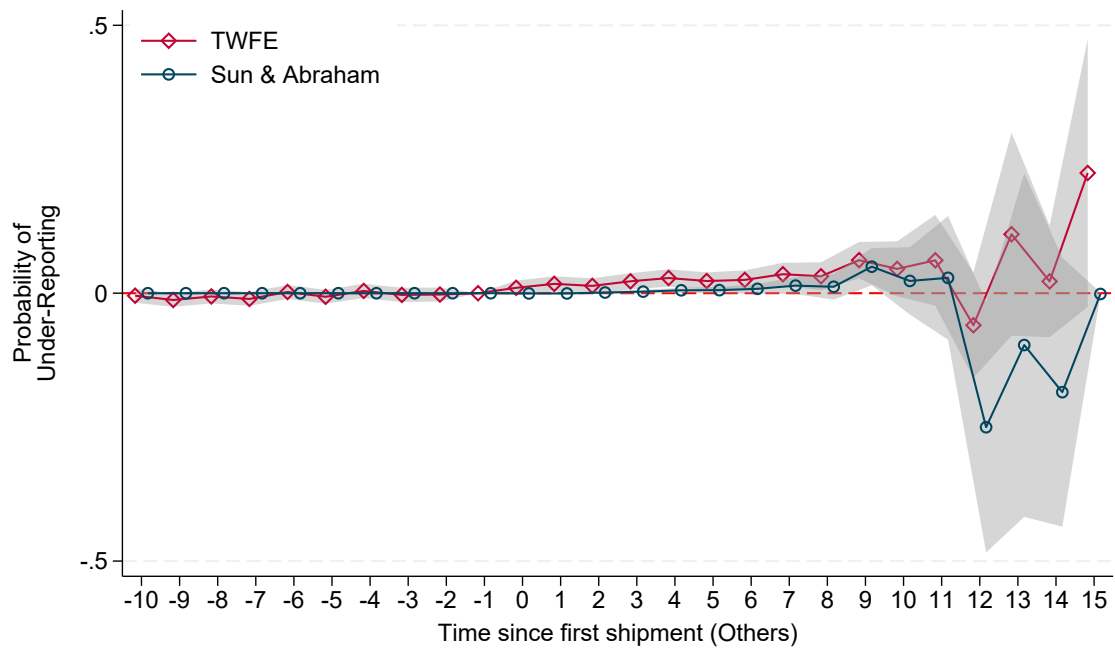


Figure C.20: The probability of under-reporting after acquiring 1033 other items.

Notes: The figure represents an event study on the time since the first shipment and the probability of under-reporting homicides by police. The panel presents coefficients after the first shipment of all others. In all instances, the control group excludes police departments that participate in the 1033 program. The red coefficients are computed with a two-way fixed effect estimator, and the Sun & Abraham estimator is in blue (Sun and Abraham, 2021). The regression controls for population size, labor force, unemployment, per capita personal income, the police department's total budget, the number of full-time sworn officers, the total number of employees, year fixed effects, and agency fixed effects. Clustered standard errors are computed at the commuting zone and year interaction. The graph plots the 95 percent confidence interval.

Event studies after the largest shipment

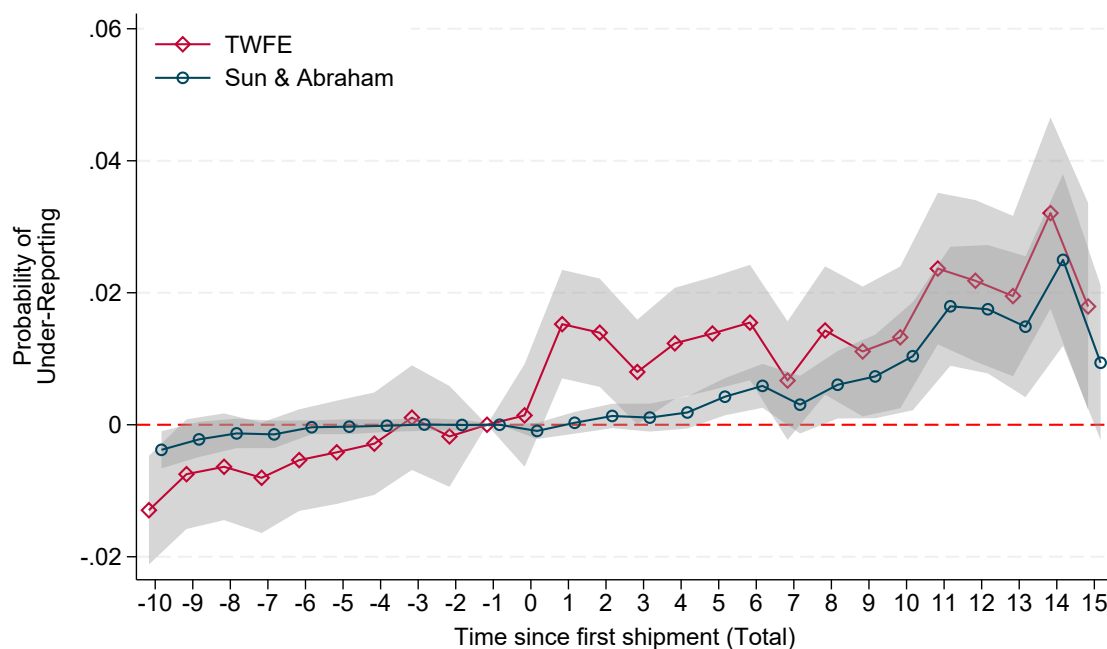


Figure C.21: The probability of under-reporting after acquiring 1033 items (All items).

Notes: The figure represents an event study on the time since the largest shipment and the probability of under-reporting homicides by police. The panel presents coefficients after the largest shipment of any 1033 item. In all instances, the control group excludes police departments that participate in the 1033 program. The red coefficients are computed with a two-way fixed effect estimator, and the Sun & Abraham estimator is in blue (Sun and Abraham, 2021). The regression controls for population size, labor force, unemployment, per capita personal income, the police department's total budget, the number of full-time sworn officers, the total number of employees, year fixed effects, and agency fixed effects. Clustered standard errors are computed at the commuting zone and year interaction. The graph plots the 95 percent confidence interval.

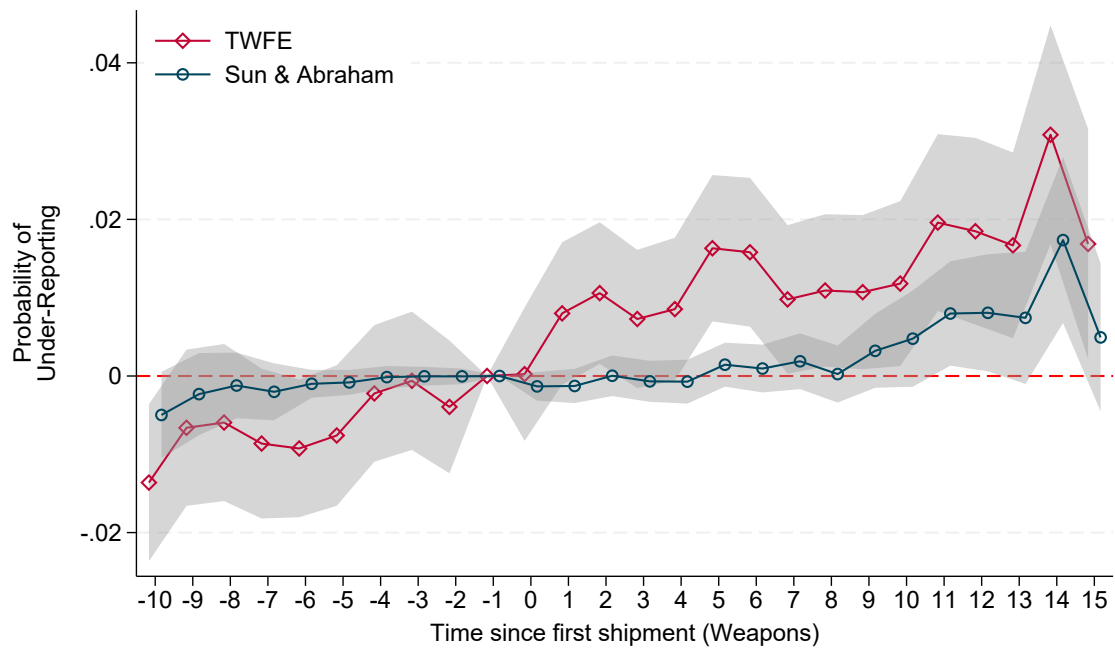


Figure C.22: The probability of under-reporting after acquiring 1033 weapons.

Notes: The figure represents an event study on the time since the largest shipment and the probability of under-reporting homicides by police. The panel presents coefficients after the largest shipment of weapons. In all instances, the control group excludes police departments that participate in the 1033 program. The red coefficients are computed with a two-way fixed effect estimator, and the Sun & Abraham estimator is in blue (Sun and Abraham, 2021). The regression controls for population size, labor force, unemployment, per capita personal income, the police department's total budget, the number of full-time sworn officers, the total number of employees, year fixed effects, and agency fixed effects. Clustered standard errors are computed at the commuting zone and year interaction. The graph plots the 95 percent confidence interval.

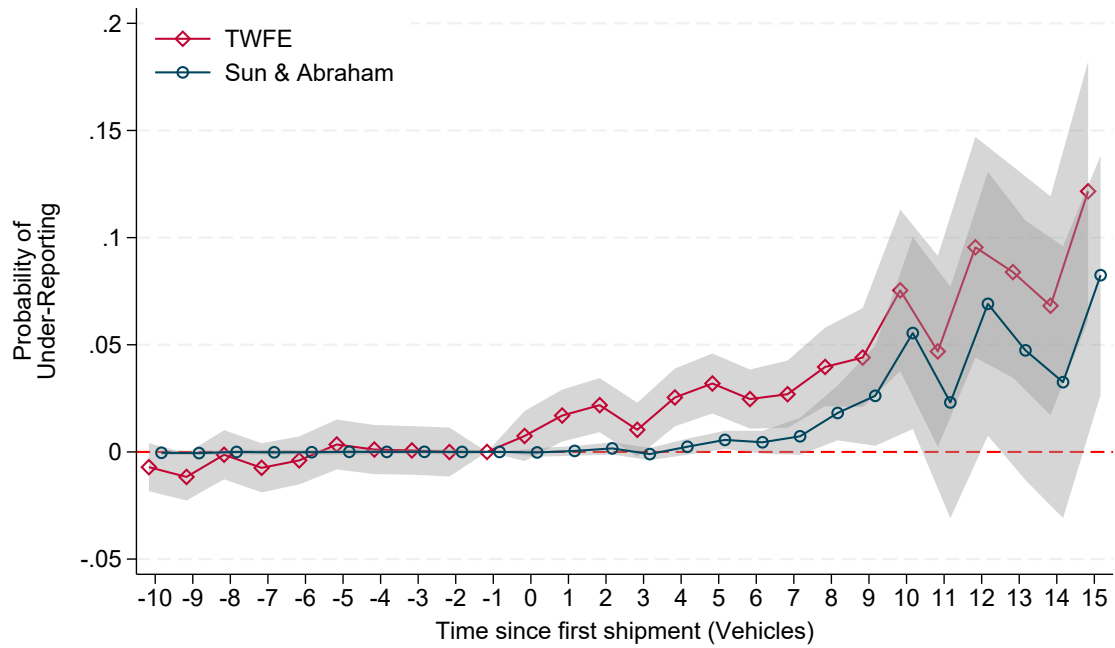


Figure C.23: The probability of under-reporting after acquiring 1033 vehicles.

Notes: The figure represents an event study on the time since the largest shipment and the probability of under-reporting homicides by police. The panel presents coefficients after the largest shipment of vehicles. In all instances, the control group excludes police departments that participate in the 1033 program. The red coefficients are computed with a two-way fixed effect estimator, and the Sun & Abraham estimator is in blue (Sun and Abraham, 2021). The regression controls for population size, labor force, unemployment, per capita personal income, the police department's total budget, the number of full-time sworn officers, the total number of employees, year fixed effects, and agency fixed effects. Clustered standard errors are computed at the commuting zone and year interaction. The graph plots the 95 percent confidence interval.

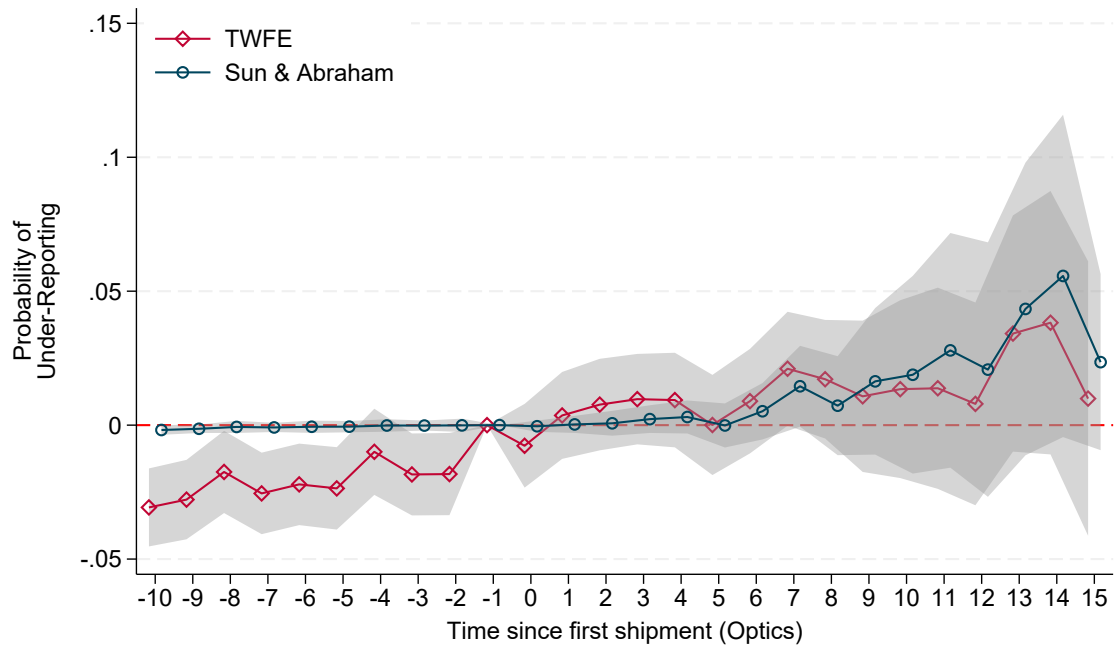


Figure C.24: The probability of under-reporting after acquiring 1033 optics.

Notes: The figure represents an event study on the time since the largest shipment and the probability of under-reporting homicides by police. The panel presents coefficients after the largest shipment of optics. In all instances, the control group excludes police departments that participate in the 1033 program. The red coefficients are computed with a two-way fixed effect estimator, and the Sun & Abraham estimator is in blue (Sun and Abraham, 2021). The regression controls for population size, labor force, unemployment, per capita personal income, the police department's total budget, the number of full-time sworn officers, the total number of employees, year fixed effects, and agency fixed effects. Clustered standard errors are computed at the commuting zone and year interaction. The graph plots the 95 percent confidence interval.

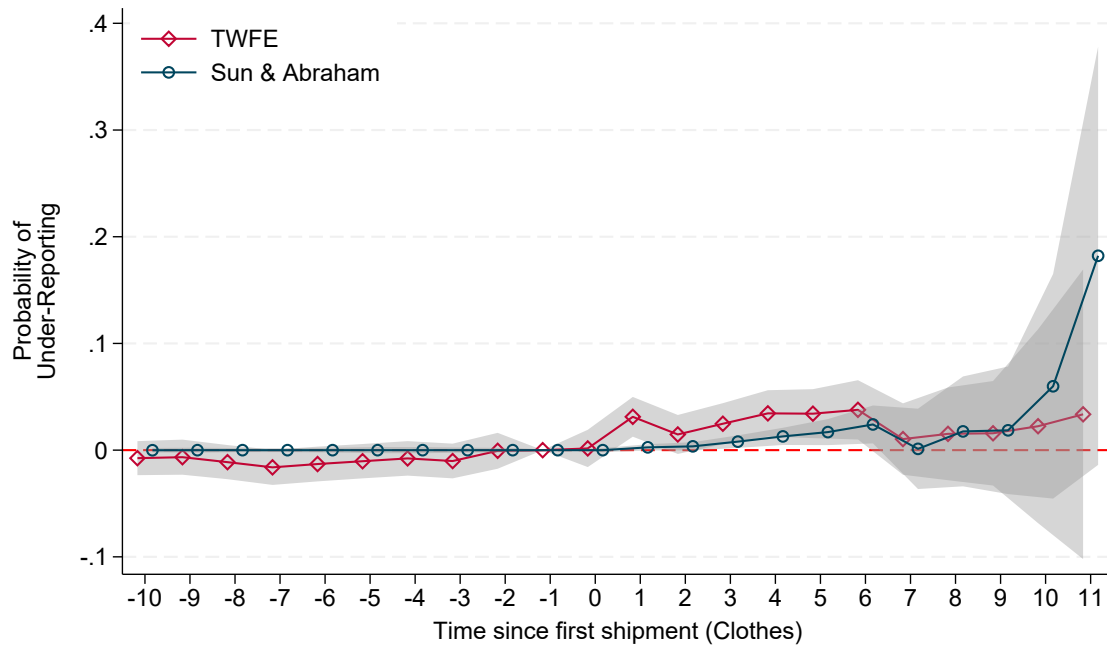


Figure C.25: The probability of under-reporting after acquiring 1033 clothes.

Notes: The figure represents an event study on the time since the largest shipment and the probability of under-reporting homicides by police. The panel presents coefficients after the largest shipment of clothes. In all instances, the control group excludes police departments that participate in the 1033 program. The red coefficients are computed with a two-way fixed effect estimator, and the Sun & Abraham estimator is in blue (Sun and Abraham, 2021). The regression controls for population size, labor force, unemployment, per capita personal income, the police department's total budget, the number of full-time sworn officers, the total number of employees, year fixed effects, and agency fixed effects. Clustered standard errors are computed at the commuting zone and year interaction. The graph plots the 95 percent confidence interval.

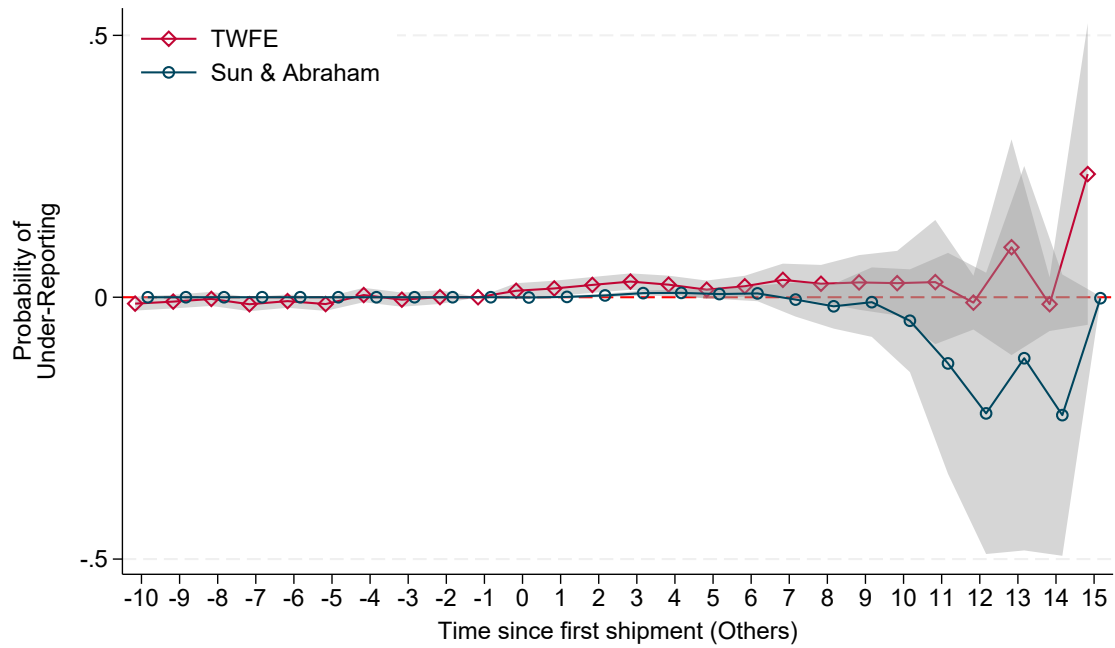


Figure C.26: The probability of under-reporting after acquiring 1033 other items.

Notes: The figure represents an event study on the time since the largest shipment and the probability of under-reporting homicides by police. The panel presents coefficients after the largest shipment of all other items. In all instances, the control group excludes police departments that participate in the 1033 program. The red coefficients are computed with a two-way fixed effect estimator, and the Sun & Abraham estimator is in blue (Sun and Abraham, 2021). The regression controls for population size, labor force, unemployment, per capita personal income, the police department's total budget, the number of full-time sworn officers, the total number of employees, year fixed effects, and agency fixed effects. Clustered standard errors are computed at the commuting zone and year interaction. The graph plots the 95 percent confidence interval.

Vita

Diego A. Guerrero was born in Caracas, Venezuela. He earned a Bachelor's degree in Liberal Studies at the Universidad Metropolitana where he taught public policy and economic development as a Lecturer soon after graduation. His career in public policy and economic analysis started working as a consultant and researcher at the IESA Business School in Caracas. At IESA, he developed expertise in the political economy of oil revenue, oil and gas taxation, and social welfare programs funded with oil windfalls. Later, Diego earned a M.Sc. in Applied Economics at St. Cloud State University. In 2019, he chose to attend the University of Tennessee and pursue a doctoral degree in Economics, specializing in Public Economics and Labor Economics. His research centers on the analysis of intergenerational mobility, poverty alleviation, human capital formation, and labor markets. His dissertation investigates the effects of government programs, particularly the Earned Income Tax Credit, on the adult income of children and policy interactions with maternal labor supply. As a Graduate Research Assistant at the Boyd Center for Business and Economic Research, Diego evaluates the effectiveness of K-12 and college subsidies aimed at improving academic achievement and labor outcomes among Tennesseans. He has also published research on elementary school attendance in developing economies. With nearly a decade of teaching experience, he is a dedicated educator and committed to improve policy effectiveness through collaboration with state and local governments and multilateral organizations. After graduation, he will begin a new position as a Lead Economist at Northwestern University's School of Education and Social Policy.