THE UNIVERSITY OF CHICAGO

ESSAYS ON POVERTY AND SAFETY-NET PROGRAMS

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE IRVING B. HARRIS

GRADUATE SCHOOL OF PUBLIC POLICY STUDIES

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

BY

DANKE DEREK WU

CHICAGO, ILLINOIS

DECEMBER 2021

For Mom, Dad, & Delia

Table of Contents

LIST OF FIGURES vii

LIST OF TABLES viii

ACKNOWLEDGMENTS ix

ABSTRACT x

1 CERTIFICATION AND RECERTIFICATION IN WELFARE PROGRAMS: WHAT HAPPENS WHEN AUTOMATION GOES WRONG? 1
- 1.1 Introduction 1
- 1.2 Background on Programs and Natural Experiment 7
  - 1.2.1 Description of SNAP, TANF, and Medicaid Programs 7
  - 1.2.2 Application and Recertification Process 11
  - 1.2.3 Automating the Administration of Welfare Services in Indiana 13
- 1.3 Data Sources and Summary Statistics 18
  - 1.3.1 Data Sources 18
  - 1.3.2 Characteristics of Treated and Untreated Counties in Indiana 21
- 1.4 Empirical Strategy 23
  - 1.4.1 Regression Specification 23
  - 1.4.2 Raw Trends in Welfare Receipt 25
- 1.5 Effects of IBM Automation on Enrollment 27
  - 1.5.1 Regression Specification 27
  - 1.5.2 Disaggregating Enrollment Effects by Entry and Exit 31
- 1.6 Who is Screened Out by IBM Automation? 34
  - 1.6.1 Heterogeneous Effects by County-Level Characteristics 35
  - 1.6.2 Effects on Targeting Efficiency 38
- 1.7 Additional Analyses and Robustness Checks 45
  - 1.7.1 Differences in Short-Run Enrollment Effects Across Programs 46
  - 1.7.2 Mechanisms Underlying Long-Run Enrollment Declines 49
  - 1.7.3 Robustness Checks 54
- 1.8 Conclusion 63

2 DOES GEOGRAPHICALLY ADJUSTING POVERTY THRESHOLDS IMPROVE POVERTY MEASUREMENT AND PROGRAM TARGETING? 67
- 2.1 Introduction 67
- 2.2 Theory and Previous Literature 72
  - 2.2.1 Simple Model of Local Poverty Thresholds and Prices 72
  - 2.2.2 Literature on Prices and Amenities Across Geography 74
  - 2.2.3 Geographic Price Indices 78
  - 2.2.4 Evidence on Association with Deprivation 79
- 2.3 Data and Poverty Measures 81
  - 2.3.1 Data 81

2.3.2 Poverty Measures ... 83
2.4 Summary Statistics ... 88
2.4.1 Changes in Poverty Rates by Geography ... 88
2.4.2 Changes in Poverty Rates for Other Groups ... 92
2.5 Methods: Deprivation Measures ... 95
2.5.1 Measures of Material Well-Being ... 95
2.5.2 Analytical Methods for Comparing Well-Being Across Poverty Measures ... 98
2.5.3 Shares and Counts by Geographic Poverty Category ... 100
2.6 Main Regression Results ... 102
2.6.1 Estimates for the Supplemental Poverty Measure ... 102
2.6.2 Estimates for the Comprehensive Income Poverty Measure ... 110
2.7 Extensions and Robustness Checks ... 115
2.7.1 Scaling the Geographic Adjustment Factor ... 115
2.7.2 Using Regional Price Parities Rather Than the MRI Price Index ... 120
2.7.3 Deep and Near Poverty ... 125
2.7.4 Additional Analyses ... 130
2.8 Empirical Explanations for Results ... 134
2.9 Conclusion ... 137
2.A Appendix ... 140
2.A.1 Detailed Description of Geographic Price Indices ... 140
2.A.2 Detailed Description of Data Sources ... 142
2.A.3 Methods for Constructing Poverty Measures ... 146

3 THE USE AND MISUSE OF INCOME DATA AND EXTREME POVERTY IN THE UNITED STATES ... 156
3.1 Introduction ... 156
3.2 Literature ... 162
3.2.1 Past Claims of Extreme Poverty and Conflicting Evidence ... 162
3.2.2 Why are Extreme Poverty Rates from Survey-Reported Cash So High? ... 166
3.3 Data ... 169
3.3.1 Survey Data ... 169
3.3.2 Administrative Data ... 170
3.3.3 Linking Survey and Administrative Data ... 173
3.4 Methods ... 174
3.4.1 Defining Extreme Poverty and Sample Construction ... 174
3.4.2 Corrections and Adjustments Using Survey and Administrative Data ... 176
3.4.3 Validating the Survey-Only Adjustments ... 181
3.5 Main Results ... 183
3.5.1 Extreme Poverty After Adjustments ... 183
3.5.2 Corrections and Adjustments Bringing in the Administrative Data First ... 186
3.5.3 Extreme Poverty by Household Type ... 189
3.5.4 Distribution of Household Types ... 192
3.6 Validation of Survey-Only Adjustments ... 194
3.6.1 Administrative Income Data ... 194
3.6.2 Survey-Reported Material Well-Being ... 197
3.6.3 Survey-Reported Demographics ... 203

3.7 Comparison to Current Population Survey Results 204
3.7.1 Data and Methods 204
3.7.2 Results 206
3.8 Robustness Checks and Caveats 209
3.8.1 Robustness Checks 209
3.8.2 Caveats 213
3.9 Conclusion 217
3.A Appendix 220
3.A.1 Survey-Only Adjustments in the SIPP 220
3.A.2 Survey-Only Adjustments in the CPS 226
3.A.3 Inverse Probability Weighting Adjustment for Incomplete PIK Linkage 228
3.A.4 Household Types 231

REFERENCES 233

List of Figures

1.1 Baseline Hazard Rates and Survival Probabilities by Spell Length................................ 12
1.2 Rollout of IBM's Automated System Across Indiana Counties 15
1.3 Raw Trends in Welfare Receipt Rates.. 26
1.4 Dynamic Treatment Effects on Overall Enrollment 28
1.5 Dynamic Treatment Effects on Enrollment for Existing Recipients................................ 47
1.6 Dynamic Treatment Effects on Enrollment After Excluding Groups 51
1.7 Synthetic Control Estimates for Log SNAP Receipt.. 62

2.1 Proportions of OPM Threshold Used to Anchor Poverty Rates at Official Levels 87
2.2a Percent Change in Poverty Rates After Geographic Adjustments (State)........................ 89
2.2b Percent Change in Poverty Rates After Geographic Adjustments (CBSA) 90
2.3 Regression Estimates of Well-Being by Scaling Factor (SPM) 117
2.4 Regression Estimates of Well-Being by Scaling Factor (CIPM).................................. 118
2.5 Summary of Geographic Adjustment Effects on Well-Being by Model 131

3.1 Share of Reported Cash Extreme Poor Households Raised Above Thresholds 188
3.2 Household Type Distribution of Extreme Poor Subgroups After Adjustments 193
3.3 Share of Households in Extreme Poor Subgroups Raised Above Thresholds 196
3.4 Mean Number of Material Hardships for Extreme Poor Subgroups.............................. 198
3.5 Share of Households with Any Material Hardships for Extreme Poor Subgroups......... 200
3.6 Mean Number of Home Problems for Extreme Poor Subgroups 201
3.7 Mean Number of Appliances Owned by Extreme Poor Subgroups.............................. 201
3.8 Expenditures on Four Largest Transfer Programs Not in Administrative Data 215

List of Tables

1.1 Characteristics of SNAP, TANF, and Medicaid Recipients (2005-2007) 9
1.2 Characteristics of Untreated and Treated Counties (2007) 22
1.3 Treatment Effects on Overall Program Enrollment 30
1.4 Treatment Effects on Monthly Entry and Exit Rates 33
1.5 Treatment Effects on Enrollment by County-Level Characteristics 37
1.6a Treatment Effects on the Characteristics of SNAP Recipients (Targeting) 41
1.6b Treatment Effects on the Characteristics of TANF Recipients (Targeting) 42
1.6c Treatment Effects on the Characteristics of Medicaid Recipients (Targeting) 43
1.7 Treatment Effects on Enrollment Among Existing Recipients 48
1.8 Treatment Effects on Enrollment After Excluding Certain Types of Recipients 52
1.9 Treatment Effects on Enrollment (Different Specifications) 56
1.10 Treatment Effects for Automated vs. Non-Automated Programs 57
1.11 Treatment Effects Using Alternative Difference-in-Differences Estimators 59

2.1a Poverty Rates with and without Geographic Adjustments (SPM) 93
2.1b Poverty Rates with and without Geographic Adjustments (CIPM) 94
2.2 Shares and Counts by Geographic Poverty Category 101
2.3a Regression Estimates of Well-Being on Geographic-Only Poor (SPM) 103
2.3b Regression Estimates of Well-Being on Geographic-Only Poor (SPM) 111
2.4 Shares and Counts by Geographic Poverty Category (using RPP Adjustments) 121
2.5a Regression Estimates of Well-Being on Geographic-Only Poor Using RPP (SPM) 122
2.5b Regression Estimates of Well-Being on Geographic-Only Poor Using RPP (CIPM) 124
2.6 Shares and Counts by Geographic Poverty Category (Deep and Near Poverty) 126
2.7a Regression Estimates of Well-Being on Geographic-Only Deep & Near Poor (SPM) 127
2.7b Regression Estimates of Well-Being on Geographic-Only Deep & Near Poor (CIPM) 128
2.8a Elasticities of Wage and Non-Wage Income with Respect to Price Indices 134
2.8b Elasticities of Per-Capita State Spending with Respect to Price Indices 136

3.1 Administrative Data Sources 171
3.2a Percentage of Households in Extreme Poverty 184
3.2b Percentage of Individuals in Extreme Poverty 185
3.3 Income Receipt Rates for Remaining Extreme Poor Households After Adjustments 191
3.4 Comparison of CPS and SIPP Extreme Poverty Estimates for All Households 207
3.5 Robustness Checks (Percentage of Households in Extreme Poverty) 210

Acknowledgments

I am indebted to my committee – Bruce Meyer, Dan Black, and Manasi Deshpande – for their generous and continuous mentorship. Bruce gave me a unique opportunity to work on the Comprehensive Income Dataset and to dive into research on poverty and safety net programs, and he has taught me a great deal about the nuts and bolts of conducting careful and clear research. Ever since the first year of my PhD, Dan has always been a fountain of creative ideas and a great source of encouragement, providing helpful feedback at any stage of the research process. Manasi – through her class on social insurance – helped to inspire my interest in studying the effects of administrative burdens, and has never failed to provide insightful and thought-provoking feedback.

I am also grateful to many other professors at the University of Chicago from whom I took illuminating classes and with whom I had valuable discussions. My professional and personal well-being were further buoyed by the many kind and supportive graduate students whom I met, learned from, and befriended – I am particularly thankful to the members of my first-year study group (Mariella Gonzales, Justin Holz, Seunghoon Lee, and Miguel Morales), the Harris Student Working Group, and my roommate for many years Neil Cholli. Numerous staff members – especially Cynthia Cook-Conley – were instrumental in steering me in the right direction throughout my graduate school experience.

Finally, I would like to thank my partner Breanna, my sister Delia, and Mom and Dad – for their love and support and for always encouraging me to make a positive impact on this world.

Abstract

This dissertation consists of three chapters that examine policy questions relevant to the low-income population in the United States. In the first chapter "Certification and Recertification in Welfare Programs: What Happens When Automation Goes Wrong?" (joint with Bruce Meyer), I examine how administrative burdens influence enrollment in different welfare programs and who is screened out at a given stage. I study the impacts of complications arising from the automation of welfare services, leveraging a unique natural experiment in Indiana in which the IBM Corporation remotely processed applications for two-thirds of all counties. Using linked administrative records covering nearly 3 million program recipients, I find that SNAP, TANF, and Medicaid enrollment fall by 15%, 24%, and 4% one year after automation, with these heterogeneous declines largely attributable to cross-program differences in recertification costs. Earlier-treated and higher-poverty counties experience larger declines in welfare receipt. More needy individuals are screened out at exit while less needy individuals are screened out at entry, a novel distinction that would be missed by typical measures of targeting which focus on average changes overall or at a single margin. The decline in Medicaid enrollment exhibits considerable permanence after IBM's automated system was disbanded, suggesting potential long-term consequences of increases in administrative burdens.

In the second chapter "Does Geographically Adjusting Poverty Thresholds Improve Poverty Measurement and Program Targeting?" (joint with Brian Curran and Bruce Meyer), we assess the desirability of geographic cost-of-living adjustments to poverty measures by examining how well they achieve a central objective of a poverty measure: identifying the least advantaged population. We compare an exhaustive list of material well-being indicators – drawn from survey and administrative data and including material hardships, appliances owned, home

quality issues, food security, public services, health, education, assets, permanent income, and mortality – of those classified as poor with and without a geographic adjustment. For nine of the ten domains of well-being indicators, we find that incorporating a geographic adjustment identifies a less deprived poor population. These results are broadly consistent across different poverty measures, various ways of implementing a geographic adjustment, and multiples of the poverty line.

In the third chapter "The Use and Misuse of Income Data and Extreme Poverty in the United States" (joint with Carla Medalia, Bruce Meyer, and Victoria Mooers), we re-examine the rate of extreme poverty – defined as living on less than $2/person/day – in the U.S. by linking survey data to administrative tax and program for 2011. Of the 3.6 million non-homeless households with survey-reported cash incomes below $2/person/day, we find that more than 90% are not in extreme poverty once we include in-kind transfers, replace survey reports of earnings and transfer receipt with administrative records, and account for the ownership of substantial assets. Of the households remaining in extreme poverty, 90% consist of a single individual. These results caution against taking survey incomes in the far left tail at face value.

CHAPTER 1: Certification and Recertification in Welfare Programs: What Happens When Automation Goes Wrong?

1.1 Introduction

Safety net programs play a key role in insuring individuals against economic risk, but enrollment often requires overcoming numerous administrative burdens (Herd & Moynihan 2018).[1] Crucially, many of these burdens appear not only at initial application but also at multiple points throughout one's receipt spell, as most programs in the United States require periodic recertification of eligibility. While researchers and policymakers have long been interested in the effects of administrative burdens on program take-up, prior studies typically examine changes only on a single margin (initial take-up or retention) and few explicitly distinguish between effects on both margins (see, e.g., Currie 2006, Bhargava & Manoli 2015, Barr & Turner 2018, Gray 2019).

At the same time, administrative burdens may serve a useful screening purpose by targeting benefits to the neediest individuals (Kleven & Kopczuk 2011). Neoclassical theory suggests that "ordeals" are likely to screen out less needy individuals with a higher opportunity cost of time (Nichols et al. 1971, Nichols & Zeckhauser 1982, Besley & Coate 1992). However, more recent studies conclude that enrollment barriers may screen out more needy individuals if application costs are negatively correlated with cognitive ability (Deshpande & Li 2019) or if needier

[1] For example, information about program eligibility may be difficult to acquire (Daponte et al. 1999, Manoli & Turner 2014, Armour 2018). Application forms may be onerous to complete and require the submission of numerous documents to prove eligibility (Currie & Grogger 2001, Bettinger et al. 2012), and applicants may also have to physically travel to a local welfare office to conduct an interview (Rossin-Slater 2013). Individuals may also face psychological costs such as stigma (Moffitt 1983).

individuals are prone to present bias (Bertrand et al. 2004, Mani et al. 2013, Mullainathan & Shafir 2013). Empirical studies of targeting efficiency also yield mixed results (see, e.g., Alatas et al. 2019, Deshpande & Li 2019, Finkelstein & Notowidigdo 2019, Homonoff & Somerville 2019, Shepard & Wagner 2021, Unrath 2021). Yet, these findings potentially mask important differences in targeting across initial enrollment and reenrollment, which may be further obscured by differences in their research settings.

This study investigates how barriers to enrollment affect the take-up and targeting of three important safety net programs – the Supplemental Nutrition Assistance Program (SNAP, formerly food stamps), Temporary Assistance for Needy Families (TANF), and Medicaid - in both the short and long run. Importantly, it decomposes and compares changes in overall take-up along the entry and exit margins and further disentangles the types of individuals screened out at each stage. Medicaid and SNAP are among the largest means-tested transfers in the U.S., reaching 23% and 10% of the nation's population at a collective federal cost of $465 billion in 2019.[2] TANF is a smaller cash welfare program aimed at families with children (reaching less than 1% of the population with federal expenditures of $14 billion in 2019).[3] This paper focuses on complexities in the enrollment and reenrollment process triggered by the automation of welfare services, which states and localities have increasingly adopted to allow individuals to apply and recertify virtually.[4] These changes are often thought to provide greater convenience for program applicants and lower administrative costs for program operators (Needels et al. 2000, Rowe et al. 2010). Yet, by

[2] See https://www.cms.gov/newsroom/fact-sheets/medicaid-facts-and-figures for the number of Medicaid recipients in October 2019, https://sgp.fas.org/crs/misc/R42640.pdf (p. 9) for federal Medicaid expenditures in 2019, and https://fns-prod.azureedge.net/sites/default/files/resource-files/SNAPsummary-9.pdf for SNAP recipients and benefit dollars in 2019.
[3] See Table A.1 in https://www.acf.hhs.gov/ofa/data/tanf-financial-data-fy-2019.
[4] By 2016, every state had either set up online applications, established call centers, or enabled telephone interviews for SNAP applicants - up from ten states in 2002.

processing and evaluating applications in a more mechanized fashion, automation may lead to greater inflexibility and an inability to tailor services to individual circumstances.

The research setting is the state of Indiana, which in 2007 automated the administration of SNAP, TANF, and Medicaid by outsourcing their management to the IBM Corporation. IBM used online and phone platforms to replace face-to-face interactions with local caseworkers, resulting in a lack of personalized caseworker assistance, a zero-tolerance policy for errors, and long wait times and unanswered calls at overwhelmed call centers. IBM's automated system rolled out to only two-thirds of Indiana's counties (covering 48% of the state's caseload) before unexpectedly halting in 2009 due to performance problems.[5] By assigning those counties receiving the rollout to treated areas and all other counties to untreated areas, this natural experiment facilitates a comparison of outcomes between these areas over time. While the rollout sought to initially reach less populous counties (believed to have an easier time transitioning to the new system), a myriad of validity checks establish that outcomes in treated and untreated counties evolved in parallel prior to treatment and would have continued to evolve in parallel in the absence of treatment.

To enable a novel analyses of changes in enrollment and targeting along the entry and exit margins, this paper's data sources draw from administrative longitudinal welfare records spanning nearly 3 million program recipients in Indiana between 2005 and 2014, which are linked to Internal Revenue Service (IRS) microdata and other administrative data sources. Leveraging a generalized difference-in-differences approach, the results show that overall SNAP, TANF, and Medicaid enrollments fall by 15%, 24%, and 4% (respectively) one year after the rollout of IBM's automated system. All three programs experience statistically significant declines in entry rates that are similar in magnitude, but differential increases in exit rates that are largest for TANF and smallest

[5] State of Indiana's Amended Complaint for Damages and Declaratory Relief at 67, Indiana v. International Business Machines Corp., No. 49D10-1005-PL-021451, 2010 WL 5677110 (Ind. Super. Nov. 4, 2010).

for Medicaid. For Medicaid, the overall enrollment reduction also exhibits considerable permanence – four years after the automated system was disbanded (and six years after the initial rollout), enrollment in treated counties remains 2% lower than in untreated counties.

This paper then considers who is screened out by IBM's automated system, in terms of both county-level characteristics and individual well-being. Enrollment reductions are most pronounced in earlier-treated and higher-poverty counties, as well as in lower-unemployment counties whose residents may have more earnings to verify. Turning to individual well-being – measured using more than a dozen proxies linked from administrative records or imputed from survey data – the results show that IBM automation slightly, but statistically significantly, improves overall targeting efficiency. Program recipients remaining in treated counties typically have lower pre-treatment incomes, less education, higher per-person benefits (based on a progressive benefit formula), and higher disability levels. These overall targeting effects, however, mask stark differences across application stages. IBM's automated system appears to be an efficient screen at initial application (screening out entrants who are less needy) but an inefficient screen at recertification (severing benefits to those who are more needy).

Finally, this paper tests the role of various mechanisms in explaining short- and long-run effects on enrollment. Recertification costs are an important contributor to heterogeneous effects across programs in the short run, as the initial reductions in enrollment continue to be larger for TANF and smaller for Medicaid after holding fixed recipient composition. In the longer run, this paper provides suggestive evidence on the role of various factors associated with permanent withdrawal from program rolls – and which may contribute to the persistence of the Medicaid enrollment reduction. Four years after IBM's automated system was disbanded, approximately 60% and 75% of the 2% long-run Medicaid enrollment decline is associated with an increased likelihood of dying and receiving Medicare (perhaps as an imperfect substitute for Medicaid), respectively, in the treated counties.

This paper makes several distinct contributions. First, it decomposes changes in overall take-up and targeting across entrants and exiters, in contrast to prior studies that tend to focus on changes via a single margin. In doing so, it distinguishes enrollment effects across initial application and recertification and assesses the sources of potential differences. This paper offers a new explanation for why existing empirical studies of targeting (particularly those examining SNAP) find seemingly inconsistent results. Finkelstein and Notowidigdo (2019), for example, find that SNAP application costs lead to improved targeting when studying initial applicants. Homonoff and Somerville (2019) and Gray (2019) find the opposite result when studying the effects of administrative burdens on SNAP recertifiers. By explicitly demonstrating that targeting efficiency may vary based on the stage at which individuals are applying for benefits, this study suggests a potential channel through which to reconcile the theoretical tension between neoclassical and behavioral models of targeting.[6]

This paper also analyzes a setting in which a common set of administrative burdens affects access to multiple programs, while prior studies typically evaluate effects for a single program. It is often the case that a single government agency administers multiple programs,[7] and administrative burdens may very well exist at the agency level. Whereas cross-program comparisons in the literature often embed large differences in research settings, this study compares effects across programs in a controlled setting. It focuses specifically on differences

[6] The differences in targeting efficiency across application stages could, in turn, reflect differences in the nature of the costs at each stage or differential selection of applicants into each stage.

[7] For example, a state welfare agency may administer programs ranging from SNAP, TANF, and Medicaid to the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) and the Low Income Home Energy Assistance Program (LIHEAP). The Social Security Administration also administers both Social Security and Supplemental Security Income.

across programs in transaction costs, which have been shown to be important determinants of take-up for well-established programs that require periodic recertification (Currie 2006).

Furthermore, this is one of the first papers to empirically examine the drawbacks to automating social services, in contrast to prior studies focusing on the benefits of automatic enrollment (Madrian & Shea 2001, Shepard & Wagner 2021) or technological changes to enrollment procedures (Kopczuk & Pop-Eleches 2007, Ebenstein & Stange 2010, Schwabish 2012). As these changes become widely adopted and imperfectly administered, they are likely to also induce complexities that are less well understood. This paper shows, for example, that even a short-lived automation spell – when poorly implemented – can lead to potentially long-term adverse effects. This relates more broadly to the intended and unintended consequences of governments engaging in policy experimentation (Callander & Harstad 2015). This paper also contributes to a growing literature analyzing the effects of privatizing services typically offered by the public sector, given that the automated system in Indiana was handled by a private organization with potentially differing goals. Prior studies have examined contexts ranging from health care (Duggan et al. 2017) to prisons (Mukherjee 2021).

The analyses in this paper rely on a combination of public- and restricted-use administrative records – the latter of which are part of the Comprehensive Income Dataset (CID) at the U.S. Census Bureau (Medalia et al. 2018). By using administrative data sources to measure welfare receipt, one can circumvent issues of measurement error pervading survey reports (Meyer & Mittag 2019). The longitudinal nature of the administrative panel data also enables the disentangling of enrollment and targeting effects along the entry and exit margins, which would not be possible using cross-sectional data. By linking to other microdata sources (including IRS tax records), this paper mobilizes a battery of well-being outcomes derived from survey and administrative data to rigorously measure program targeting.

While this paper focuses on a single state during the Great Recession, its policy significance has a far more extensive reach. Systems similar to IBM's automated system have been implemented in other states, and many states have moved towards having computers or call center workers make eligibility determinations – a trend that has accelerated and received renewed attention during the COVID-19 pandemic. On the flip side, caseworkers have seen their roles diminish over time in response to perceived inefficiencies and complaints about fraud. In addition, policies that seek to alter enrollment procedures are typically agnostic about the application stage to which they apply, but this paper shows that the welfare implications of these policies depend critically on the application stage. More generally, this paper speaks to the importance of reforming the administration of social service programs as a policy lever for expanding the safety net.

The rest of this paper is structured as follows. Section 1.2 provides background information about the welfare programs and IBM's automated system, drawing from sources that include court documents. Section 1.3 discusses data sources and compares the baseline characteristics of treated and untreated counties. Section 1.4 formalizes the empirical strategy. Sections 1.5 and 1.6 show regression results on enrollment and targeting, and Section 1.7 presents additional analyses and robustness checks. Section 1.8 concludes.

1.2 Background on Programs and Natural Experiment

1.2.1 Description of SNAP, TANF, and Medicaid Programs

This paper focuses on three important safety net programs for low-income individuals: the Supplemental Nutrition Assistance Program (SNAP), Temporary Assistance for Needy Families

(TANF), and Medicaid. A key feature of these federally-funded programs is that they are administered by the individual states. SNAP is the largest food assistance program, whose benefits are widely available to all low-income households and consist of in-kind vouchers – typically in the form of Electronic Benefit Transfer (EBT) cards – that can be used to purchase items from grocery stores and related outlets (Currie 2003). Households are usually eligible for SNAP if their gross and net monthly incomes fall below 130% and 100%, respectively, of the federal poverty line and their countable assets are below $2,250 ($3,500 if elderly or disabled).[8] In 2019, the maximum monthly benefit for a family of four was $642, amounting to 30% of the total income for a two-adult/two-child family at the poverty line.[9]

TANF, formerly known as Aid to Families with Dependent Children (AFDC), is a cash transfer targeted to low-income families with children. Families receiving TANF typically have either single parents or no parents at all.[10] After welfare reform in the late 1990s, a number of restrictions were placed on TANF recipients with the goal of helping them achieve economic self-sufficiency. Adults cannot receive TANF payments for more than sixty months over their lifetimes (although states can exempt this requirement for some families), and 50% of all TANF families must work for at least 30 hours per week (Ziliak 2016).[11] The current maximum monthly benefit for a family of four in Indiana is $346.[12]

---

[8] Under the Broad-Based Categorical Eligibility (BBCE) policy, many states have also increased the income limit to 200% of the poverty line and eliminated the asset test. Indiana is not one of these states.
[9] See https://www.fns.usda.gov/snap/allotment/COLA.
[10] Children who are part of TANF cases with no parents (i.e., child-only cases) generally live with another relative or guardian (such as a grandparent) or with a parent who does not qualify for TANF for certain non-financial reasons (Golden & Hawkins 2011). In October 2007, 46% of TANF cases nationwide had no parents while 25% of TANF cases in Indiana had no parents.
[11] In Indiana, adults cannot receive TANF for more than 24 months and children cannot receive TANF for more than 60 months.
[12] See https://www.in.gov/fssa/dfr/files/TANF-Brochure-English.pdf.

Table 1.1. Characteristics of SNAP, TANF, and Medicaid Recipients (2005-2007)

Characteristics	United States				Indiana			
	All (1)	SNAP (2)	TANF (3)	Med. (4)	All (5)	SNAP (6)	TANF (7)	Med. (8)
CPS (2005-2007)								
Has Children (%)	38.5	62.7	82.3	64.1	37.1	63.4	88.9	70.1
Single Parent (%)	3.8	20.7	32.2	11.6	4	25.9	44.7	16.9
Has Elderly (%)	22.1	15.3	7.8	19.3	19.1	9.9	2	10
Has Disabled (%)	15.9	41.1	37	34.4	13.9	45.2	38	35
Total Income ($)*	105,798	25,628	29,280	50,848	96,466	28,372	19,975	47,902
Total Earnings ($)*	85,096	14,779	17,089	36,534	78,736	17,304	10,436	36,021
Total Asset Inc. ($)*	5,285	147	533	1,496	3,757	82	69	546
Move in Last Year (%)	11.9	25.1	28.5	17.5	10.3	34.8	41.3	21.5
Homeowner (%)	72.4	30.2	25.8	48.9	78.4	23.2	17.4	46
Receive Hous. Sub. (%)	3.1	25.1	29	12.6	3.7	32.5	45.5	17.9
White (non-Hisp) (%)	72.3	48.4	43	52.8	88	69	56.2	73.2
Black (non-Hisp) (%)	10.4	26.5	29.8	17.5	7	23.7	30.2	15.6
Hispanic (%)	11.5	19.7	20.7	23.2	3.6	4.6	9.1	9.1
Has HS Diploma (%)	86.9	63.5	66.3	70.9	88.1	59.3	58.8	69.2
Citizen (%)	92.2	90	88.7	85.4	97.4	98.3	97.7	94.5
Unemployed (%)	2.6	8.4	11.7	5	3.2	13.5	18.5	7.9
Sample Size	361,692	21,249	5,039	60,359	5,511	400	99	786
SIPP (2005)								
Material Hardships	0.52	1.53	1.6	1.05	0.46	1.41	1.34	0.9
Unowned Appliances	1.35	2.67	2.79	2.17	1.31	2.31	2.63	1.86
Home Problems	0.24	0.53	0.69	0.42	0.16	0.37	0.46	0.3
Food Problems	0.37	1.34	1.46	0.91	0.32	1.2	1.29	0.81
Any Health Prob. (%)	22.1	48.1	45	39.8	20.6	38.2	33	33
Sample Size	37,368	3,315	537	7,869	1,332	116	21	240
NLSY97 (2005-2007)								
AFQT Score (Pctile)	51.1	32.9	25.6	31				
AFQT < Army Min (%)	30	55.4	67.8	59.4				
Sample Size	5,990	1,058	256	835				

* Equivalized for 2-adult, 2-child household

Data Sources (public-use): Current Population Survey Annual Social and Economic Supplement (2006-2008), Survey of Income and Program Participation (2004 Panel, Wave 5), National Longitudinal Survey of Youth 1997

Notes: This table shows average characteristics of the entire population and SNAP, TANF, and Medicaid recipients in the United States and in Indiana, using data for reference years 2005-2007. Columns 1, 2, 3, and 4 show estimates for the entire population and for SNAP, TANF, and Medicaid recipients (respectively) in the entire nation, and Columns 4, 5, 6, and 7 show estimates for the entire population and for SNAP, TANF, and Medicaid recipients (respectively) in Indiana. CPS characteristics are calculated for households and/or household heads pooled over reference years 2005-2007 (weighted using survey household weights), SIPP characteristics are calculated for households in reference year 2005 (weighted using survey household weights), and NLSY characteristics are calculated for individuals in reference years 2005-2007 (weighted using survey individual weights for 2007).

Medicaid is the single largest means-tested transfer in the U.S. and pays for medical costs accrued by low-income individuals. Prior to its expansion under the Affordable Care Act (ACA) in 2014 (which extended eligibility to single adults without dependents), eligible groups mainly included low-income families with children, elderly individuals, and disabled individuals. Indiana extended eligibility to single adults with low incomes starting in 2008, being one of a few states to do so prior to the ACA expansion.[13]

Drawing from various survey data sources, Table 1.1 shows characteristics of SNAP, TANF, and Medicaid recipients in both the U.S. and Indiana measured between 2005-2007. In general, households receiving TANF are more likely to have children, Medicaid recipients are more likely to be elderly, and SNAP recipients are more likely to have a work-limiting disability. Furthermore, Medicaid recipients appear to be more well-off than SNAP or TANF recipients across a range of characteristics. They have higher levels of income, education, and homeownership, lower levels of unemployment and disability, and fewer material hardships. In contrast, households receiving TANF appear to be less well-off than SNAP or Medicaid recipients (particularly in Indiana) on a variety of dimensions – including cognitive ability based on the Armed Forces Qualification Test (AFQT). Compared to those in the rest of the country, program recipients in Indiana are more likely to be single parents, disabled, and unemployed and are less likely to be elderly, homeowners, and high-school educated.[14]

[13] Most beneficiaries in Indiana also received Medicaid benefits on a fee-for-service basis prior to 2005, before Indiana shifted to offering the majority of Medicaid benefits through managed care plans.

[14] The comparisons of Indiana recipients to those in the broader U.S. follow a similar pattern when using the SNAP Quality Control administrative microdata.

1.2.2 Application and Recertification Processes

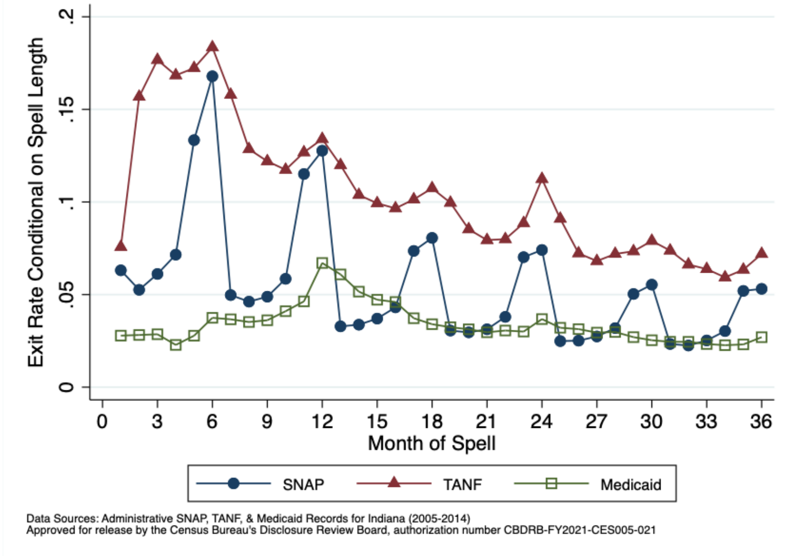
The enrollment processes for SNAP, TANF, and Medicaid each involve undergoing an initial application process and recertifying eligibility at periodic intervals, although the exact processes differ across states. The initial application process requires applicants to provide demographic information on every individual in the assistance unit and an exhaustive list of income sources and expenses. In many states, applicants are also required to submit numerous supporting documents, ranging from proofs of identity to proofs of all income sources and deductions. In many states (including Indiana), individuals can apply for SNAP and TANF using a single application, while one almost always fills out a separate application to obtain Medicaid.[15]

After submitting their applications, applicants are usually contacted by their local welfare office to conduct an interview to verify eligibility. Successful applicants typically start receiving benefits within 60 days of first submitting their applications. Each program requires that its recipients periodically recertify their eligibility, which involves completing a renewal form listing changes in one's income and household structure, providing supporting documentation, and conducting another interview if needed. Medicaid recipients typically recertify every 12 months, SNAP recipients recertify their eligibility every 6 to 12 months,[16] and TANF recipients have the shortest recertification intervals (6 months). TANF recipients often interact with their welfare office even more frequently, as they usually have to report all income changes (which occur regularly given work and job search requirements).

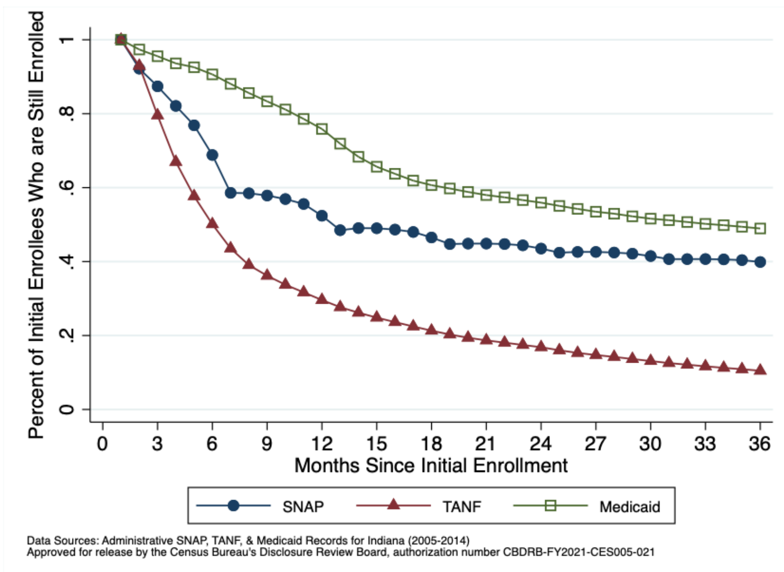
[15] For Indiana, the Medicaid application is similar to the SNAP/TANF application in many ways, although it asks for some additional information (e.g., health coverage from current employer if applicable).

[16] Indiana requires that its SNAP recipients provide updates on unit structure and incomes/expenses every 6 months, while renewal interviews may happen every 6 or 12 months. Prior to 2003, Indiana required that SNAP recipients report any income changes immediately to their welfare office so that benefit amounts could be adjusted accordingly. After 2003, however, Indiana enacted "simplified reporting" rules requiring SNAP recipients only to report income changes outside of the six-month recertification interval when their new incomes exceed the gross income limit and therefore terminate their eligibility (USDA 2019).

Figure 1.1. Baseline Hazard Rates and Survival Probabilities by Spell Length



(a) Hazard Rates



(b) Survival Probabilities

Data Sources: Administrative SNAP, TANF, and Medicaid records for Indiana (2005-2014)
Notes: These figures show baseline hazard rates and survival probabilities calculated over all individual recipients of SNAP, TANF, or Medicaid in Indiana. Panel A shows hazard rates (average exit rates conditional on a given length of one's receipt spell) by program. Panel B provides a transformation of Panel A by showing the share of recipients (by program) that are enrolled at each month relative to their entry month. The samples consist of all recipients who initially enroll between February 2005 and October 2011. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization number: CBDRB-FY2021-CES005-021.

Figure 1.1a shows the rate of exiting SNAP, TANF, or Medicaid for each month of a program spell (i.e., hazard rate), calculated over all program recipients in Indiana initially enrolling between February 2005 and October 2011. Hazard rates are consistently high for TANF – particularly during the first six months of the program spell – but peak slightly every 6 months (corresponding to TANF recertification intervals). In contrast, exits for SNAP are dramatically concentrated around recertification periods (which also occur every 6 months) but are otherwise relatively infrequent. Hazard rates for Medicaid tend to be lower than those for SNAP or TANF, peaking somewhat every 12 months (corresponding to Medicaid recertification) but remaining relatively uniform across spell length. As a result of these exit patterns, 76% of Medicaid recipients are still enrolled 12 months after initial enrollment, compared to only 52% of SNAP recipients and 30% of TANF recipients (Figure 1.1b).[17] Recent studies have indicating that many who exit SNAP do so because of missed deadlines rather than ineligibility (Mills et al. 2014, Gray 2019), suggesting that administrative burdens in the re-enrollment process may partly explain the magnitudes of the exit rates in Figure 1.1.

1.2.3 Automating the Administration of Welfare Services in Indiana

In 2006, Indiana Governor Mitch Daniels announced an effort to modernize the state's welfare system, which he described as “plagued by high error rates, fraud, wasted dollars, poor conditions for its employees, and very poor service to its clients.”[18] This initiative followed prior endeavors by Governor Daniels to privatize several of the state's public services, including toll

[17] Recent studies have also found that approximately half of all SNAP recipients leave the program within their first year of receipt (Gray 2019, Unrath 2021).
[18] Indiana v. Int'l Bus. Machs. Corp., 4 N.E.3d 696, 703 (Ind. Ct. App. 2014).

roads and meal services in prisons.[19] State officials believed that transitioning from a face-to-face casework system to a virtual platform would streamline the processing of applications and recertifications. There were also concerns that the existing system, which relied on caseworkers developing personal relationships with recipients, invited fraud.[20]

In December 2006, after seeking various vendors for the project, Indiana awarded a 10-year, $1.3 billion contract to the IBM Corporation.[21] The primary goals of the new system were to combat fraud, lower administrative costs, improve the welfare-to-work pipeline, and improve access to services through a more flexible platform.[22] While the state retained operational and policymaking control over SNAP, TANF, and Medicaid (including making final eligibility determinations), IBM would assist in processing initial applications and recertifications.[23] Modeled after similar systems implemented in Florida and Texas, the new system in Indiana would allow citizens to apply for benefits online or through a call center, without the need for a face-to-face meeting with a local caseworker (Aman 2013). Approximately 1,500 of the 2,200 employees of Indiana's Family and Social Services Administration (FSSA) were transferred to IBM's private call centers, and eligibility determinations would be made on a centralized, statewide basis rather than in each county's welfare office.[24]

---

[19] See https://www.latimes.com/archives/la-xpm-2011-jun-24-la-na-indiana-privatize-20110624-story.html and https://finance.yahoo.com/news/ind-court-sets-hearing-ibm-172824630.html.

[20] State officials dealt with several high-profile fraud cases in the mid-2000s, including one in which a caseworker colluded with members of an Indianapolis church congregation to defraud the state out of $62,000 in food stamps and cash assistance (see https://www.wthr.com/article/news/church-leaders-charged-with-food-stamp-fraud/531-4351d507-4cf0-42d0-91ca-32ea302abb17). Yet, Indiana's SNAP error rates only slightly exceeded national averages in both under- and over-payment during the early-to-mid 2000s. Even during fiscal year 2007, Indiana's above-average SNAP error rates were still below those of neighboring states Ohio and Michigan.

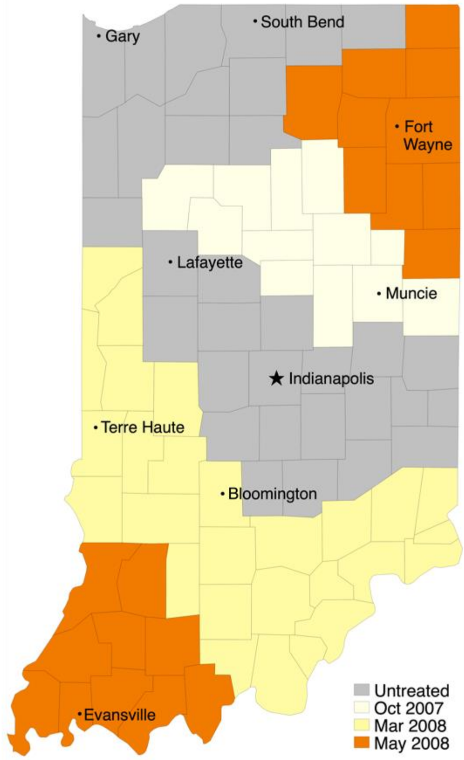
[21] Intl' Bus. Machs. Corp. v. Indiana, 112 N.E.3d 1088, 1093 (Ind. Ct. App. 2018).

[22] Indiana v. Int'l Bus. Machs. Corp., 51 N.E.3d 150, 154 (Ind. 2016).

[23] Id. at 154.

[24] See https://www.wthr.com/article/news/daniels-signs-1-billion-welfare-outsourcing-deal/531-e74baf2c-5662-40bd-8c89-1c5215beb77b.

Figure 1.2. Rollout of IBM's Automated System Across Indiana Counties



Notes: This map illustrates the rollout of IBM's automated system across Indiana's counties. Counties in light yellow received the first wave of automation treatment (in October 2007), counties in dark yellow received the second wave of automation treatment (in March 2008), and counties in dark orange received the third wave of automation treatment (in May 2008). Counties in gray never received the automation treatment.

The rollout of IBM's automated system proceeded in several stages. In March 2007, Indiana informed program recipients about the upcoming changes and began transitioning state employees to IBM.[25] Starting in October 2007, Indiana rolled out IBM's automated system to counties in several waves (Figure 1.2). The rollout sought to initially avoid the most populous counties, as they were likely to have a harder time transitioning to the new system.[26] In the first wave (October 2007), IBM was rolled out to 12 relatively rural counties in north-central Indiana.[27] The second rollout wave (March 2008) reached 27 more counties in southern and central Indiana, including

[25] Am. Compl. at 54.
[26] Id. at 67.
[27] Id. at 58.

the cities of Bloomington and Terre Haute.[28] The third rollout wave (May 2008) reached 20 more counties in southwest and northeast Indiana, covering cities such as Fort Wayne and Evansville.[29] IBM's system was originally intended to roll out to all 92 counties, but the expansion was halted due to performance issues.[30]

Call centers were quickly overwhelmed after the initial rollout, leading to tens of thousands of unanswered calls.[31] Moreover, call center operators – many of whom were not adequately trained on how to handle the new system – often served as the sole points of contact for clients in the automated counties (Eubanks 2018, p. 50). Under the previous system, local caseworkers were assigned to a docket of families with whom they interacted face-to-face and guided through the full application process. Under the new automated system, caseworkers were assigned to tasks dropped into a computerized queue, and applicants spoke with a different operator every time they called.[32] The lack of consistent guidance from a dedicated caseworker led to more mistakes on application and recertification forms.

Exacerbating these issues was a rigid zero-tolerance policy on application errors, which arose partly out of the mechanized processing of applications. Any mistake was assumed to be the fault of the client and interpreted as a refusal to cooperate in establishing eligibility (Eubanks 2018, p. 42-43). This often led to a notification of a denied application or the imminent expiration of benefits. As these notifications never specified the nature of the underlying mistakes, clients were often unaware of the reasons for denial. IBM also required that all records be electronically scanned as part of the move to a "paperless" system, placing the burden on clients – regardless of

[28] Id. at 62.
[29] Id. at 63.
[30] Id. at 68.
[31] Indiana, 4 N.E.3d at 708.
[32] See https://www.thenation.com/article/archive/want-cut-welfare-theres-app/.

how long they had been receiving benefits – to resubmit all documents that were previously available only as hard copies.[33] This led to a number of incomplete recertifications for recipients who could not locate documents in a timely manner.

Technical glitches on the IBM platform and a backlog of unanswered phone calls further led to numerous individuals being improperly denied benefits.[34] Moreover, the verification documents that IBM required clients to submit often went unprocessed, with a single missing document being enough to warrant the denial of an application. Approximately 11,000 documents remained unprocessed in December 2007, with this number surging to nearly 283,000 lost documents by early 2009 (Eubanks 2018, p. 50). Finally, call center workers would be so far behind in processing applications that they sometimes recommended premature denials simply to make their performance appear more timely (Eubanks 2018, p. 51).[35]

As a result of these performance problems, the rollout of IBM's automated system was halted in January 2009.[36] Ultimately, 59 out of Indiana's 92 counties (covering 48% of the state's caseload) were transitioned to the automated system, although some of the most populous areas in Indiana – including Indianapolis, northwest Indiana, and South Bend - were never automated.[37] This created a natural experiment, where counties receiving the rollout were assigned to treated areas and all other counties were assigned to untreated areas whose outcomes can be compared

[33] Indiana, 51 N.E.3d at 153. See also p. 50 of Eubanks (2018).
[34] Indiana, 51 N.E.3d at 156. In a subsequent court case in 2010, a number of individuals provided anecdotes about how they were adversely impacted by the automated system. One individual was denied Medicaid benefits - and subsequently SNAP benefits - because she missed her scheduled recertification interview while in the hospital suffering from terminal cancer. Even though she called IBM's call center to inform them of the circumstances, her benefits were cut off anyway (Am. Compl. at 114). In another case, a deaf individual was unable to comply with IBM's requirement that she have a telephone interview, causing considerable delay in granting her benefits (Am. Compl. at 115).
[35] Applicants were then recommended to either reapply or appeal the decision, with either avenue potentially taking months if pursued.
[36] Am. Compl. at 71.
[37] Id. at 67.

over time. After the rollout was halted, Indiana asked IBM in March 2009 to formulate a "Corrective Action Plan" to show consistent and needed improvement in the 59 automated counties.[38] In July 2009, Indiana and IBM agreed on the terms of a Corrective Action Plan.[39] Indiana eventually decided to terminate its contract with IBM in December 2009 in favor of a hybrid modernization approach, which moved eligibility determinations back to local welfare offices and gave individuals the option to have either face-to-face or phone interactions with caseworkers.[40]

1.3 Data Sources and Summary Statistics

1.3.1 Data Sources

This paper relies on two sets of data sources: 1) public-use data sources on county-level program enrollment and various county-level characteristics, and 2) restricted-use administrative microdata on the universe of program participants linked to microdata on various outcomes available at the U.S. Census Bureau.

Public-Use Data (County-Level)

Monthly county-level enrollment data for SNAP, TANF, and Medicaid (starting in October 2002) are obtained from Indiana's Family and Social Services Administration (FSSA). The SNAP and TANF data cover the number of enrolled cases (assistance units), individuals, and benefit dollars, while the Medicaid data cover the number of enrolled individuals.[41] Also available are

[38] Indiana, 4 N.E.3d at 709.
[39] Indiana, 4 N.E.3d at 710.
[40] Int'l Bus. Machs. Corp., 124 N.E.3d at 1094.
[41] Monthly county-level Medicaid enrollment records are available online starting in September 2013, and earlier records were obtained using a Freedom of Information Act request.

county-level data sources covering other demographic and economic characteristics, including monthly unemployment rates from the Bureau of Labor Statistics (BLS) and quarterly data on jobs, establishments, and wages by industry from the BLS Quarterly Census of Employment and Wages.[42] Annual data on county-level population counts (by age, race/ethnicity, and gender categories), poverty rates, and median incomes are available from the U.S. Census Bureau. Dollar amounts are adjusted for inflation using the Personal Consumption Expenditure (PCE) price index.[43]

This paper also brings in annual county-level enrollment data for a number of other programs not administered by IBM's automated system (for placebo tests). These include the numbers of individuals enrolled in both Social Security and Supplemental Security Income (SSI), measured by the Social Security Administration in December of every year. County-level enrollment counts for Medicare come from the Centers for Medicare and Medicaid Services (for 2007 onward). In addition, the Department of Education's Common Core of Data contains annual school-level counts of students receiving free and reduced meals, which can be aggregated to the county level.

Restricted-Use Data (Individual-Level)

Accompanying the public-use data sources are a number of restricted-use administrative data sources linked together as part of the Comprehensive Income Dataset (CID). Administrative SNAP, TANF, and Medicaid records are available from Indiana and contain information on

[42] Note that county-level unemployment rates are not directly measured by the BLS but are instead imputed based on state-level employment rates and county-level Unemployment Insurance records. Consequently, these county-level rates are likely subject to some measurement error (Ganong & Liebman 2018).
[43] The PCE deflator is better than the Consumer Price Index (CPI) in accounting for the ability of consumers to substitute between broad product categories when prices change (Bullard 2013, Winship 2016).

monthly benefits paid out to the universe of SNAP and TANF recipients (starting in July 2004) and monthly enrollment for the universe of Medicaid recipients (starting in January 2005). These program data also contain information on incomes (as observed by Indiana's FSSA), county of residence, eligibility characteristics, and some limited demographic information (e.g., gender, race, education, etc.). A key advantage of these longitudinal welfare records is that they measure a case's receipt history over time and thereby enable analyses along the entry and exit margins.

Administrative microdata are also available for various outcomes that can be used to examine either targeting and/or the effects of program receipt. Tax records from the Internal Revenue Service (IRS) provide information on various income measures, including formal sector wages from IRS Form W-2s, retirement income from IRS Form 1099-Rs, and adjusted gross income from IRS Form 1040s. The 2010 Decennial Census and the 1099 Information Return Master File (covering Forms W-2, 1098, 1099-DIV, 1099-G, 1099-INT, 1099-MISC, 1099-R, 1099-S, or SSA-1099) contain information on addresses over time, allowing one to construct indicators of migration across counties or states. Medicare enrollment records from the Centers for Medicare and Medicaid Services (CMS) provide annual enrollment indicators for the universe of Medicare recipients. The Numident file from the Social Security Administration (SSA) can also be used to measure migration since birth and mortality. These data are linked using individual identifiers called Protected Identification Keys (PIKs).[44]

[44] PIKs can be approximately thought of as anonymized Social Security Numbers. Nearly 100% of individuals in all administrative data sources link to a PIK.

1.3.2 Characteristics of Treated and Untreated Counties in Indiana

Table 1.2 compares untreated and treated counties across a range of characteristics measured in September 2007 (the month prior to treatment).[45] The summary statistics for the 33 untreated counties are calculated for all such counties (Column 1) and also for a "less populous" subset that omits Marion and Lake counties (Column 2), the two most populous counties in Indiana. The summary statistics for the 59 treated counties are also calculated across all such counties (Column 3) and subdividing by treatment wave (Columns 4-6). Compared to the "less populous" set of untreated counties, treated counties have higher rates of SNAP and Medicaid receipt and similar rates of TANF receipt. However, the counties treated in Wave 1 tend to have the highest rates of welfare receipt for each program, while the counties treated in Wave 3 tend to have the lowest rates of welfare receipt for each program (except for TANF). Treated counties are also more rural and more white than their untreated counterparts (even after omitting Marion and Lake counties from the untreated group), in addition to having fewer children and more elderly adults.

Turning next to economic characteristics, treated counties have a higher share of individuals employed in the manufacturing sector and fewer shares of workers in the construction, wholesale/retail trade, and professional/administrative services sectors. Additionally, treated counties have slightly higher unemployment rates than untreated counties. Finally, treated counties have median household incomes below those of untreated counties. Poverty rates are therefore considerably higher in treated counties, which is consistent with these counties having higher rates of program receipt. In summary, the counties exposed to IBM's automated system are on average

[45] Annual numbers for population counts and poverty/income measures correspond to calendar year 2007, quarterly numbers for employment characteristics correspond to the third quarter of 2007, and monthly numbers for welfare receipt and unemployment rates correspond to September 2007.

Table 1.2. Characteristics of Untreated and Treated Counties (2007)

Characteristics	Untreated All (1)	Untreated Less Pop. (2)	Treated All (3)	Treated Wave 1 (4)	Treated Wave 2 (5)	Treated Wave 3 (6)
Welfare Receipt						
SNAP (%)	9.6	7	9	11.2	8.7	8
TANF (%)	2	1.1	1.3	1.7	1.1	1.2
Medicaid (%)	13.8	11.1	13.5	15.2	13.4	12.6
Population						
Rural (%)	20.2	31.2	40.4	35.1	49.3	35.7
White (%)	83.6	92.2	92.8	92.2	94.5	91.7
Black (%)	12.7	4.2	4.4	5.3	2.8	5.4
Hispanic (%)	7.2	5.1	3	3.1	2	3.9
Ages 0-17 (%)	26	26.2	24.2	23.2	23.1	25.7
Ages 18-64 (%)	62.2	61.8	62.3	61.8	63.8	61.2
Ages 65+ (%)	11.8	12	13.5	15	13.1	13.1
Employment & Industry						
Manufacturing (%)	17.4	20.5	21.9	21	19.1	24.8
Construction (%)	5.9	5.7	4.8	4.1	4.7	5.3
Wholesale & Retail Trade (%)	15.8	16.4	14.7	15.2	14	15.1
Prof. & Adm. Services (%)	15.9	13.9	11.2	10.8	9.9	12.6
Education & Health (%)	19.9	19.6	19.3	24	18	18
Entertain. & Hospitality (%)	10.4	10.2	9.4	10.3	9.1	9.1
Public Administration (%)	4.7	4.4	5.3	6.4	6.4	3.8
Unemployment (%)	4	3.8	4.4	5.3	4.2	4.1
Income & Poverty						
Median Household Income ($)	51,403	54,624	45,186	42,507	43,851	47,739
Poverty Rate (%)	12.1	10.1	12.7	13.8	14.2	10.9
Average County Population	109,413	72,053	46,932	51,050	36,806	58,131
Number of Counties	33	31	59	12	27	20

Data Sources (public-use): County-level program records (2007), Census population and SAIPE estimates (2007), BLS Local Area Unemployment Statistics (2007), BLS Quarterly Census of Employment and Wages (2007)

Notes: This table shows average pre-treatment demographic and economic characteristics of counties in Indiana classified based on whether they are treated (i.e., received IBM's automated system) or untreated. Column 1 shows averages for all untreated counties, and Column 2 shows averages for a subset of the untreated counties that excludes the two largest counties (Marion and Lake counties). Column 3 shows averages for all treated counties, and Columns 4-6 divide treated counties based on the wave during which they receive treatment. Characteristics measured at the monthly level (i.e., SNAP, TANF, Medicaid, and unemployment rates) correspond to September 2007 values, while characteristics measured at the annual level (all other variables) correspond to 2007. All county-level averages (with the exception of total population) are weighted by county population in 2007.

smaller, poorer, more rural, less racially and ethnically diverse, and more welfare-reliant than the counties that were never automated. For the most part, these differences are larger for counties treated in earlier waves.

1.4 Empirical Strategy

To assess the causal effects of IBM's automated system on SNAP, TANF, and Medicaid enrollment, this paper relies on a generalized difference-in-differences design. This empirical strategy exploits the fact that counties were naturally assigned to treated and untreated groups whose outcomes can be tracked over time. The identifying assumption is that treated and untreated counties, despite differing on baseline characteristics, would have had similar trends in outcomes in the absence of treatment.

1.4.1 Regression Specification

The following dynamic specification compares how enrollment in treated counties (i.e., those automated by IBM) evolves relative to enrollment in untreated counties:

$$y_{ct} = \mu_c + \lambda_t + \Sigma_k \, \gamma_k \, D^k_{ct} + \beta X_{ct} + \varepsilon_{ct}, \tag{1.1}$$

where y_{ct} is an aggregate outcome (e.g., log number of SNAP cases) for county c and year-month t. D^k_{ct} is a dummy variable equaling one if county c receives treatment and month t is k quarters after (or before, if k is negative) IBM's automated system is implemented. The coefficients of interest are the γ_k's, which measure the difference in outcomes between treated and untreated counties k quarters after automation.[46] The main estimates rely on a window of 12 quarters (3 years) before treatment and 24 quarters (6 years) after treatment, although estimates are also shown for an extended pre-treatment window of 20 quarters (5 years). County- and month-fixed effects are denoted respectively by μ_c and λ_t, and X_{ct} is a vector of county- and time-varying covariates.[47]

By measuring effects on county-level enrollment (aggregated from individual-level enrollment data), one can capture changes at both the entry and exit margins. County-month observations are weighted by enrollment volume in September 2007 (the month prior to IBM's initial rollout). Marion and Lake counties (which are part of the metropolitan areas of Indianapolis and Chicago, respectively) are excluded from the untreated group, as they are by far the two most populous counties in Indiana and unlikely to be a good counterfactual for any of the treated counties. Marion and Lake counties are outliers on a number of dimensions relevant to program enrollment, including urbanicity and the population shares of black and Hispanic individuals.[48]

While equation (1.1) allows for the decomposition of treatment effects across event-time, the following static specification can be used to summarize the dynamic estimates:

$$y_{ct} = \mu_c + \lambda_t + \gamma D_{ct} + \beta X_{ct} + \varepsilon_{ct}, \qquad (1.2)$$

where D_{ct} is a dummy variable equaling one if county c receives treatment and month t is after the rollout of the automated system. Estimates of γ use observations within a window of 12 quarters

---

[46] The dummy variable corresponding to k = -1 is omitted, meaning the estimates of γ_k should be interpreted as being relative to event-time k = -1 (the quarter immediately prior to treatment).

[47] These covariates include total population, the numbers of white, black, and Hispanic individuals, the number of females, the number of children, and the number of non-elderly adults (with ages between 18 and 64). All of these variables are logged when the outcome is also logged, and they vary by county and year (rather than county and month).

[48] Despite representing only 2% of all counties in Indiana, Marion and Lake counties constitute 22% of the overall state population and 32% of the state SNAP caseload. They are also the two most urban counties and have the highest population share of black individuals in Indiana.

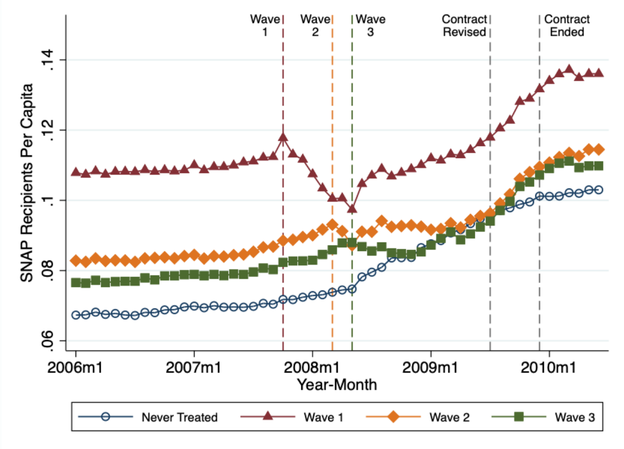
before treatment and 12 or 24 quarters after treatment (to measure both the short- and long-term effects of the automated system). As a result, the estimated γ in equation (1.2) can be roughly interpreted as a weighted average of the estimated γ_k's in equation (1.1). While a variety of recent studies have identified biases associated with traditional two-way fixed effects or event study specifications with variation in treatment timing (see, e.g., Callaway and Sant'Anna 2020, Goodman-Bacon 2021, Sun and Abraham 2020), this paper later provides evidence that these biases are small in the current context. This is because the staggered IBM rollout occurs over a short time frame (7 months) relative to the overall time frame for the analyses (72 months or 108 months).

Finally, the main estimates of the standard errors are clustered at the county level, as doing so accounts for serial correlation within counties over time. However, there may also exist spatial correlation across counties within a treatment region, since IBM's automated system was rolled out to a group of counties at a time. As a robustness check, standard errors are also clustered at the region level, where regions are defined by the presence and/or timing of treatment. A complication under this approach is that the number of clusters (i.e., regions) is only six, while accurate estimates of clustered standard errors rely on the number of clusters being large. The robustness check therefore uses a wild cluster bootstrap procedure to account for the small number of clusters (Cameron et al. 2008).

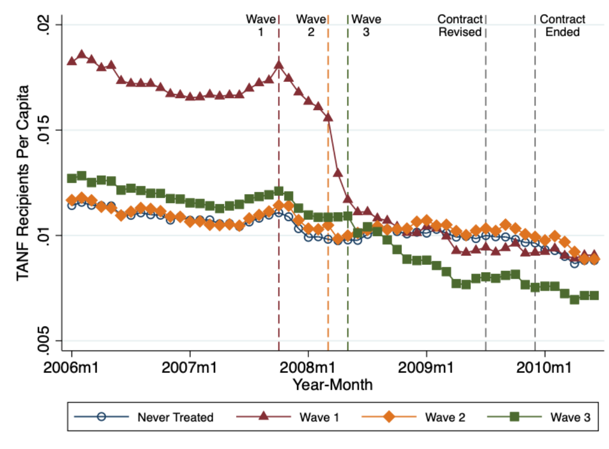
1.4.2 Raw Trends in Welfare Receipt

To provide visual evidence for the difference-in-differences strategy, Figure 1.3 shows monthly trends in SNAP, TANF, and Medicaid receipt rates for the counties automated by IBM

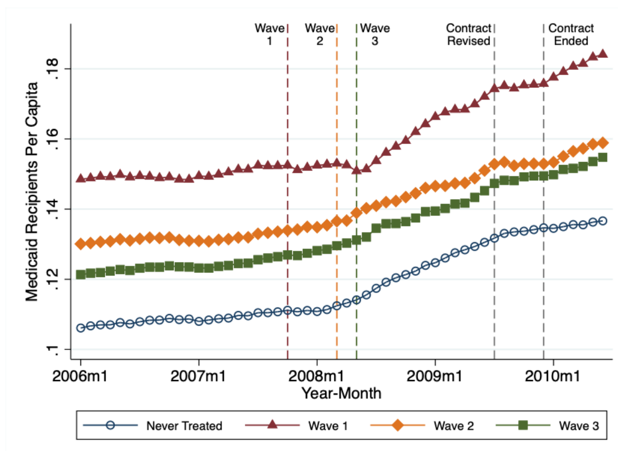
Figure 1.3. Raw Trends in Welfare Receipt Rates



a) SNAP



b) TANF



c) Medicaid

Data Sources (public-use): County-level SNAP, TANF, and Medicaid program records (2006-2010), Census population estimates (2006-2010)
Notes: These figures show raw trends in monthly receipt rates for SNAP, TANF, and Medicaid - calculated as the share of the population receiving each program - for counties exposed to the automated system (categorized by rollout wave) and counties never automated by IBM. The receipt rate for each group is calculated as the average of receipt rates (weighted by county population in September 2007) across each of the group's counties. In each of the panels, the dashed vertical lines correspond to the timing of major events, including the rollout of the automated system in Waves 1, 2, and 3, the undertaking by IBM of the "Corrective Action Plan" (in July 2009), and the termination of the IBM contract (in December 2009).

(categorized by rollout wave) and the never-automated counties.[49] Focusing first on SNAP receipt, Figure 1.3a shows that the enrollment patterns for treated counties track the patterns for untreated counties until precisely each set of treated counties is automated. Six months after each set of

[49] Average county-level receipt rates are weighted by total county population measured in September 2007.

counties is automated, the rate of SNAP receipt falls by 10% for the counties treated in Wave 1 and by 3% or less for the counties treated in Waves 2 and 3. A major reason for these raw differences in receipt rates is that treatment in Waves 2 and 3 tends to coincide with increases in SNAP applications resulting from the Great Recession, as indicated by the steadily increasing enrollment rates in the untreated counties.

Similar to the patterns for SNAP, TANF enrollment following treatment falls dramatically in the counties treated in Wave 1 while experiencing a more muted decrease in the counties treated in Waves 2 and 3 (Figure 1.3b). Unlike the patterns for SNAP, the trend in TANF enrollment remains relatively flat for the untreated counties. For Medicaid, there are no sharp drops in enrollment for any of the treated sets of counties, although they experience flatter growth rates than the untreated counties (Figure 1.3c).

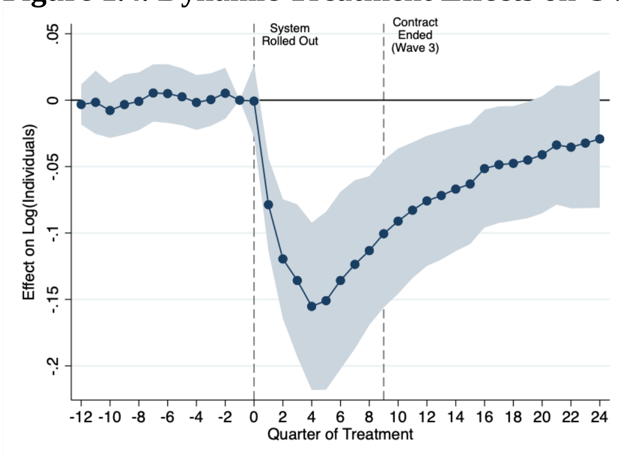
1.5 Effects of IBM Automation on Enrollment

This section describes the overall effects of IBM's automated system on SNAP, TANF, and Medicaid enrollment. The first subsection discusses dynamic effects on total enrollment – following their evolution for six years after treatment – and static treatment effects. The next subsection disaggregates changes in total enrollment along the entry and exit margins.

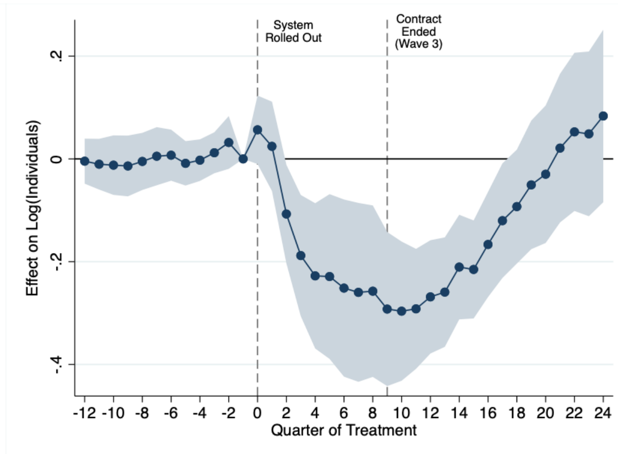
1.5.1 Overall Effects on SNAP, TANF, and Medicaid Enrollment

Focusing first on SNAP enrollment, Figure 1.4a shows no pre-trends in the twelve quarters prior to treatment before a sharp drop in SNAP enrollment immediately after the automated system is rolled out. Four quarters after the implementation of IBM's automated system, treated counties

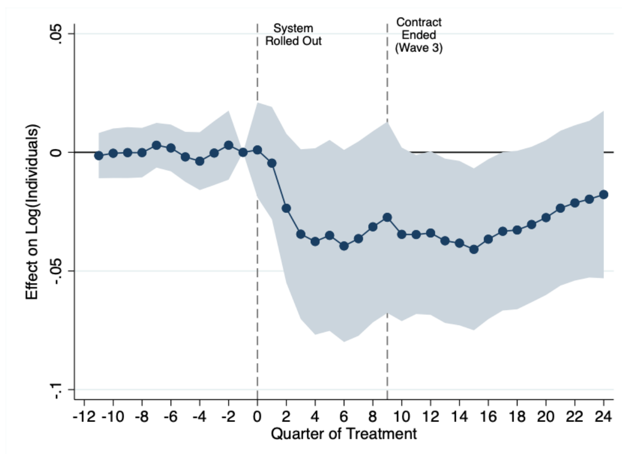
Figure 1.4. Dynamic Treatment Effects on Overall Enrollment



a) SNAP



b) TANF



c) Medicaid

Data Sources: Administrative SNAP, TANF, and Medicaid records for Indiana (2004-2014), Census pop. estimates (2004-2014)
Notes: These figures show regression estimates of log total individuals on binary indicators corresponding to the event-quarter relative to receiving IBM automation, controlling for county- and month-fixed effects as well as county- and year-varying covariates that include log total population and population subgroups by race, gender, and age. Observations are at the county-month level and are weighted by the county-specific enrollment volume in September 2007 (month prior to initial IBM rollout). Marion and Lake counties (which are untreated) are excluded from the sample given that they are outliers on a number of dimensions. Standard errors are clustered at the county level. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization number: CBDRB-FY2021-CES005-021.

have 15% fewer SNAP individuals than untreated counties. The decline in SNAP enrollment reverses in the ensuing quarters, with the turnaround corresponding to the institution of IBM's contract revision. The rebound in SNAP enrollment among the treated counties slows down after the termination of IBM's contract. Four years after IBM's automated system was disbanded (and six years after it was first implemented), SNAP enrollment in treated counties remains 3% lower than in untreated counties. These longer-run estimates, although statistically insignificant, suggest some level of permanence in the reduction of SNAP receipt in the treated counties.

Figure 1.4b shows an ever steeper initial drop in TANF receipt as a result of the automated system. Treated counties have 24% fewer TANF individuals than untreated counties after the first four quarters of automation, with this difference expanding to 27% after four more quarters. This decline is more pronounced among recipients in cases with an adult and less substantial among recipients in child-only cases – likely a result of adults facing work requirements and thus having more incomes to verify. The reduction in TANF enrollment in the treated counties hits a trough ten quarters after initial treatment, but – unlike the effects for SNAP – the differences in TANF enrollment between the treated and untreated counties disappear six years after the automated system was first rolled out. The long-term rebound in TANF enrollment can likely be explained by the continuous turnover among TANF recipients, as lifetime enrollment caps imply that most TANF recipients at the time of IBM's rollout would no longer be recipients several years later.

Medicaid receipt also noticeably drops after treatment (Figure 1.4c), although the initial reduction for Medicaid is smaller than that for SNAP or TANF. Four quarters after automation, treated counties have 4% fewer Medicaid individuals than untreated counties. Although there is a modest reversal in the decline two years after the automated system was disbanded, the estimates show a level of permanence in the reduction of Medicaid receipt. Four years after the automated system was terminated, Medicaid enrollment in the treated counties remains 2% lower than in the untreated counties.

Table 1.3. Treatment Effects on Overall Program Enrollment

Outcomes	Post-Treatment Window: 3 Years		Post-Treatment Window: 6 Years	
	Point Estimate	Standard Error	Point Estimate	Standard Error
	(1)	(2)	(3)	(4)
SNAP				
Log Individuals	-0.1014***	(0.0261)	-0.0745***	(0.0212)
Log Cases	-0.1034***	(0.0234)	-0.0800***	(0.0181)
Log Dollars	-0.0880***	(0.0272)	-0.0646***	(0.0212)
TANF				
All Recipients				
Log Individuals	-0.1895***	(0.0532)	-0.1321**	(0.0546)
Log Cases	-0.1745***	(0.0445)	-0.1053**	(0.0485)
Log Dollars	-0.1502***	(0.0483)	-0.0853*	(0.0512)
Child-Only Cases				
Log Individuals	-0.1481***	(0.0512)	-0.0747	(0.0616)
Log Cases	-0.1222***	(0.0456)	-0.0526	(0.0541)
Log Dollars	-0.1063**	(0.0469)	-0.0326	(0.0555)
Cases with Adult				
Log Individuals	-0.2119***	(0.0622)	-0.2085***	(0.061)
Log Cases	-0.2129***	(0.0526)	-0.1985***	(0.055)
Log Dollars	-0.1741***	(0.0576)	-0.1712***	(0.0616)
Medicaid				
Log Individuals	-0.0281*	(0.0153)	-0.0271*	(0.014)
Log Cases	-0.0318***	(0.0113)	-0.0334***	(0.0108)
County-Months	7,200		10,500	

*** p<0.01, ** p<0.05, * p<0.1

Data Sources: Administrative SNAP, TANF, and Medicaid records for Indiana (2004-2014), Census pop. estimates (2004-2014)
Notes: This table shows regression estimates of various log enrollment measures (total cases, individuals, or benefit dollars) on a binary indicator for receiving IBM automation and being in the post-treatment period, controlling for county- and month-fixed effects as well as county- and year-varying covariates that include log total population and population subgroups by race, gender, and age. Observations are at the county-month level and are weighted by the county-specific enrollment volume in September 2007. Marion and Lake counties (which are untreated) are excluded from the sample given that they are outliers on a number of dimensions. All regressions use a pre-treatment window of 12 quarters before automation. Columns 1 and 2 show regression estimates using a post-treatment window of 12 quarters after automation, and Columns 3 and 4 show regression estimates using a post-treatment window of 24 quarters after automation. Standard errors are clustered at the county level. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization number: CBDRB-FY2021-CES005-021.

Table 1.3 shows static estimates summarizing the effects of IBM's automated system on log total cases, individuals, and benefit dollars, using post-treatment windows of 3 years and 6 years. These estimates correspond to γ in equation (1.2). For SNAP, IBM's automated system leads to statistically significant decreases in individuals, cases, and benefit dollars of 10.1%, 10.3%, and 8.8% over the three years following treatment. Using a longer post-treatment window of six years, the treatment effects attenuate slightly to 7.4%, 8.0%, and 6.5% for individuals, cases, and benefit dollars – although they remain statistically significant at the 1% level. For TANF, the automated system leads to reductions in individuals, cases, and benefit dollars of 19.0%, 17.5%, and 15.0% over the three years after treatment and 13.2%, 10.5%, and 8.5% over the six years after treatment. The short-run enrollment declines for TANF, however, are more substantial for cases with an adult than for child-only cases, with the long-run declines driven almost entirely by reductions among cases with an adult. Finally, Medicaid individuals and cases decline by 2.8% and 3.2% in the three years following automation, with these effects remaining relatively uniform at 2.7% and 3.3% when calculated using a post-treatment window of six years.[50]

In summary, the negative effects of IBM's automated system on program enrollment are initially largest for TANF and smallest for Medicaid, while the long-run effects are most persistent for Medicaid and least persistent for TANF (a reversal of the patterns for initial enrollment). In Section 1.7, this paper explores various hypotheses underlying these patterns in the short- and long-run effects of IBM automation on program enrollment.

1.5.2 Disaggregating Enrollment Effects by Entry and Exit

Given that IBM's automated system imposed administrative burdens at multiple stages of the application process, it is useful to disaggregate the extent to which the reductions in total enrollment are driven by decreased entry versus increased exit. Let $Entrants_{ct}$ and $Exiters_{ct}$ denote

[50] After clustering standard errors at the region level using the wild cluster bootstrap, the short-run estimates for SNAP and TANF enrollment remain statistically significant at the 10% level while the short-run estimates for Medicaid enrollment become statistically insignificant. In the long-run, only the SNAP enrollment estimates remain statistically significant at the 10% level after clustering at the region level.

the number of individual recipients in county c entering and exiting a program, respectively, between year-months t and $t + 1$.[51] Monthly entry and exit rates can then be defined as follows:

$$Entry\ Rate_{ct} = \frac{Entrants_{ct}}{Population_{c,t+1} - Recipients_{ct}}$$

$$Exit\ Rate_{ct} = \frac{Exiters_{ct}}{Recipients_{ct}}$$

where $Population_{c,t+1}$ denotes total population in county c in month $t + 1$ and $Recipients_{ct}$ denotes total recipients in county c and month t. Scaling these rates by the relevant exposed populations allows one to approximate the probabilities of entering a program conditional on not being enrolled and exiting a program conditional on being enrolled, which are the transition probabilities in a two-state Markov model.

Table 1.4 shows the effects of IBM's automated system on entry and exit rates, estimated using the regression framework in equation (1.2) and a post-treatment window of 3 years.[52] First, IBM automation causes statistically significant declines in entry rates ranging from 0.03 percentage points to 0.05 percentage points - relatively similar effects - for all three programs. Because average entry rates are lowest for TANF (0.002%) and highest for SNAP (0.009%), the declines in entry are larger for TANF and smaller for SNAP when scaled by the relevant baseline means. Furthermore, even though the static percentage point effects are similar across programs,

[51] Using the longitudinal welfare records, entry and exit indicators are calculated based on PIK identifiers rather than individual IDs assigned by the welfare agency. Whereas PIKs are objective identifiers assigned by the Census Bureau (based on a recipient's Social Security Number), the individual IDs assigned by the welfare agency could have changed endogenously if IBM assigned different IDs to individuals who churned back onto a program. In practice, whether entry and exit indicators are calculated using PIKs or IDs assigned by the welfare agency yields little to no differences in the results.

[52] When entry rate is the regression outcome, the covariates include the mean shares of race/ethnicity, gender, and age groups among non-recipients in X_{ct} and county-months are weighted by the number of non-recipient individuals measured in September 2007. When exit rate is the regression outcome, the covariates include the mean shares of race/ethnicity, gender, and age groups among recipients in X_{ct} and county-months are weighted by the number of recipient individuals measured in September 2007.

Table 1.4. Treatment Effects on Monthly Entry and Exit Rates

Outcomes	Treatment Effects: Point Estimate (1)	Treatment Effects: Standard Error (2)	Pre-Treatment Mean for Treated Counties (3)	Ratio of (1) to (3) (4)
SNAP				
Entry Rate	-0.00055***	(0.00018)	0.0085	0.0642
Exit Rate	0.0074***	(0.0014)	0.059	0.1249
TANF				
All Recipients				
Entry Rate	-0.00029***	(0.00005)	0.0018	0.1625
Exit Rate	0.0105*	(0.0061)	0.1104	0.0408
Child-Only Cases				
Entry Rate	-0.00005***	(0.00001)	0.0003	0.1446
Exit Rate	-0.0089	(0.0074)	0.087	-0.1019
Cases with Adult				
Entry Rate	-0.00026***	(0.00005)	0.0015	-0.1804
Exit Rate	0.0175**	(0.0072)	0.1211	0.1443
Medicaid				
Entry Rate	-0.00036***	(0.00006)	0.00522	-0.0697
Exit Rate	0.0004	(0.0004)	0.02968	0.0143
County-Months		7,200		

*** p<0.01, ** p<0.05, * p<0.1

Data Sources: Administrative SNAP, TANF, and Medicaid records for Indiana (2004-2011), Census pop. estimates (2004-2014)
Notes: This table shows regression estimates of monthly entry and exit rates on a binary indicator for receiving IBM automation and being in the post-treatment period, controlling for county- and month-fixed effects as well as a set of county- and year-varying covariates. The entry rate in month t is defined as the number of recipients entering between month *t* and month *t* + 1 as a share of relevant non-recipients (total population in month *t* + 1 minus recipients in month *t*). The exit rate in month *t* is defined as the number of recipients exiting between month *t* and month *t* + 1 as a share of recipients in month *t*. Observations are at the county-month level. Regression estimates when entry rate is the outcome control for the mean shares of race/gender/age groups among non-recipients and weight counties by their number of non-recipients in September 2007. Regression estimates when exit rate is the outcome control for the mean shares of race/gender/age groups among recipients and weight counties by their number of recipients in September 2007. Marion and Lake counties (which are untreated) are excluded from the sample given that they are outliers on a number of dimensions. Columns 1 and 2 show the main treatment effects (with regressions using a pre-treatment window of 12 quarters before automation and a post-treatment window of 12 quarters after automation). Column 3 shows the baseline average (weighted) entry and exit rates for treated counties, measured prior to treatment (September 2007). Standard errors are clustered at the county level. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization number: CBDRB-FY2021-CES005-021.

there are differences in the dynamic patterns across programs. The entire decline in entry for SNAP is concentrated in the first two quarters of treatment, while the decline in entry for TANF and Medicaid is initially smaller and more protracted. For SNAP, the rapid bounceback in entry (and slight increase in subsequent entry rates) can likely be explained by churn among those recently exiting SNAP.

Next, the results show considerable differences across programs in the effect of IBM's automated system on exit rates. Specifically, IBM automation causes an increase of 1.05 percentage points (statistically significant at the 10% level) in the exit rate for TANF, an increase of 0.74 percentage points (statistically significant at the 1% level) in the exit rate for SNAP, and a statistically insignificant increase of 0.04 percentage points in the exit rate for Medicaid. When limiting TANF recipients to TANF cases with an adult, the results display an even more substantial increase in the exit rate of 1.75 percentage points (statistically significant at the 5% level).[53] These cross-program differences in the effects on exit rates are likely a function, in part, of differences across programs in their baseline exit rates.

1.6 Who is Screened Out by IBM Automation?

To shed light on the welfare implications of the policy change, this section investigates what types of recipients are screened out by IBM's automated system. This first subsection examines what types of counties experience the largest enrollment declines, while the second subsection explicitly analyzes how targeting efficiency (measured using proxies of individual well-being) changes as a result of IBM automation. Motivated by the analyses in the prior section on entry and exit rates, these targeting analyses focus on distinguishing between those screened out at entry versus exit.

[53] It makes sense that the increase in the TANF exit rate is driven entirely by increased exit among cases with an adult, as child-only cases are likely to face very little risk of not being recertified (given that they have little to no income information and are not subject to work requirements).

1.6.1 Heterogeneous Effects by County-Level Characteristics

The following regression specification, which extends the framework in equation (1.2), analyzes what types of counties experience the largest enrollment declines:

$$y_{ct} = \mu_c + \lambda_t + \gamma D_{ct} + \gamma^j D_{ct} A_c^j + \Sigma_{l \in J \setminus \{j\}} \gamma^l D_{ct} A_c^l + \beta X_{ct} + \varepsilon_{ct}, \qquad (1.3)$$

where y_{ct} denotes the log total number of individuals receiving a program in county c and year-month t. D_{ct} is a dummy variable equaling one if county c receives treatment and month t is after treatment. A_c^j is a time-invariant dummy variable indicating whether or not county c has characteristic j (e.g., high-poverty, low-population), given a set of county-level characteristics J.

The baseline effect of IBM's automated system on log enrollment is given by γ, and γ^j is the additional effect on enrollment for counties with a given characteristic j (holding constant the interactive effects of treatment with all other characteristics l). The union of A_c^j and the A_c^l's includes indicators for four characteristics (whether a county has an above-median unemployment rate, above-median poverty rate, below-median population size, and above-median share of non-white individuals) measured in September 2007 (prior to treatment), as well as indicators for whether a county is treated in Wave 2 and Wave 3 (relative to Wave 1). There is no need to include A_c^j and the A_c^l's as separate regressors, as they are subsumed by the county-fixed effects μ_c. X_{ct} is again a vector of county- and time-varying covariates.

Table 1.5 presents static regression estimates of γ and the γ^j's in equation (1.3) for log SNAP, TANF, and Medicaid individuals, estimated using a post-treatment window of 3 years (Columns 1-3) and a post-treatment window of 6 years (Columns 4-6). Counties with higher poverty rates have additional declines in SNAP, TANF, and Medicaid enrollment of 8.3, 14.0, and

1.8 percentage points beyond the baseline effects, respectively, during the three years after treatment. These marginal effects amount to 51%, 45%, and 31% of the baseline treatment effect, although only the differences for SNAP and TANF are statistically significant. As a share of the baseline treatment effect, these interactive effects become even larger (86% for SNAP and TANF and 49% for Medicaid) when calculated over a longer post-treatment window of six years. Counties with smaller populations and higher shares of non-white individuals also tend to experience larger declines in program enrollment, although none of these differences are statistically significant.

In contrast, counties with higher unemployment rates experience smaller declines in SNAP, TANF, and Medicaid enrollment of 8.3, 2.5, and 4.8 percentage points relative to the baseline effects over the three years after treatment. These marginal effects correspond to 51%, 8%, and 82% of the baseline treatment effect, with only the differences for SNAP and Medicaid being statistically significant. Over the six years after treatment, higher-unemployment counties experience enrollment effects that are 75%, 29%, and 117% smaller than baseline for SNAP, TANF, and Medicaid. While this result appears to qualitatively contrast with the patterns for the other characteristics (which show that more "disadvantaged" counties tend to see larger reductions in enrollment), it is likely a result of counties with higher unemployment having recipients with more income sources (who are more at risk of being affected by IBM automation since they have more information to verify).

Finally, even after controlling for interactions with other characteristics, counties treated in later waves have smaller reductions in program enrollment than those treated in Wave 1. These differences are most pronounced for TANF and also statistically significant for SNAP. One reason for these differences is that counties treated earlier in calendar time experienced the termination

Table 1.5. Treatment Effects on Enrollment by County-Level Characteristics

Regressors	Post-Treatment Window: 3 Years			Post-Treatment Window: 6 Years		
	SNAP	TANF	Medicaid	SNAP	TANF	Medicaid
	(1)	(2)	(3)	(4)	(5)	(6)
D_{ct}	-0.1624***	-0.3085***	-0.0582**	-0.0950**	-0.1845	-0.042
	(0.0478)	(0.0851)	(0.024)	(0.0476)	(0.1148)	(0.0256)
$D_{ct} \times (High\ Unemp)_c$	0.0827**	0.0251	0.0478**	0.0715*	0.0543	0.0490**
	(0.0349)	(0.0508)	(0.0197)	(0.0392)	(0.0736)	(0.0227)
$D_{ct} \times (Small\ Pop)_c$	-0.0395	-0.0289	-0.0218	-0.0445	-0.0455	-0.0196
	(0.027)	(0.0487)	(0.0137)	(0.0282)	(0.0604)	(0.0142)
$D_{ct} \times (High\ Poverty)_c$	-0.0827***	-0.1397***	-0.0181	-0.0821***	-0.1580**	-0.0206
	(0.0267)	(0.0459)	(0.0143)	(0.0303)	(0.0642)	(0.0162)
$D_{ct} \times (High\ Non\text{-}White)_c$	-0.018	0.0458	-0.0113	-0.0358	-0.0177	-0.014
	(0.0277)	(0.0539)	(0.0129)	(0.0296)	(0.0719)	(0.0144)
$D_{ct} \times (Wave\ 2)_c$	0.0881***	0.3399***	-0.0015	0.0705**	0.2841***	-0.0182
	(0.0297)	(0.0501)	(0.0155)	(0.0327)	(0.0639)	(0.0173)
$D_{ct} \times (Wave\ 3)_c$	0.1299***	0.1697***	0.0453**	0.0964***	0.1488**	0.0243
	(0.031)	(0.0606)	(0.0188)	(0.0336)	(0.0723)	(0.0204)
County-Months		7,200			10,500	

*** p<0.01, ** p<0.05, * p<0.1

Data Sources: Administrative SNAP, TANF, and Medicaid records for Indiana (2004-2014), Census pop. estimates (2004-2014)
Notes: This table shows regression estimates of various log enrollment measures (total cases, individuals, or benefit dollars) on a binary indicator for receiving IBM automation and being in the post-treatment period, controlling for county- and month-fixed effects as well as county- and year-varying covariates that include log total population and population subgroups by race, gender, and age. Observations are at the county-month level and are weighted by the county-specific enrollment volume in September 2007. Marion and Lake counties (which are untreated) are excluded from the sample given that they are outliers on a number of dimensions. All regressions use a pre-treatment window of 12 quarters before automation. Columns 1 and 2 show regression estimates using a post-treatment window of 12 quarters after automation, and Columns 3 and 4 show regression estimates using a post-treatment window of 24 quarters after automation. Standard errors are clustered at the county level. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization number: CBDRB-FY2021-CES005-021.

of IBM's contract later in event time. The counties treated in Wave 1 were thus exposed to the burdens for longer than the counties treated in Waves 2 or 3. Another reason is that earlier-treated counties were less able to foresee (and therefore more overwhelmed by) the difficulties associated with IBM's automated system, while the counties treated later were better able to anticipate and prepare for IBM automation.

1.6.2 Effects on Targeting Efficiency

To more precisely study the individual-level targeting effects of IBM's automated system, the following specification estimates how average well-being among individual recipients changes in treated counties relative to untreated counties before and after treatment:

$$\bar{y}_{ct} = \mu_c + \lambda_t + \gamma D_{ct} + \beta X_{ct} + \varepsilon_{ct}, \tag{1.4}$$

where $\bar{y}_{ct}$ is some average characteristic for recipients in county c and year-month t. Not only is this regression estimated separately for recipients of each program (SNAP, TANF, and Medicaid), but – for a given program – it is also estimated for all remaining recipients and separately for entrants and exiters.[54] Doing so allows for a novel examination of the degree to which targeting effects vary across application stages.

This section measures targeting using a variety of proxies for individual well-being.[55] The first set of outcomes encompasses various income measures linked from IRS tax records, corresponding to calendar year 2007 or earlier (thus uncontaminated by treatment). These measures include three continuous income measures: tax income in 2007, formal sector wages in

[54] When estimated over all recipients of a given program, the covariates include the mean shares of race/ethnicity, gender, and age groups among the entire population and county-months are weighted by the number of recipients in September 2007. When estimated over entrants, the covariates include the mean shares of race/ethnicity, gender, and age groups among non-recipients and county-months are weighted by the number of non-recipients in September 2007. When estimated over exiters, the covariates include the mean shares of race/ethnicity, gender, and age groups among recipients and county-months are weighted by the number of recipients in September 2007.

[55] In an ideal world, targeting would be measured based upon some perfect metric of ability or earnings capacity. however, such a measure is unobservable.

2007, and tax income spanning 2005-2007.[56] They also include binary indicators for having earnings in 2007 and asset income in 2007.[57]

A variety of outcomes also come directly from the administrative program records. For SNAP and TANF, average benefit dollars per person can be thought of as a measure of need from the welfare agency's perspective (as higher-benefit recipients are typically considered to be needier as they have lower incomes and/or greater expenses). The length of one's spell (namely the upcoming spell for entrants and the terminating spell for exiters) can also implicitly serve as a proxy for disadvantage. Program records also contain various demographic indicators for whether an assistance unit has an elderly member, a disabled member, a single parent, multiple parents, and – in the case of TANF – no adults. The administrative records for SNAP and TANF also contain information on a recipient's highest education level, which is converted into years of education.[58]

Another targeting outcome is a composite "deprivation index", constructed using predicted values from the SIPP based upon demographic and income variables in the administrative program records.[59] Higher values of the deprivation index signify greater need (and vice-versa). Whether

[56] Tax income is defined as adjusted gross income from Form 1040 for filers and wages and retirement income from Forms W-2 and 1099-R for non-filers. Formal sector wages are defined as wages, tips, and other compensation (Box 1) from Form W-2. To account for differences among assistance units in size and composition, an equivalence scale of the form $(A + PK)^F$ is applied, where A and K respectively designate the number of adults and children in the assistance unit (Citro and Michael 1995). Following Meyer and Sullivan (2012b), P and F are set to equal 0.7. Incomes are then re-scaled to be representative of an assistance unit with two adults and two children.

[57] One is defined as having earnings if they have either wages reported on a 1040 or W-2 or self-employment income reported on a 1040 or 1099-MISC. One is defined as having asset income if they receive either a Form 1099-INT or Form 1099-DIV.

[58] When examining targeting using income or education measures, the covariates in equation (1.4) additionally include the average age and age-squared of the assistance unit head (at the county- and month-level).

[59] This deprivation index is constructed based upon 32 measures of material hardships, home quality problems, food security issues, health problems, and lack of appliances taken from the topical modules in Wave 5 of the 2004 SIPP Panel and Wave 6 of the 2008 SIPP Panel. An advantage of looking at these measures is that they are relatively objective measures of material need (see, e.g., Meyer Wu, Mooers et al. 2021, Meyer, Wu, & Curran 2021). A disadvantage of using this deprivation index is that variation in the index tends to be small, as the R-squared from the prediction equation is about 0.1.

or not a recipient is non-white may also shed light on racial inequities within the welfare application and recertification process (even if such an indicator is not necessarily an objective proxy for "disadvantage"). Finally, for Medicaid, an indicator for whether or not a recipient receives "home care" can provide a gauge of being in poor health.

Focusing on SNAP, Table 1.6a presents targeting estimates for all remaining recipients, entrants, and exiters. The coefficients in Columns 1, 5, and 9 pertain to γ in equation (1.4), calculated over a post-treatment window of 3 years. Examined over all remaining recipients, IBM automation leads to slightly better targeting for 12 out of 13 outcomes, with estimates statistically significant at the 5% level for 10 outcomes. For example, those remaining on SNAP have 2.8% lower 3-year incomes, 0.8% fewer years of education, 1.5% higher benefit dollars per recipient, and 2.5% higher disability rates in the treated counties.

Yet, an interesting pattern appears when disaggregating targeting effects across entrants and exiters. IBM automation appears to be an effective screen at the entry stage for all 14 outcomes, as those entering SNAP in the treated counties tend to be less well-off than their untreated counterparts (with estimates statistically significant at the 5% level for 13 outcomes).[60] For example, those entering SNAP have 8.9% lower 3-year incomes, 2.0% fewer years of education, 8.1% higher benefit dollars per recipient, 6.7% higher disability rates, and 12.4% longer spells in the treated counties. Conversely, IBM automation appears to be an ineffective screen at the exit stage for all 14 outcomes, as those exiting SNAP in the treated counties tend to be less well-off than their untreated counterparts (with estimates statistically significant at the 5% level

[60] Spell length is an additional outcome examined for entrants and exiters, even though it is not examined for all remaining recipients.

Table 1.6a. Treatment Effects on the Characteristics of SNAP Recipients (Targeting)

Outcomes	All Remaining Recipients				Entrants				Exiters			
	Point Estimate	Standard Error	Base Mean		Point Estimate	Standard Error	Base Mean		Point Estimate	Standard Error	Base Mean	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Log Benefit $s/Person	0.0149***	(0.0034)	103	+	0.0814***	(0.0097)	90	+	0.0355***	(0.005)	100	–
Spell Length (mos.)					1.874***	(0.2311)	15.1	+	1.125***	(0.1369)	12.1	–
Log 3-Year Tax Inc.[1]	-0.0277**	(0.0135)	38,620	+	-0.0892***	(0.017)	65,150	+	-0.0789***	(0.0133)	58,870	–
Log Tax Inc.[1]	-0.0335**	(0.0153)	14,900	+	-0.0915***	(0.0201)	25,200	+	-0.0807***	(0.0153)	24,540	–
Log Wages (W-2)[1]	-0.0444***	(0.0152)	6,735	+	-0.1112***	(0.0191)	12,160	+	-0.1017***	(0.0162)	12,310	–
Has Earnings[1]	-0.0067**	(0.0028)	0.794	+	-0.0138***	(0.0036)	0.869	+	-0.0109***	(0.0023)	0.85	–
Has Asset Inc.[1]	-0.0007	(0.0012)	0.074	+	-0.0072***	(0.0019)	0.101	+	-0.0046***	(0.0018)	0.086	–
Years of Education	-0.0871***	(0.0268)	10.9	+	-0.2259***	(0.0426)	11.1	+	-0.0918***	(0.0269)	11.1	–
Has Elderly Member	0.0061***	(0.0016)	0.105	+	0.0046***	(0.0014)	0.033	+	0.0030*	(0.0016)	0.046	–
Has Disabled Member	0.0078***	(0.0027)	0.316	+	0.0114***	(0.0027)	0.169	+	0.0107***	(0.0023)	0.211	–
Single Parent	0.0076**	(0.0033)	0.328	+	0.0254***	(0.0051)	0.337	+	0.0162***	(0.0044)	0.333	–
Multiple Parents	-0.0075***	(0.0027)	0.161	+	-0.0108***	(0.0035)	0.188	+	-0.0044	(0.0028)	0.197	–
Deprivation (SIPP)	-0.0002	(0.0002)	0.195	–	0.0015***	(0.0003)	0.191	+	0.0023***	(0.0003)	0.185	–
Non-White	0.0035	(0.0046)	0.178	+	0.005	(0.0038)	0.168	+	0.0061**	(0.0028)	0.173	–
County-Months		7,200				7,200				7,200		

*** p<0.01, ** p<0.05, * p<0.1

[1] All tax income measures are calculated for 2007 (the year before treatment was fully rolled out) and 3-year tax income is calculated for 2005-2007

Data Sources: Administrative SNAP records for Indiana (2004-2014), Census population estimates (2004-2014), IRS Forms 1040, W-2, and 1099-R (2005-2007), Survey of Income and Program Participation (2004 Panel, Wave 5 and 2008 Panel, Waves 6 and 9)

Notes: This table shows regression estimates of average SNAP recipient characteristics on a binary indicator for receiving IBM automation and being in the post-treatment period, controlling for county- and month-fixed effects as well as a set of county- and year-varying covariates. Observations are at the county-month level. Regression estimates for all remaining recipients (Columns 1-3) control for the mean shares of race/gender/age groups for the entire population and weight counties by the number of recipients in September 2007. Regression estimates for entrants control for the mean shares of race/gender/age groups among non-recipients and weight counties by the number of non-recipients in September 2007. Regression estimates for exiters control for the mean shares of race/gender/age groups among recipients and weight counties by their number of recipients in September 2007. The continuous income outcomes are equivalized using the NAS equivalence scale to reflect the number of children and adults in the assistance unit. For income/education outcomes, regression estimates additionally control for average age and age-squared of the case head. Marion and Lake counties (which are untreated) are excluded from the sample given that they are outliers on a number of dimensions. Columns 1-2, 5-6, and 9-10 show the main treatment effects for all remaining recipients, entrants, and exiters, respectively (with regressions using a pre-treatment window of 12 quarters before automation and a post-treatment window of 12 quarters). Standard errors are clustered at the county level. Columns 3, 7, and 11 show the (weighted) baseline means for each outcome among all recipients, entrants, and exiters (respectively) in treated counties in September 2007. Columns 4, 8, and 12 show indicators for whether IBM automation improves (+) or worsens (–) targeting efficiency. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization number: CBDRB-FY2021-CES005-021.

Table 1.6b. Treatment Effects on the Characteristics of TANF Recipients (Targeting)

Outcomes	All Remaining Recipients				Entrants				Exiters			
	Point Estimate	Standard Error	Base Mean		Point Estimate	Standard Error	Base Mean		Point Estimate	Standard Error	Base Mean	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Log Benefit $s/Person	0.0421***	(0.006)	97	+	0.1537***	(0.0123)	107	+	0.0327***	(0.0077)	93	–
Spell Length (mos.)					0.13	(0.2103)	7.9	+	0.754***	(0.2251)	8.7	–
Log 3-Year Tax Inc.[1]	-0.0443	(0.0384)	21,230	+	-0.0598	(0.043)	33,370	+	-0.0249	(0.0253)	24,780	–
Log Tax Inc.[1]	-0.0474	(0.0344)	6,368	+	-0.0679	(0.0436)	10,030	+	-0.0411	(0.0347)	8,369	–
Log Wages (W-2)[1]	-0.0738*	(0.04)	3,243	+	-0.0645	(0.0534)	6,093	+	-0.0461	(0.0308)	4,941	–
Has Earnings[1]	-0.0294***	(0.0101)	0.531	+	-0.017	(0.0125)	0.731	+	-0.0052	(0.0054)	0.618	–
Has Asset Inc.[1]	-0.0011	(0.0018)	0.04	+	-0.0014	(0.0035)	0.062	+	0.0011	(0.0021)	0.049	+
Years of Education	-0.3366**	(0.1526)	11.1	+	-0.2402**	(0.1184)	11.3	+	-0.0686**	(0.0315)	11.1	–
Has Elderly Member	0.0040***	(0.001)	0.024	+	-0.001	(0.0012)	0.018	–	0.0011	(0.0018)	0.037	–
Has Disabled Member	0.0267***	(0.0095)	0.416	+	0.0037	(0.008)	0.224	+	-0.0207***	(0.008)	0.375	+
No Adults	-0.0010**	(0.0004)	0.193	–	-0.0008	(0.0005)	0.197	–	-0.0002	(0.0003)	0.187	+
Deprivation (SIPP)	0.0117	(0.0084)	0.271	+	0.0082	(0.0057)	0.206	+	0.0008	(0.0028)	0.266	–
Non-White	0.0421***	(0.006)	97	+	0.1537***	(0.0123)	107	+	0.0327***	(0.0077)	93	–
County-Months		7,200				7,200				7,200		

*** p<0.01, ** p<0.05, * p<0.1

[1] All tax income measures are calculated for 2007 (the year before treatment was fully rolled out) and 3-year tax income is calculated for 2005-2007

Data Sources: Administrative TANF records for Indiana (2004-2014), Census population estimates (2004-2014), IRS Forms 1040, W-2, and 1099-R (2005-2007), Survey of Income and Program Participation (2004 Panel, Wave 5 and 2008 Panel, Waves 6 and 9)

Notes: This table shows regression estimates of average TANF recipient characteristics on a binary indicator for receiving IBM automation and being in the post-treatment period, controlling for county- and month-fixed effects as well as a set of county- and year-varying covariates. Observations are at the county-month level. Regression estimates for all remaining recipients (Columns 1-3) control for the mean shares of race/gender/age groups for the entire population and weight counties by the number of recipients in September 2007. Regression estimates for entrants control for the mean shares of race/gender/age groups among non-recipients and weight counties by the number of non-recipients in September 2007. Regression estimates for exiters control for the mean shares of race/gender/age groups among recipients and weight counties by their number of recipients in September 2007. The continuous income outcomes are equivalized using the NAS equivalence scale to reflect the number of children and adults in the assistance unit. For income/education outcomes, regression estimates additionally control for average age and age-squared of the case head. Marion and Lake counties (which are untreated) are excluded from the sample given that they are outliers on a number of dimensions. Columns 1-2, 5-6, and 9-10 show the main treatment effects for all remaining recipients, entrants, and exiters, respectively (with regressions using a pre-treatment window of 12 quarters before automation and a post-treatment window of 12 quarters). Standard errors are clustered at the county level. Columns 3, 7, and 11 show the (weighted) baseline means for each outcome among all recipients, entrants, and exiters (respectively) in treated counties in September 2007. Columns 4, 8, and 12 show indicators for whether IBM automation improves (+) or worsens (–) targeting efficiency. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization number: CBDRB-FY2021-CES005-021.

Table 1.6c. Treatment Effects on the Characteristics of Medicaid Recipients (Targeting)

Outcomes	All Remaining Recipients				Entrants				Exiters			
	Point Estimate	Standard Error	Base Mean		Point Estimate	Standard Error	Base Mean		Point Estimate	Standard Error	Base Mean	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Spell Length (mos.)					1.771***	(0.3027)	31.2	+	1.343***	(0.2426)	20.1	–
Log 3-Year Tax Inc.[1]	-0.0171	(0.0133)	21,520	+	-0.0128	(0.0161)	40,540	+	-0.0344**	(0.0155)	25,200	–
Log Tax Inc.[1]	-0.0241*	(0.0135)	6,812	+	-0.0131	(0.0182)	12,700	+	-0.0537***	(0.0179)	8,803	–
Log Wages (W-2)[1]	-0.0209	(0.0193)	2,998	+	-0.0083	(0.0217)	7,416	+	-0.0372*	(0.0227)	4,460	–
Has Earnings[1]	0.0017	(0.0031)	0.576	–	0.0099***	(0.0044)	0.622	–	0.0104***	(0.0036)	0.476	+
Has Asset Inc.[1]	-0.0034***	(0.0011)	0.078	+	-0.0039	(0.0026)	0.118	+	-0.0044***	(0.0021)	0.095	–
Has Elderly Member	-0.0021	(0.0021)	0.161	–	-0.003	(0.003)	0.124	–	0.0106***	(0.0027)	0.12	–
Has Disabled Member	-0.0062***	(0.002)	0.3	–	-0.0103***	(0.0037)	0.122	–	0.0045	(0.0029)	0.173	–
Single Parent	0.0088***	(0.003)	0.221	+	0.0054	(0.0055)	0.187	+	0.0123***	(0.0026)	0.159	–
Multiple Parents	-0.0007	(0.0011)	0.047	+	0.0015	(0.0016)	0.035	–	0.0012	(0.001)	0.023	+
Deprivation (SIPP)	-0.0002	(0.0001)	0.154	–	-0.0001	(0.0002)	0.149	–	0.0001	(0.0001)	0.15	–
Non-White	0.0061*	(0.0031)	0.171	+	0.0140***	(0.0033)	0.17	+	0.0011	(0.0019)	0.156	–
County-Months		7,200				7,200				7,200		

*** p<0.01, ** p<0.05, * p<0.1

[1] All tax income measures are calculated for 2007 (the year before treatment was fully rolled out) and 3-year tax income is calculated for 2005-2007

Data Sources: Administrative Medicaid records for Indiana (2004-2014), Census population estimates (2004-2014), IRS Forms 1040, W-2, and 1099-R (2005-2007), Survey of Income and Program Participation (2004 Panel, Wave 5 and 2008 Panel, Waves 6 and 9)

Notes: This table shows regression estimates of average Medicaid recipient characteristics on a binary indicator for receiving IBM automation and being in the post-treatment period, controlling for county- and month-fixed effects as well as a set of county- and year-varying covariates. Observations are at the county-month level. Regression estimates for all remaining recipients (Columns 1-3) control for the mean shares of race/gender/age groups for the entire population and weight counties by the number of recipients in September 2007. Regression estimates for entrants control for the mean shares of race/gender/age groups among non-recipients and weight counties by the number of non-recipients in September 2007. Regression estimates for exiters control for the mean shares of race/gender/age groups among recipients and weight counties by their number of recipients in September 2007. The continuous income outcomes are equivalized using the NAS equivalence scale to reflect the number of children and adults in the assistance unit. For income outcomes, regression estimates additionally control for average age and age-squared of the case head. Marion and Lake counties (which are untreated) are excluded from the sample given that they are outliers on a number of dimensions. Columns 1-2, 5-6, and 9-10 show the main treatment effects for all remaining recipients, entrants, and exiters, respectively (with regressions using a pre-treatment window of 12 quarters before automation and a post-treatment window of 12 quarters). Standard errors are clustered at the county level. Columns 3, 7, and 11 show the (weighted) baseline means for each outcome among all recipients, entrants, and exiters (respectively) in treated counties in September 2007. Columns 4, 8, and 12 show indicators for whether IBM automation improves (+) or worsens (–) targeting efficiency. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization number: CBDRB-FY2021-CES005-021.

for 12 outcomes). Those exiting SNAP have 7.9% lower 3-year incomes, 0.8% fewer years of education, 3.6% higher benefit dollars per recipient, 5.1% higher disability rates, and 9.3% longer spells in the treated counties. Exiters in the treated counties are also 3.5% more likely to be non-white.

Tables 1.6b and 1.6c show similar patterns in targeting efficiency for TANF and Medicaid, although the estimates are noisier than those for SNAP (likely because there is less variation across characteristics for TANF recipients and smaller enrollment effects for Medicaid recipients). Moreover, the set of targeting outcomes examined differs slightly across programs, due either to a lack of data availability (e.g., years of education for Medicaid) or relevance (e.g., multiple parents or elderly member for TANF). For TANF, IBM automation again leads to better targeting overall for 10 out of 11 outcomes, with estimates statistically significant at the 5% level for 6 outcomes. Targeting efficiency improves among entrants for 10 out of 12 outcomes with 2 outcomes being statistically significant, and worsens among exiters for 9 out of 12 outcomes with 3 outcomes being statistically significant. Turning to Medicaid, IBM automation improves targeting among all remaining recipients for 8 out of 12 outcomes, with estimates being statistically significant at the 5% level for only 2 outcomes. Once again, targeting efficiency improves among entrants for 8 out of 13 outcomes with 3 outcomes being statistically significant, and worsens among exiters for 11 out of 13 outcomes with 7 outcomes being statistically significant.

Taken together, these results - suggesting that the effects on targeting efficiency may differ across entry and exit - offer a potential explanation for conflicting results in prior empirical studies (particularly those focusing on SNAP). Finkelstein and Notowidigdo (2019), for example, find that SNAP application costs lead to improved targeting when studying individuals applying for the first time. Homonoff and Somerville (2019) and Gray (2019), on the other hand, find the opposite result when studying the effects of administrative burdens at SNAP recertification.

What factors could potentially explain these stark differences in targeting efficiency across application stages? One possible explanation pertains to differences in the nature of administrative burdens - and therefore the salience of IBM's automated system - across application stages. At the initial application stage, individuals may face a variety of costs ranging from transaction costs to stigma costs and information costs. Thus, IBM's automated system (which affects only transaction costs) may play a bigger role in screening out less needy individuals who may already face stigma costs, for example. Conversely, the burdens from IBM's automated system are likely to be more salient at the recertification stage - where information and stigma costs are much less relevant - and thus screen out more needy individuals who face higher cognitive costs. A second possible explanation deals with differential selection into each application stage. Initial applicants tend on average to be better-off than recertifiers, since means-tested programs are well-targeted on average. Then, assuming that there are non-monotonic effects of IBM's automated system by well-being (i.e., the least and most needy individuals are screened out to the greatest extent), one would expect to see better targeting among entrants if the least needy individuals appear at initial application and worse targeting among exiters if the most needy individuals appear at recertification.

1.7 Additional Analyses and Robustness Checks

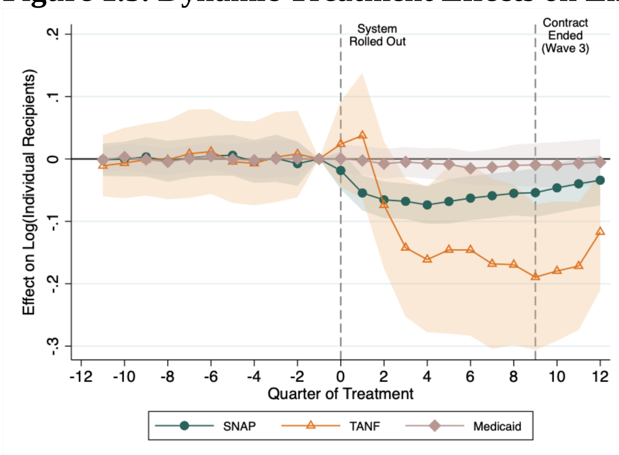
This section discusses the results of additional analyses and robustness checks. The first subsection explores explanations for cross-program differences in the short-run effects of IBM's automated system on enrollment. The next subsection provides suggestive evidence on various mechanisms underlying the persistent enrollment decline for Medicaid (and, to a lesser degree, for SNAP) in the long run. The final subsection presents a series of robustness checks that validate the difference-in-differences design and address other potential threats to identification and inference.

1.7.1 Differences in Short-Run Enrollment Effects Across Programs

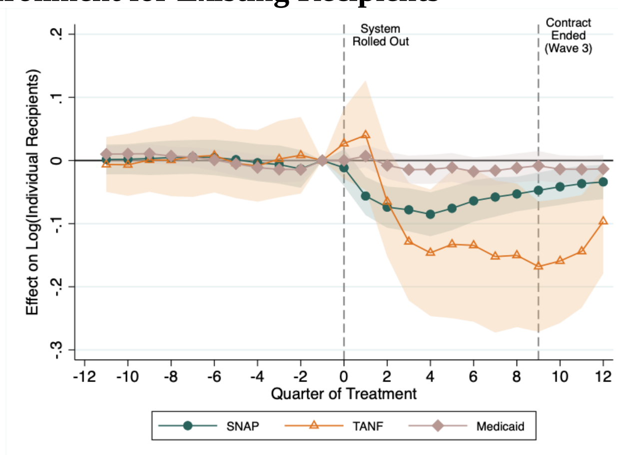
The rollout of IBM's automated system led to reductions in program enrollment that are initially largest for TANF and smallest for Medicaid. One reason for these cross-program differences may be that enrollment (particularly recertification) costs are highest for TANF and lowest for Medicaid. Recertification intervals are shortest for TANF and longest for Medicaid. TANF recipients may also receive less assistance than SNAP or Medicaid recipients to stay on the program, as TANF is designed to wean its recipients off of welfare. Conversely, those eligible for Medicaid may have more avenues for application assistance, including from healthcare providers who may have an incentive to get uninsured patients enrolled on Medicaid. Another reason for these cross-program differences in enrollment effects may be that TANF recipients tend to be more needy than Medicaid recipients (see Table 1.1). If those who are more needy are less able to cope with administrative burdens, then the differences in enrollment effects could also be driven by cross-program differences in recipient composition.

To explore these hypotheses, program-specific enrollment changes are calculated for existing recipients of multiple programs (receiving as of September 2007, the month prior to IBM rollout) and compared to the enrollment changes for existing recipients of each individual program. Similar enrollment effects across programs after holding fixed recipient composition would suggest that the original cross-program differences in enrollment effects are largely a result of differences in recipient composition. Differences in enrollment effects among multiple-program recipients would point to differing enrollment costs playing a role in explaining the original cross-

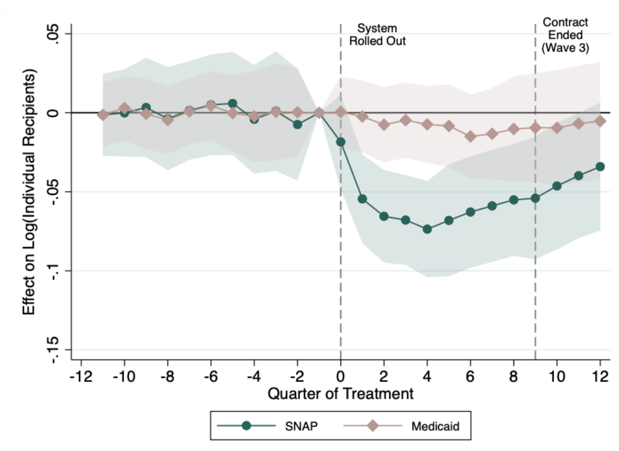
Figure 1.5. Dynamic Treatment Effects on Enrollment for Existing Recipients



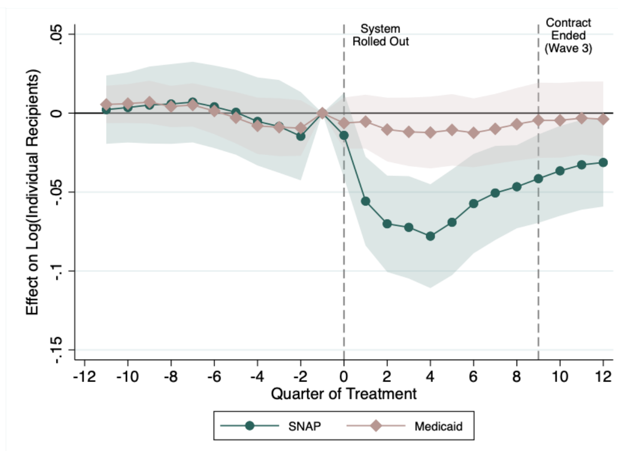
a) Recipients of Each Individual Program



b) Recipients of All Three Programs



c) Recipients of Each Individual Program



d) Recipients of All Three Programs

Data Sources: Administrative SNAP, TANF, and Medicaid records for Indiana (2005-2011), Census pop. estimates (2004-2014)
Notes: These figures show regression estimates of log total individuals receiving SNAP, TANF, or Medicaid (conditional on receiving some or all of these programs in September 2007) on binary indicators corresponding to the event-quarter relative to receiving IBM automation and being in the post-treatment period, controlling for county- and month-fixed effects as well as a series of county- and year-varying covariates that include log total population and population subgroups by race, gender, and age. Observations are at the county-month level and are weighted by the number of individuals receiving any program in September 2007. Marion and Lake counties (which are untreated) are excluded from the sample given that they are outliers on a number of dimensions. Panels A and C shows estimates conditional on receiving each individual program (regardless of multiple program receipt) in September 2007, Panel B shows estimates conditional on receiving all three programs together (SNAP, TANF, and Medicaid) in September 2007, and Panel D shows estimates conditional on receiving SNAP and Medicaid together in September 2007. Standard errors are clustered at the county level. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization number: CBDRB-FY2021-CES005-021.

program differences in enrollment effects. Moreover, conditioning on receipt prior to treatment implicitly focuses on the exit margin, which was shown to yield larger cross-program differences in treatment effects.

Table 1.7. Treatment Effects on Enrollment Among Existing Recipients

Outcomes	**Treatment Effects** **Point Estimate** **(1)**	 **Standard Error** **(2)**	**Total Treated Recipients (Sep. '07)** **(3)**
		Recipients of Each Individual Program	
Log SNAP Enrollment	-0.0418***	(0.0153)	251,000
Log TANF Enrollment	-0.1147***	(0.0385)	36,000
Log Medicaid Enrollment	-0.0002	(0.0126)	352,000
		Recipients of SNAP, TANF, and Medicaid	
Log SNAP Enrollment	-0.0451***	(0.0134)	
Log TANF Enrollment	-0.1007***	(0.0359)	31,500
Log Medicaid Enrollment	-0.0047	(0.0085)	
		Recipients of SNAP and Medicaid	
Log Individuals	-0.0420***	(0.0128)	194,000
Log Cases	-0.0035	(0.0087)	
County-Months		7,200	

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Data Sources: Administrative SNAP, TANF, and Medicaid records for Indiana (2005-2011), Census pop. estimates (2004-2014)
Notes: This table shows regression estimates of log total individuals receiving SNAP, TANF, or Medicaid (conditional on receiving some or all of these programs in September 2007) on a binary indicator for receiving IBM automation and being in the post-treatment period, controlling for county- and month-fixed effects as well as a series of county- and year-varying covariates that include log total population and population subgroups by race, gender, and age. Observations are at the county-month level and are weighted by the number of individuals receiving any program in September 2007. Marion and Lake counties (which are untreated) are excluded from the sample given that they are outliers on a number of dimensions. Panel A shows estimates conditional on receiving each individual program (regardless of multiple program receipt) in September 2007, Panel B shows estimates conditional on receiving all three programs together (SNAP, TANF, and Medicaid) in September 2007, and Panel C shows estimates conditional on receiving SNAP and Medicaid together in September 2007. All regressions use a pre-treatment window of 12 quarters before automation and a post-treatment window of 12 quarters after automation. Column 3 shows the number of unique individuals underlying the samples for each panel in treated counties. Standard errors are clustered at the county level. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization number: CBDRB-FY2021-CES005-021.

Figures 1.5a and 1.5b illustrate the effects of IBM automation on SNAP, TANF, and Medicaid enrollment for existing recipients of each individual program and of all three programs, respectively. Note that the vast majority (88%) of TANF recipients receive all three programs, compared to 13% of SNAP recipients and 9% of Medicaid recipients in treated counties. The patterns in enrollment effects remain remarkably similar regardless of whether one conditions or does not condition on multiple program receipt. Calculated over a post-treatment window of 3 years, TANF enrollment falls by 10.1% and 11.5%, SNAP enrollment falls by 4.5% and 4.2%, and Medicaid enrollment falls by a statistically insignificant 0.5% and 0.02% for all existing recipients and multiple-program recipients, respectively (Table 1.7). Focusing on existing recipients of both SNAP and Medicaid (who constitute 77% of all existing SNAP recipients and 55% of all existing Medicaid recipients), the results again show declines in SNAP and Medicaid enrollment of 4.2% and 0.4% that are very similar to the declines of 4.5% and 0.5% among all existing recipients of SNAP and Medicaid. These results suggest that differences across programs in recertification costs – rather than recipient composition – likely explain the bulk of the differences in enrollment effects across programs.

1.7.2 Mechanisms Underlying Long-Run Enrollment Declines

Contrary to the short-run patterns in enrollment effects, IBM's automated system causes long-run enrollment declines that are highly persistent for Medicaid and, to a lesser extent, SNAP - suggesting potential behavioral responses or effects on well-being among those cut off from each program. For example, those who are cut off from a program may be more likely to die, either directly as a result of not being able to access the program or indirectly in that those with higher mortality risk are more likely to be cut off. Individuals who lose access to a program may also be more likely to migrate, either to an untreated county in Indiana or to another state for which program enrollment is not observed. The lack of a full enrollment rebound could also be driven by an increase in incomes (e.g., through increased employment) that is sufficiently large to remove some individuals from program eligibility. Finally, individuals who are cut off from Medicaid may rely on Medicare (if they are elderly or disabled) to imperfectly substitute for their lost Medicaid coverage.

To test the role of each hypothesis in explaining the persistent long-run enrollment declines, the effects of IBM's automated system on enrollment are re-estimated after removing certain groups from the enrollment counts for all counties (treated and untreated) and months. One then examines the degree to which the persistence in the enrollment declines diminishes. For example, to test the role of mortality, all individuals who die between 2005-2014 (identified using death dates linked from the SSA Numident) are removed from the sample and the effects of IBM automation are re-estimated among those who are alive during the entire sample period. A less persistent enrollment decline among this group would suggest that mortality plays some role in the permanence of overall enrollment reduction.

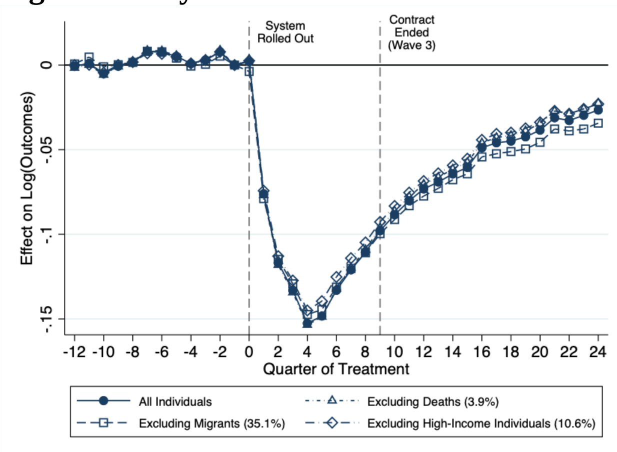
To test the role of migration, all individuals who appear in multiple counties or states between 2005-2014 are removed.[61] To test the role of increased income, all individuals with at least 3 consecutive years of tax income above twice the individual poverty line between 2005-2014 are removed. To test the role of Medicare reliance (for Medicaid), all individuals receiving Medicare between 2005-2014 are removed. It is worth noting that these analyses are not meant to provide *causal* estimates of losing program access on downstream outcomes. Rather, they shed light on whether these factors can serve as *mechanisms* in explaining why program enrollment in treated counties (relative to untreated counties) does not recover to pre-treatment levels long after IBM automation is disbanded.

Figure 1.6 presents the effects of IBM automation on program enrollment after excluding relevant groups from the sample. Figure 1.6a reveals that the increased likelihoods of mortality

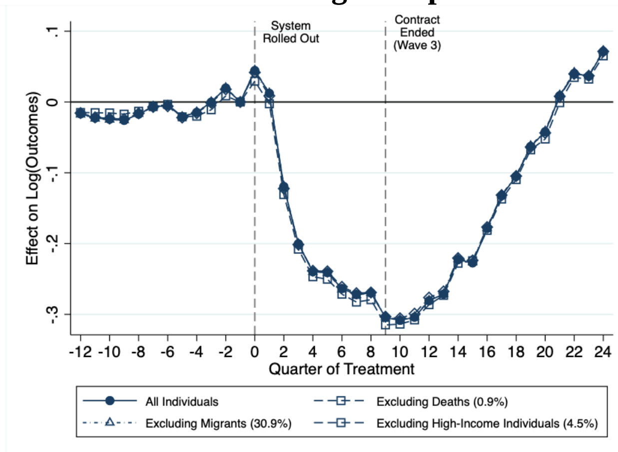
[61] Migrants are identified using information on county and state of residence linked from the 2010 Decennial Census and IRS 1099 Information Return Master File.

and Medicare receipt in the treated counties play a large role in explaining the persistence of the Medicaid enrollment decline. Six years after IBM's rollout (and four years after its termination),

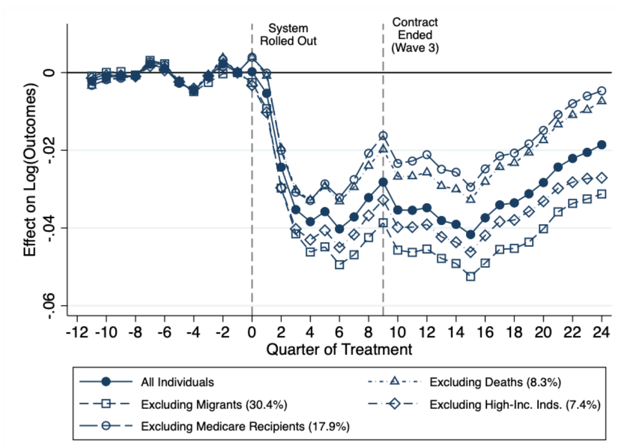
Figure 1.6. Dynamic Treatment Effects on Enrollment After Excluding Groups



a) SNAP



b) TANF



c) Medicaid

Data Sources: Administrative SNAP, TANF, and Medicaid records for Indiana (2004-2014), Census population estimates (2004-2014), IRS Forms 1040, W-2, and 1099-R (2005-2014), IRS information returns master file (2005-2014), SSA Numident, 2010 Decennial Census, CMS Medicare enrollment data (2005-2014)

Notes: These figures show regression estimates of log individuals on binary indicators corresponding to the event-quarter relative to receiving IBM automation, controlling for county- and month-fixed effects as well as county- and year-varying covariates that include log total population and population subgroups by race, gender, and age. Observations are at the county-month level and are weighted by the number of recipient individuals in September 2007 (month prior to initial IBM rollout). Effects are calculated for log total individuals as well as log individuals after removing certain groups of individuals from the enrollment counts for all counties and months. Marion and Lake counties (which are untreated) are excluded from the sample given that they are outliers on a number of dimensions. Standard errors are clustered at the county level. The shares of total unique recipients in all counties (except Marion and Lake) between 2005-2014 that are removed by each category are reported in parentheses in the legend. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization number: CBDRB-FY2021-CES005-021.

Table 1.8. Treatment Effects on Enrollment After Excluding Certain Types of Recipients

Outcomes	Post-Period: 3 Years Point Estimate (1)	Std. Error (2)	Post-Period: 6 Years Point Estimate (3)	Std. Error (4)	24th Quarter After Treatment Point Estimate (5)	Std. Error (6)	Share Excl. (7)
SNAP							
Log Individuals	-0.1014***	(0.0261)	-0.0745***	(0.0212)	-0.0266	(0.0264)	
Excl. Deaths[1]	-0.1011***	(0.0261)	-0.07262***	(0.021)	-0.0222	(0.0268)	0.039
Excl. Migrants[2]	-0.1030***	(0.0273)	-0.07689***	(0.0224)	-0.0318	(0.0274)	0.351
Excl. High-Inc. Inds.[3]	-0.0966***	(0.025)	-0.06997***	(0.0206)	-0.0242	(0.0256)	0.106
Excl. Deaths & High-Inc.	-0.0964***	(0.025)	-0.06812***	(0.0203)	-0.0198	(0.026)	0.144
TANF							
Log Individuals	-0.1895***	(0.0532)	-0.1321**	(0.0546)	0.0713	(0.0856)	
Excl. Deaths[1]	-0.1893***	(0.0532)	-0.1316**	(0.0547)	0.0728	(0.0857)	0.009
Excl. Migrants[2]	-0.2019***	(0.0519)	-0.1385***	(0.0525)	0.0709	(0.0866)	0.309
Excl. High-Inc. Inds.[3]	-0.1894***	(0.0532)	-0.1315**	(0.0547)	0.0716	(0.0861)	0.045
Medicaid							
Log Individuals	-0.0281*	(0.0153)	-0.0271*	(0.014)	-0.0186	(0.018)	
Excl. Deaths[1]	-0.0224	(0.0152)	-0.01919	(0.014)	-0.007	(0.019)	0.083
Excl. Migrants[2]	-0.0356**	(0.0145)	-0.03574***	(0.0127)	-0.0292	(0.0163)	0.304
Excl. High-Inc. Inds.[3]	-0.0325**	(0.0162)	-0.03178**	(0.0147)	-0.0264	(0.0166)	0.074
Excl. Medicare Rec.[4]	-0.0203	(0.0166)	-0.01705	(0.0152)	-0.0043	(0.0196)	0.179
Excl. Deaths & Medicare	-0.019	(0.0164)	-0.01525	(0.015)	-0.0011	(0.0197)	0.198
County-Months	7,200		10,500		10,500		

*** p<0.01, ** p<0.05, * p<0.1

[1] "Deaths" are all recipients dying between 2005 and 2014 (measured using death dates in the SSA Numident)

[2] "Migrants" are all recipients migrating between counties or across states between 2005 and 2014 (based on 2010 Decennial or IRS information returns)

[3] "High-Income Individuals" are all recipients with 3+ consecutive years of tax income above twice the individual poverty line between 2005-2014

[4] "Medicare Recipients" are all recipients who receive Medicare at some point between 2005 and 2014 (based on CMS enrollment records)

Data Sources: Admin. SNAP, TANF, & Medicaid records for Indiana (2004-2014), Census pop. estimates (2004-2014), IRS Forms 1040, W-2, & 1099-R (2005-2014), IRS info. returns master file (2005-2014), SSA Numident, 2010 Decennial Census, CMS Medicare enrollment data (2005-2014)

Notes: This table shows regression estimates of log individuals on a binary indicator for receiving IBM automation and being in the post-treatment period, controlling for county- and month-fixed effects as well as county- and year-varying covariates. Observations are at the county-month level and are weighted by the number of recipient individuals in September 2007. Effects are calculated for log total individuals as well as log individuals after removing certain groups of individuals from the enrollment counts for all counties and months. Marion and Lake counties (which are untreated) are excluded from the sample as they are outliers on a number of dimensions. All regressions use a pre-treatment window of 12 quarters before automation. Columns 1 and 2 show regression estimates using a post-treatment window of 12 quarters after automation, Columns 3 and 4 show regression estimates using a post-treatment window of 24 quarters after automation, Columns 5 and 6 show event-time regression estimates corresponding to the 24th quarter after automation, and Column 7 shows the share of total unique recipients between 2005-2014 removed by each category. Standard errors are clustered at the county level. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization number: CBDRB-FY2021-CES005-021.

accounting for mortality can explain 60% of the enrollment decline (which decreases from 1.86% to 0.74%) and accounting for Medicare receipt can explain 75% of the enrollment decline (which decreases from 1.86% to 0.47%). Taken together, mortality and Medicare receipt can explain nearly all of the long-run enrollment decline, which decreases from 1.86% to 0.16% (Table 1.8). Importantly, treatment effects for total individuals and the relevant subgroups begin to diverge when treatment is rolled out (with pre-trends being non-existent). In contrast, the enrollment reductions are slightly larger after removing migrants and those with high incomes.[62]

Turning next to SNAP, Figure 1.6b shows that the increased likelihoods of mortality and higher incomes play a small role in explaining the persistence of the decline in SNAP enrollment. Six years after IBM's rollout, accounting for mortality can explain 15% of the enrollment decline (which decreases from 2.66% to 2.25%) and accounting for higher incomes can explain 14% of the enrollment decline (which decreases from 2.66% to 2.3%). Taken together, mortality and higher incomes explain about 30% of the long-run enrollment decline, which decreases from 2.66% to 1.89% (Table 1.8). Once again, the enrollment reductions are slightly larger after removing migrants. Finally, for TANF, Figure 1.6c shows little to no differences in treatment effects measured either for all individuals or the relevant subgroups. This is unsurprising given that TANF enrollment, unlike that of Medicaid or SNAP, bounces back fully in the long run

[62] While this does not necessarily mean that individuals did not migrate or increase their labor supply in response to losing program access, it does suggest that migration and increased incomes do not appear to be mechanisms behind the lack of a full enrollment rebound for Medicaid.

1.7.3 Robustness Checks

This subsection presents a series of robustness checks that validate the difference-in-differences design and therefore a causal interpretation of the effects of IBM's automated system on program enrollment. Specifically, the analyses affirm from a variety of angles that treated counties would have had similar trends in outcomes as untreated counties in the absence of treatment.

Validity of Parallel Trends Assumption

First, the lack of pre-trends in SNAP, TANF, and Medicaid receipt persists even after extending the pre-treatment period from three years to five years. Yet, the existence of parallel trends between treated and untreated counties in the pre-treatment period does not guarantee that these trends would have persisted in the absence of treatment (Kahn-Lang & Lang 2018). For example, time-varying shocks that influence enrollment in welfare programs may affect treated and untreated counties differentially during the post-treatment period.

One example of such a shock is the Great Recession. Enrollment in welfare programs typically rises during recessions, and research has found evidence of spatial variation in Great Recession severity across local areas in the U.S. (see, e.g., Yagan 2019). Since treated counties in Indiana tend on average to be smaller, poorer, and more rural than untreated counties, one may be concerned that these counties experienced differential patterns in economic conditions before and after treatment. Yet, treated and untreated counties have similar trends in overall unemployment rates and employment-to-population ratios, as well as employment shares corresponding to the two largest economic sectors in Indiana (manufacturing and wholesale/retail trade).

A caveat with analyzing unemployment rates and employment shares is that these measures may follow different trends in treated and untreated counties as a direct result of treatment.[63] However, behavioral responses on the part of treated individuals should not be large enough to affect overall employment patterns, since those affected constitute a small minority of all working-age adults. Examining previous recessions in the early 1990s and 2000s also yields largely parallel trends in both SNAP receipt rates and unemployment rates between the counties that later received IBM's automated system and those that did not. This provides further evidence that, in the absence of treatment, trends in welfare receipt would have likely evolved in a similar fashion across treated and untreated counties during the Great Recession.

Table 1.9 shows estimates using a number of additional specifications to validate the robustness of the main difference-in-differences results. Columns 1-2 start by showing that the treatment effects on SNAP, TANF, and Medicaid enrollment are nearly identical when calculated using either the public- or restricted-use data sources. Relying on the public-use data, Columns 3-5 then show that the treatment effects on SNAP, TANF, and Medicaid enrollment are highly stable with respect to the inclusion of county- and time-varying covariates as well as to the choice of covariates used.[64] Column 6 shows estimates that exclude the five most populous counties in Indiana outside of Marion and Lake counties (two of which are untreated) and find little to no differences compared to the main effects on program enrollment. Column 7 shows estimates that exclude the 26 counties in Indiana severely impacted by floods in September 2008 (which could

[63] For example, individuals cut off from welfare by IBM's automated system may seek increased employment to offset the loss of welfare benefits.

[64] Estimates that control for no covariates beyond county- and month-fixed effects (Column 3) and estimates that control for log county population (Column 4) are similar to the preferred estimates that control for population and demographic covariates. Estimates remain largely stable after controlling for unemployment rates (Column 5), which importantly reflect local economic conditions but have the caveat that they may not be pre-determined.

Table 1.9. Treatment Effects on Enrollment (Different Specifications)

Outcomes	Main Estimate (Public-Use) (1)	Main Estimate (Restricted-Use) (2)	No Controls (3)	Population Controls Only (4)	Pop. + Demog. + Unemp. Controls (5)	Exclude Biggest Counties (6)	Exclude Flooded Counties (7)
Post-Treatment Window: 3 Years							
Log SNAP Individuals	-0.0990***	-0.1014***	-0.1085***	-0.0953***	-0.0656***	-0.0995***	-0.1049***
	(0.0258)	(0.0261)	(0.031)	(0.0302)	(0.0184)	(0.0166)	(0.0331)
Log TANF Individuals	-0.1851***	-0.1895***	-0.1856**	-0.1712**	-0.1583***	-0.1739***	-0.2109***
	(0.0403)	(0.0532)	(0.0714)	(0.0739)	(0.036)	(0.0394)	(0.0549)
Log Medicaid Individuals	-0.0297*	-0.0281*	-0.0440**	-0.0281	-0.0152	-0.0432***	-0.0226
	(0.0156)	(0.0153)	(0.0214)	(0.0198)	(0.0137)	(0.0097)	(0.0188)
County-Months	7,200	7,200	7,200	7,200	7,200	6,800	5,200
Post-Treatment Window: 6 Years							
Log SNAP Individuals	-0.0730***	-0.0745***	-0.0881***	-0.0721***	-0.0531***	-0.0809***	-0.0761***
	(0.0209)	(0.0212)	(0.0281)	(0.0254)	(0.0185)	(0.0182)	(0.0267)
Log TANF Individuals	-0.1306**	-0.1321**	-0.1375*	-0.1247	-0.1209**	-0.1697***	-0.1491**
	(0.0509)	(0.0546)	(0.0774)	(0.0793)	(0.0513)	(0.0475)	(0.0698)
Log Medicaid Individuals	-0.0310**	-0.0271*	-0.0496**	-0.0306*	-0.0234	-0.0490***	-0.0234
	(0.015)	(0.014)	(0.0225)	(0.0183)	(0.0146)	(0.0113)	(0.0181)
County-Months	10,500	10,500	10,500	10,500	10,500	9,800	7,500

*** p<0.01, ** p<0.05, * p<0.1

Data Sources: Administrative restricted- and public-use SNAP, TANF, and Medicaid records for Indiana (2004-2014), Census pop. estimates (2004-2014)

Notes: This table shows regression estimates of log total individuals on a binary indicator for receiving IBM automation and being in the post-treatment period, controlling for county- and month-fixed effects as well as county- and year-varying covariates. Columns 1 and 2 show the primary estimates from the public- and restricted-use data sources, respectively. Column 3 shows estimates without any controls, Column 4 controls only for log total population, and Column 5 controls for log total population (along with subgroups by race gender, and age) and unemployment rate. Column 6 shows estimates after removing the five largest counties in Indiana (after Marion and Lake counties). Column 7 shows estimates that exclude the 26 Indiana counties that were severely impacted by floods in September 2008. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization number: CBDRB-FY2021-CES005-021.

have led to increased disaster payments to affected areas) and again find small differences compared to the main effects.

As placebo tests, Table 1.10 compares enrollment changes between treated and untreated counties for a number of other government programs. The analyses focus on Social Security, SSI, Medicare, and free and reduced-price school meals (FARM), since these programs were not administered under IBM's automated system but have eligible populations that overlap with those eligible for SNAP, TANF, or Medicaid.[65] The placebo estimates rely on annual county-level

Table 1.10. Treatment Effects for Automated vs. Non-Automated Programs

Outcomes	Point Estimate (1)	Standard Error (2)
Automated Programs		
Log SNAP Individuals	-0.1001***	(0.0287)
Log TANF Individuals	-0.2115***	(0.0511)
Log Medicaid Individuals	-0.0308*	(0.0173)
Other Programs		
Log Social Security Recipients	-0.0031	(0.0041)
Log SSI Recipients	-0.0283**	(0.0127)
Log Medicare Recipients	-0.0017	(0.0037)
Log Free and Reduced Meals Recipients	-0.0289	(0.0194)
County-Years	630	

*** p<0.01, ** p<0.05, * p<0.1

Data Sources (public-use): County-level public-use SNAP, TANF, and Medicaid program records (2005-2011), SSA annual estimates of Social Security and SSI recipients (2005-2011), CMS Medicare enrollment counts (2007-2011), Department of Education CCD Farm counts (2005-2011), Census population estimates (2005-2011)

Notes: This table shows regression estimates of log total individuals receiving various programs - including SNAP, TANF, and Medicaid (which are automated) and Social Security, SSI, Medicare, and free and reduced meals (which are not automated) - on a binary indicator for receiving IBM automation and being in the post-treatment period, controlling for county- and month-fixed effects as well as county- and year-varying covariates that include log total population and population subgroups by race, gender, and age. Annual individuals for SNAP, TANF, and Medicaid are averaged across months in a calendar year. Annual individuals for Social Security and SSI are measured by the number of recipients in December of a given calendar year. Annual individuals receiving free and reduced meals are measured by the number of recipients in a given school year. The calendar year corresponding to the implementation of the automated system is set to 2008. All regressions also use a pre-treatment window of 2005-2007 (except for Medicare, which uses a pre-treatment window of only 2007 given data availability issues) and a post-treatment window of 2008-2011. Observations are at the county-month level and are weighted by the county-specific enrollment volume in September 2007. Marion and Lake counties (which are untreated) are excluded from the sample given that they are outliers on a number of dimensions. Standard errors are clustered at the county level.

[65] For example, children receiving SNAP are "categorically eligible" for free and reduced-price school meals, and SSI and Medicaid tend to have overlapping recipient populations.

enrollment data for these programs spanning 2005-2011, with 2008 set as the calendar year corresponding to the first year of treatment.[66] The results show statistically insignificant effects of IBM's automated system on participation in each of the non-automated programs, with the exception of SSI. It is worth noting, however, that the observed effect on SSI may actually reflect spillover effects of treatment itself (e.g., if SSI recipients are also those with higher mortality risk).

Robustness to Alternative Difference-in-Difference Estimators

A flurry of recent studies have identified issues associated with TWFE or event study specifications with variation in treatment timing (see, e.g., Callaway & Sant'Anna 2020, de Chaisemartin & D'Haultfoeuille 2020, Sun & Abraham 2020, Goodman-Bacon 2021, Borusyak et al. 2021). These issues include the potential for negative weights to arise on certain treatment effect parameters and for some earlier-treated units to be improperly used as control units for later-treated units. In light of these concerns, this subsection discusses comparisons of the main two-way fixed effects (TWFE) estimates to difference-in-differences estimates calculated using alternative methods proposed in recent studies.

Table 1.11 compares the main static TWFE estimates against two alternative estimators: a stacked difference-in-differences (DiD) estimator (see, e.g, Cengiz et al. 2019, Deshpande & Li 2019) and the estimator from Callaway and Sant'anna (CS). Columns 1 and 2 first show that the TWFE point estimates are virtually indistinguishable regardless of whether the panel is balanced on event time or calendar time. The stacked DiD estimates (in Column 3) are obtained by comparing each group of treated counties to the untreated counties and "stacking" these comparisons on top of one another. Because each unit is always treated or untreated in a given

[66] Because county-level Medicare data are available only starting from 2007, there is only a single year covering the pre-treatment period for Medicare.

Table 1.11. Treatment Effects Using Alternative Difference-in-Differences Estimators

Outcomes	Two-Way Fixed Effects	Two-Way Fixed Effects	Stacked Diff-in-Diff	Callaway-Sant'anna Estimator
	(1)	(2)	(3)	(4)
		Post-Treatment Window: 3 Years		
Log SNAP Individuals	-0.0990***	-0.0989***	-0.1145***	-0.1242***
	(0.0258)	(0.0257)	(0.0286)	(0.0279)
Log TANF Individuals	-0.1851***	-0.1857***	-0.2195***	-0.2177***
	(0.0403)	(0.04)	(0.0492)	(0.069)
Log Medicaid Individuals	-0.0297*	-0.0303*	-0.0335**	-0.0365***
	(0.0156)	(0.0156)	(0.0159)	(0.0108)
County-Months	7,200	7,400	11,400	7,400
		Post-Treatment Window: 6 Years		
Log SNAP Individuals	-0.0730***	-0.0726***	-0.0816***	-0.0995***
	(0.0209)	(0.0209)	(0.0222)	(0.0307)
Log TANF Individuals	-0.1306**	-0.1285**	-0.1635***	-0.149
	(0.0509)	(0.0515)	(0.0585)	(0.106)
Log Medicaid Individuals	-0.0310**	-0.0309**	-0.0326**	-0.0380***
	(0.015)	(0.0149)	(0.0144)	(0.0143)
County-Months	10,500	10,500	17,000	10,500
Balanced on	Event Time	Calendar Time	Event Time	Calendar Time

*** p<0.01, ** p<0.05, * p<0.1

Data Sources (public-use): County-level SNAP, TANF, and Medicaid program records (2004-2014), Census population estimates (2004-2014)

Notes: This table shows regression estimates of log total individuals receiving SNAP, TANF, or Medicaid on a binary indicator for receiving IBM automation and being in the post-treatment period, controlling for county- and month-fixed effects as well as county- and year-varying covariates that include log total population and population subgroups by race, gender, and age. Observations are at the county-month level and are weighted by the county-specific enrollment volume in September 2007. Marion and Lake counties (which are untreated) are excluded from the sample given that they are outliers on a number of dimensions. All regressions use a pre-treatment window of 12 quarters before automation. Panel A shows regression estimates using a post-treatment window of 12 quarters after automation, and Panel B shows regression estimates using a post-treatment window of 24 quarters after automation. Column 1 shows standard two-way fixed effects estimates with panel observations balanced on event time (12 quarters before treatment and 12 or 28 quarters after treatment). Column 2 shows two-way fixed effects estimates with panel observations balanced on calendar time (from October 2004 to July 2011 for Panel A and from October 2004 to July 2014 for Panel B). Column 3 shows stacked difference-in-differences estimates where the counties affected by each treatment wave are compared against never-treated counties (balanced on event time), and estimates are obtained after stacking these comparisons. Column 4 shows aggregated difference-in-differences effects from the Callaway-Sant'anna (2020) estimator that is robust to settings with more than two time periods and variation in treatment timing. Standard errors are clustered at the county level.

"stack", this approach circumvents issues with negative weights. The stacked DiD and TWFE estimates are very similar to each other (generally within 10% of each other), although the stacked

DiD estimates are always larger in magnitude. Alternatively, the CS estimator - which relies on assumptions of parallel trends and limited treatment anticipation - robustly identifies the average treatment effect at time t among the units first treated at time g. Column 4 shows the overall CS treatment effects aggregated over t and g. The patterns continue to be highly consistent with the TWFE estimates in both sign and statistical significance.

To provide some context for these patterns, we can use the methodology from Goodman-Bacon (2021) to decompose the main TWFE estimate. Depending on the length of the post-treatment period, 82-87% of the TWFE estimate comes from comparisons of treated counties to never-treated counties and only 8-11% of the TWFE estimate comes from comparisons of treated units with different treatment times.[67] The small weight on the timing group comparisons makes sense given that the staggered IBM rollout occurs over a short time frame (7 months) relative to the overall time frame for my analyses (72 or 108 months). Moreover, the treatment effect from the treated vs. never-treated comparisons is always larger in magnitude than the main TWFE estimate. This can explain why the stacked DiD and CS estimates – which rely heavily on "clean" comparisons between treated and never-treated units – are usually larger in magnitude than the TWFE estimates.

Comparison to Other States

This final part discusses alternative estimates of treatment effects involving comparisons to other states. These estimates focus on SNAP enrollment, given that public-use data on county- and state-level program receipt for the entire U.S. are more easily accessible for SNAP than for TANF or Medicaid. The first alternative specification calculates the effects of IBM's automated

[67] The remaining 4-8% of the TWFE estimate comes from within-group variation as a result of covariates.

system on SNAP enrollment in Indiana's treated counties, compared to SNAP enrollment both in Indiana's untreated counties and in border counties in neighboring states (Illinois, Kentucky, Ohio, and Michigan). These estimates (comparing to untreated counties in Indiana and counties in neighboring states) are very similar to baseline estimates that only compare to untreated counties in Indiana. These similarities suggest that SNAP enrollment in Indiana's untreated counties follow similar trends as SNAP enrollment in neighboring states during the post-treatment period.

The second specification uses a synthetic control method to compare changes in SNAP enrollment in Indiana (the single treated state) to those of a "synthetic control" state constructed using a weighted combination of untreated states (see, e.g., Abadie et al. 2010, Cunningham & Shah 2018, Jones & Marinescu 2018).[68] The weights are calculated by matching all other states to Indiana on observable pre-treatment covariates.[69] The identification strategy assumes that the synthetic control group for Indiana yields outcome patterns that would accurately represent Indiana's outcomes in the absence of treatment.

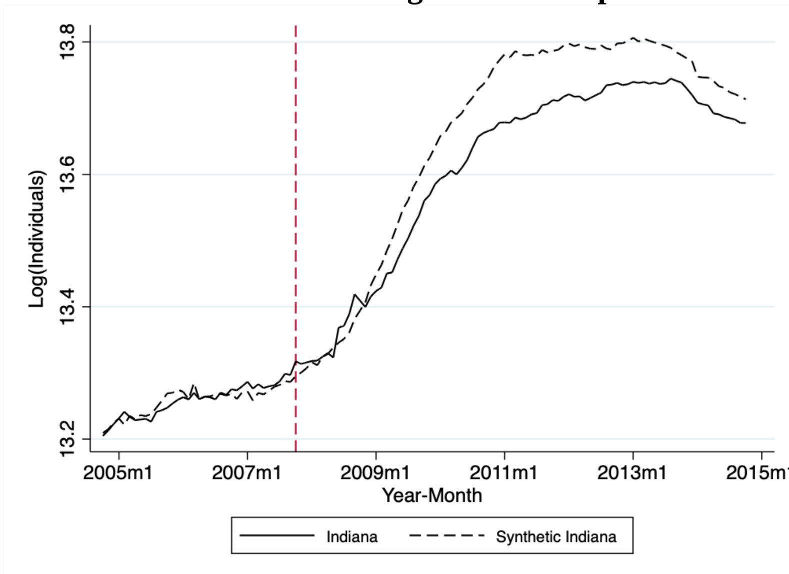
Figure 1.7 shows the change in log SNAP cases over time for Indiana and for the "synthetic control" group. Four states collectively constitute 99% of the weight given to "synthetic Indiana": Ohio (29%), Utah (26%), Wisconsin (23%), and Kentucky (21%).[70] Notably, three of these states are in the same geographic region as Indiana, even though geographic proximity did not explicitly

[68] This state-level analysis is carried out only for SNAP, as monthly state-level administrative TANF records are of questionable quality and monthly state-level administrative Medicaid records are not easily obtainable for earlier years from a single national source.

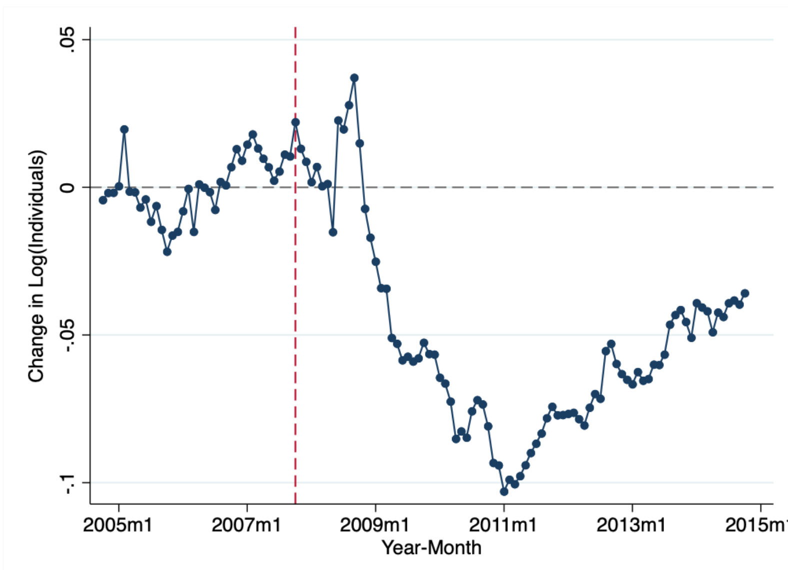
[69] Specifically, the weights are constructed by matching on the following state- and time-varying covariates: the number of SNAP recipients for each of the 12 quarters preceding treatment, the numbers of males and females (logged), the numbers of white, black, and other race individuals (logged), the numbers of individuals aged 0-4, 5-17, 18-24, 25-44, 45-64, and 65 plus (logged), median income (logged), unemployment rate, and a policy index taking values between 0 and 3 for the number of other selected policy changes (simplified reporting, broad-based categorical eligibility, and the exclusion of vehicles from the asset test) adopted by a state during a particular month.

[70] Alabama, Mississippi, Louisiana, and Florida are omitted from the donor pool of states for the synthetic control group, as they experienced abnormally large increases in SNAP enrollment in 2005 as a result of Hurricane Katrina. Given that at least one of these states received non-trivial weights via the synthetic control method when included, including them in the donor pool led to non-parallel trends in enrollment prior to treatment.

Figure 1.7. Synthetic Control Estimates for Log SNAP Receipt



(a) Indiana vs. Synthetic Indiana Trends



(b) Difference in Trends

Data Sources (public-use): USDA SNAP state-level enrollment data (2004-2015), BLS Unemployment Statistics (2004-2015), Census population and income estimates (2004-2015), USDA SNAP Policy Database (2004-2015)
Notes: These figures use the synthetic control method to compare changes in the number of individuals receiving SNAP in Indiana to changes in other states. Panel (a) compares the monthly trends in log SNAP recipients between Indiana and synthetic Indiana, with the synthetic control group constructed based on matching to Indiana on the following set of pre-treatment covariates: SNAP recipients for each of the 12 quarters preceding treatment, population cuts by gender, race, and age, median income, unemployment rate, and the number of SNAP policies adopted by a state during a particular month. Panel (b) plots the monthly differences in log SNAP recipients between Indiana and synthetic Indiana.

factor into the matching process. Panel (a) compares the trends in log SNAP enrollment between Indiana and synthetic Indiana, showing that the levels and trends in enrollment are very similar prior to October 2007. However, the trends diverge after Indiana received IBM's automated system, with enrollment growing much faster in synthetic Indiana as a result of the Great Recession. Panel (b) plots the difference in log enrollment between Indiana and synthetic Indiana, showing a gap of approximately 10% at the height of the enrollment decline that only incompletely bounces back in subsequent years. One should expect the synthetic control estimates to be somewhat smaller in magnitude than the county-level difference-in-differences estimates, since the former estimates are based on state-level changes in SNAP enrollment and only a fraction of Indiana's population is treated. These cross-state estimates consequently provide a useful benchmark that further validates the cross-county difference-in-differences estimates.

1.8 Conclusion

In 2006, Indiana awarded a 10-year, $1.3 billion contract to the IBM Corporation to manage the automation of the state's welfare services. Expected to lower administrative costs and increase convenience for clients and operators alike, IBM's automated system began rolling out to counties in late 2007. However, due to performance problems, the system was terminated two years later after reaching only 59 out of Indiana's 92 counties. Leveraging the natural experiment in this setting that distinguishes treated counties (receiving the rollout) to untreated counties (not receiving the rollout), this paper develops a framework to compare the effects of barriers to enrollment on take-up and targeting among initial applicants and recertifiers. This paper concludes that IBM's automated system created a number of burdens associated with application and recertification, leading to sharp declines in SNAP, TANF, and Medicaid enrollment in the treated counties.

Using administrative welfare records spanning nearly 3 million recipients, this paper finds statistically significant declines in TANF (24%), SNAP (15%), and Medicaid (4%) enrollments one year after the rollout of IBM's automated system. These effects are largely due to insufficient personalized assistance from caseworkers, a zero-tolerance policy for application errors, and staggering delays and technical glitches at overwhelmed call centers. Disaggregating the overall enrollment effects across application stages, this paper finds decreases in entry rates for all three programs that are similar in magnitude and increases in exit rates that are largest for TANF and smallest (and statistically insignificant) for Medicaid. The larger effects on exit rates - and thus overall enrollment - for TANF can be attributed to higher transaction costs stemming from shorter TANF recertification intervals and fewer avenues for assistance (vice-versa for Medicaid).

Linking these program participation records to IRS microdata and other administrative sources enables an analysis of heterogeneous effects on different types of counties and individuals with different characteristics. At the county level, enrollment reductions are largest in earlier-treated and higher-poverty counties, as well as in lower-unemployment counties whose residents may have more earnings to verify. This paper also finds that IBM's automated system overall screens out individuals who appear to be less needy, as those remaining on the program rolls typically have lower pre-treatment incomes, fewer years of education, higher per-person benefits, and higher disability levels. Yet, these overall effects conceal striking and novel differences across application stages. IBM’s automated system appears to be an efficient screen at initial application (screening out entrants who are less needy) but an inefficient screen at recertification (severing benefits to those who are more needy). These differences may be due either to variation in the nature of administrative burdens across stages or to differential selection of individuals into each application stage.

The natural experiment setting and groundbreaking data sources in this paper can be further used to examine the downstream outcomes and well-being of those individuals cut off from welfare programs. This includes examining whether or not welfare cuts induce individuals to increase their short-run labor supply and potentially induce greater self-sufficiency in the long run. One can also examine the extent to which individuals turn to other forms of insurance as a result of sharp negative income shocks, ranging from other government programs to informal transfers from other relatives or non-relatives. Changes in earnings and other transfer income can serve as mediators in helping to understand the broader effects of welfare cuts on financial solvency, health, and education. These analyses can provide more direct evidence on the mechanisms underlying the long-term effects on SNAP and Medicaid enrollment - identified in this paper - that persist beyond the life of IBM's automated system.

States and localities are increasingly adopting features of automated welfare services, a trend that has accelerated during the COVID-19 pandemic. While automation is often thought to make the enrollment and recertification processes more convenient, it may also induce complexities – when imperfectly administered – that are less well understood. Many features of IBM's automated system continue to persist today in other contexts, potentially precluding eligible individuals from accessing benefits and undermining the effectiveness of policies that expand eligibility or benefit amounts. For example, during the early stages of the COVID-19 pandemic, Unemployment Insurance applicants faced staggering delays in receiving their benefits as a result of overwhelmed field offices and call centers.

As more governments engage in policy experimentation, it is important to understand the combination of intended and unintended consequences that may result from such efforts (e.g., Callander & Harstad 2015). This paper shows that "one size fits all policies" may not be optimal when applied across different stages of the program certification process. Furthermore, a policy that introduces a common set of administrative burdens can still differentially affect programs if they differ in their baseline transaction costs (e.g., length of recertification interval). Policymakers would also be well-advised to consider the potentially long-term adverse effects of short-lived burdens.

CHAPTER 2:
Does Geographically Adjusting Poverty Thresholds Improve Poverty Measurement and Program Targeting?

2.1 Introduction

Prices, especially housing costs, differ greatly across geography. This pattern implies that a fixed market basket of goods costs more in some places than in others. For example, in FY 2021, the median rental price for a 2-bedroom apartment in New York City is $2,263, more than three times as large as the median rent for a 2-bedroom apartment in rural Mississippi ($684).[71] These differences raise the question as to whether poverty thresholds should be adjusted to reflect geographic differences in the cost of living. Perhaps more importantly, should payments under government transfer programs also reflect these geographic differences? Answers to these questions have enormous policy implications. In classifying fewer people as poor in lower-cost areas and more people as poor in higher-cost areas, geographic adjustments to poverty thresholds would dramatically change how researchers characterize poverty and how policymakers allocate anti-poverty efforts.

Much of the recent policy literature presumes that geographic adjustments are justified.[72] The arguments for such adjustments are often based on the wide geographical variation in prices, in particular housing costs. In its "Measuring Poverty" report, Citro and Michael (1995) recommended a number of changes to the Official Poverty Measure, including that "poverty

[71] These amounts are derived from the Department of Housing and Urban Development's estimates of the 50th percentile of rent by geographic location. See https://www.huduser.gov/portal/datasets/50per.html#2021.
[72] On more than one occasion, members of Congress – including Rep. Jim McDermott (D-WA) and Sen. Chris Dodd (D-CT) in 2009 and Rep. Alexandria Ocasio-Cortez (D-NY) in 2019 – have introduced bills proposing that the official poverty line in the U.S. be adjusted for geographic variation in cost-of-living.

thresholds should be adjusted for differences in the cost of housing across geographic areas of the country" (p. 183). The Census Bureau's Supplemental Poverty Measure (SPM) now adjusts its thresholds for geographic differences in housing prices (Fox 2019). Canada uses low-income cutoffs that vary geographically by city size and urban or rural residence (Baker et al. 2019).[73] Moreover, current federal policy frequently differentiates between geographic areas in both benefit levels and reimbursement rates. For example, maximum benefit amounts for the Supplemental Nutrition Assistance Program (SNAP) and reimbursement rates for free and reduced-price school meals are higher in Alaska and Hawaii than in the 48 continental states. Eligibility and benefits for federal housing assistance are determined based on incomes and fair market rents varying across metropolitan areas and non-metropolitan counties.

Yet, researchers remain divided as to whether geographic adjustments are conceptually desirable. A long literature in economics suggests that spatial differences in prices reflect differences in what purchases obtain for those prices. In one of the most commonly used spatial equilibrium frameworks, consumers are modeled as willing to pay more in certain areas to consume higher-quality amenities (Rosen 1974, Haurin 1980, Roback 1982). Relatedly, Tiebout (1956) hypothesizes that individuals sort themselves across geographies according to their preferences for public goods. Empirically, this sorting leads to increased housing prices in communities with higher levels of public goods provision (Oates 1969, Epple 2008, Brueckner 2011). As a result, the variation in housing prices across locations – which the SPM relies upon to geographically adjust poverty thresholds – may simply reflect variation in locational desirability. Furthermore, an earlier set of government reports highlighted conceptual and data limitations of geographic adjustments and noted that such adjustments probably would not reflect other regional

[73] See also Statistics Canada's page on low-income cutoffs:
https://www150.statcan.gc.ca/n1/pub/75f0002m/2012002/lico-sfr-eng.htm

differences such as the level of assistance to low-income families (U.S. Department of Health, Education, and Welfare 1976, General Accounting Office 1995).

In this paper, we shed light on the consequences of geographic adjustments to poverty measures and other measures of well-being. We do so by examining whether geographic adjustments currently in use – and others that have been proposed – help to achieve what several key studies have deemed to be the central goal of a poverty measure: identifying the most disadvantaged population (see, e.g., Ruggles 1990, Citro and Michael 1995).[74] A wide variety of programs determine benefit eligibility based upon either a poverty cutoff or some multiple of a poverty line, constructed using a resource measure that is conceptually similar to that of the Official Poverty Measure (OPM).[75] Consequently, a poverty measure that identifies the most disadvantaged can help to target government transfers to those who are most needy. This goal is also consistent with how researchers and the broader public often think about poverty measures, which are used as indicators of disadvantage and predictors of various negative outcomes. Even if one conceives of being in poverty has having income below some minimum standard, there are statistical difficulties associated with accurately measuring income. Thus, utilizing other indicators of material well-being that are correlated with true income can aid in measuring the truth.

We use two Census surveys for our analyses: the Current Population Survey's Annual Social and Economic Supplement (CPS), the source of official poverty estimates, and the Survey of Income and Program Participation (SIPP), thought to have the most accurate U.S. income information. In each survey, we compute both the SPM and the new Comprehensive Income

[74] For example, the NAS Panel in *Measuring Poverty* sought to produce a "measure that will more accurately identify the poor population today" (p. 1) and went on to define poverty as "material deprivation." (p. 19).
[75] These programs include the Supplemental Nutrition Assistance Program (SNAP), free and reduced price school meals, Head Start, and the Low Income Home Energy Assistance Program (LIHEAP). For a full list of federal programs that use poverty guidelines in determining eligibility, see https://www.irp.wisc.edu/resources/what-are-poverty-thresholds-and-poverty-guidelines/.

Poverty Measure (CIPM) – the latter of which measures incomes more accurately using linked survey and administrative data from the Comprehensive Income Dataset (CID) – with and without a geographic adjustment.[76] To compare different poverty measures on an equal footing, we keep the share of individuals in poverty under alternative measures the same, proportionately adjusting poverty cutoffs as needed. We analyze a wide variety of measures of well-being from survey and administrative sources, including survey reports of material hardships, appliances owned, home quality issues, food security, public services, health, education, and assets, as well as permanent income and mortality from administrative records. In total, we examine 71 well-being indicator spanning these ten domains.

For each of the well-being indicators, we compare those who are poor under the non-geographically-adjusted SPM or CIPM but not the geographically-adjusted versions (i.e., the "non-geographic-only poor") to those who are poor under the geographically-adjusted SPM or CIPM but not the non-geographically-adjusted measures (i.e., the "geographic-only poor"). These are the only two groups in a cross-classification of the two poverty definitions that matter for the comparison of the two approaches, as those classified as poor (or non-poor) by both measures do not enter the comparison. More broadly, this strategy builds upon previous work showing that one can assess the desirability of a poverty adjustment by examining the extent of deprivation among those classified as poor with and without an adjustment (see Meyer and Sullivan 2003, 2011, 2012a, Renwick 2018, Fox and Warren 2018, Renwick 2019, Meyer, Wu, Mooers et al. 2021).

For nine of the ten domains of well-being indicators that are available, we find that the majority of outcomes point to the geographic-only poor under the SPM or CIPM being more deprived than the non-geographic-only poor under either poverty measure. Among eight of these

[76] While we focus in this paper on the SPM and CIPM, our results are likely to be more general, as they can be thought of as comparing levels of well-being between high- and low-cost areas, holding measures of nominal income constant.

nine domains, at least two measures indicate that geographic adjustments statistically significantly identify a less deprived poor population. These patterns hold after a variety of extensions and robustness checks, including partial geographic adjustments, using Regional Price Parities (which cover a broader bundle of goods) as the geographic adjustment index, focusing on those switched in and out of deep and near poverty by a geographic adjustment, and using the OPM rather than the SPM or CIPM as the poverty measure.

Our results can be rationalized by the empirical fact that prices at the state or sub-state level are strongly associated with many characteristics that are important to those with low incomes. Wages have been found to rise almost one for one with prices (DuMond et al. 1999, Hirsch 2011), and we confirm this result in the CPS. Many other characteristics differ across local areas and have been shown to be reflected in home prices or rents. These include public goods such as schools (Tiebout 1956, Oates 1969, Black 1999, Epple 2008), pollution (Davis 2004, Chay and Greenstone 2005), and cash welfare (Glaeser 1998). Many categories of state and local spending are strongly associated with prices. We find that the elasticity of spending with respect to prices exceeds one for state and local expenditures on welfare, elementary and secondary education, environment and housing, and police. These characteristics have the potential to offset some or all of the increases in required resources needed to maintain a given standard of living in the face of higher prices.

The remainder of the paper is structured as follows. Section 2.2 presents a simple theoretical model of how local poverty thresholds should change with local prices and other characteristics, reviews the theoretical literature on the desirability of geographic adjustments, and discusses empirical methods previously used to geographically adjust poverty thresholds. Section 2.3 describes the survey and administrative data used and explains how we calculate poverty measures with and without a geographic adjustment. Section 2.4 presents descriptive statistics showing how poverty rates change across characteristics after incorporating a geographic adjustment. Section 2.5 discusses our measures of material well-being as well as the analytical methods used to compare well-being across poverty methods. Section 2.6 contains the main regression results comparing measures of material well-being among those moved in and out of poverty by a geographic adjustment. Section 2.7 discusses several extensions and robustness checks, and Section 2.8 provides some empirical explanations for our main results. Section 2.9 concludes.

2.2 Theory and Previous Literature

2.2.1 Simple Model of Local Poverty Thresholds and Prices

We begin with a simple formal model of how local poverty thresholds should change as local prices change, based on Glaeser (2011). Assume that households have a well-defined indirect utility function $V(Y, P, \mathbf{A})$, where Y is income, P is prices, and $\mathbf{A}$ is a vector of amenities. Assume that income and amenities enter positively into indirect utility ($V_Y \geq 0$, $V_{A_j} \geq 0 \; \forall A_j \in \mathbf{A}$), while prices enter negatively ($V_P \leq 0$). For simplicity, let us assume in this setting that P is a scalar (e.g., P may be the price for a single good like housing or a composite index of prices for a basket of goods). Following Glaeser (2011), we define a household as poor if its value of $V(Y, P, \mathbf{A})$ is below that of some minimum deprivation level $\underline{V}$. Let Y_0^* designate the income level in location 0 such that $V(Y_0^*, P_0, \mathbf{A}_0) = \underline{V}$, where P_0 and $\mathbf{A}_0$ are the price and amenity levels in location 0, respectively. In other words, Y_0^* can be thought of as the "poverty threshold" in location 0 – i.e., for a given set of prices and amenities in location 0, incomes below Y_0^* will lead to deprivation levels below $\underline{V}$ and incomes above Y_0^* will lead to deprivation levels above $\underline{V}$.

Using this setup, we analyze how the appropriate poverty threshold changes as other determinants of utility change. Before proceeding, we make an additional modification to the framework in Glaeser (2011). Namely, we include hourly wages *w* and unmeasured income *N* as additional inputs into indirect utility as they are key determinants of utility that vary geographically. *N* can include income that is often under-reported in various data sources, such as housing assistance, child support, and workers' compensation. It can also include sources ranging from medical in-kind transfers like Medicaid to uncompensated hospital care and food pantries, which are typically omitted from standard income measures. Hourly wages and unmeasured income enter positively into indirect utility ($V_w \geq 0$, $V_N \geq 0$). We can then totally differentiate $V(Y^*, P, \mathbf{A}, w, N) = \underline{V}$ to obtain:

$$V_Y dY^* + V_P dP + \sum_j V_{A_j} dA_j + V_w dw + V_N dN = 0. \quad (2.1)$$

Rearranging terms and dividing by *dP* yields:

$$\frac{dY^*}{dP} = -\frac{V_P}{V_Y} - \sum_j \frac{V_{A_j}}{V_Y}\frac{dA_j}{dP} - \frac{V_w}{V_Y}\frac{dw}{dP} - \frac{V_N}{V_Y}\frac{dN}{dP}, \quad (2.2)$$

which gives an expression for how the appropriate poverty threshold changes as prices change. Using Roy's Identity, we can rewrite equation (2.2) as follows:

$$\frac{dY^*}{dP} = \underbrace{X}_{\geq 0} - \underbrace{\sum_j \frac{V_{A_j}}{V_Y}\frac{dA_j}{dP}}_{\geq 0} - h\underbrace{\frac{dw}{dP}}_{\geq 0} - \underbrace{\frac{V_N}{V_Y}\frac{dN}{dP}}_{\geq 0}, \tag{2.3}$$

where X is some base consumption level and h is some base level of hours worked.

Consider first the unrealistic case in which amenities, hourly wages, and non-labor income are uncorrelated with prices. Under these assumptions, the expression on the right side of equation (2.3) boils down to X. This makes intuitive sense – in the absence of all other terms, higher prices must lead to a higher poverty threshold in order to maintain the same level of consumption. However, the presence of the amenity, wage, and unmeasured income terms will tend to counteract and potentially reverse the naïve price correction. We know from an abundance of evidence that amenities such as school quality and clean air tend to be positively correlated with prices. We also empirically find that areas with higher prices have higher hourly wages as well as higher levels of non-labor income sources. Thus, if da/dP, dw/dP, and dN/dP are sufficiently positive, then the appropriate poverty threshold Y^* could be decreasing in prices when we allow other factors like amenities, hourly wages, and unmeasured income (which enter positively into utility) to change with prices as well. Importantly, note that we are not assuming that the spatial equilibrium assumption holds for those near the poverty line (in such a case, increased amenities and hourly wages would exactly offset higher nominal prices). However, it is likely to be the case that local amenities and incomes are strongly correlated with local prices for those in poverty.

2.2.2. Literature on Prices and Amenities Across Geography

A long literature going back nearly seventy years discusses how geographic differences in prices reflect local characteristics. Tiebout (1956) famously argued that a household – under a set of assumptions including costless migration across areas – will sort into a community providing public good levels most closely aligned to its preferences. Oates (1969) built upon Tiebout's argument by reasoning that this type of sorting will result in higher housing prices in areas with higher levels of public goods provision. In arguably the classic approach to spatial equilibrium, Rosen (1974), Haurin (1980) and Roback (1982) demonstrated that wages and rents in equilibrium must adjust so that workers – who are assumed to care about amenities in addition to wages and cost-of-living – are indifferent between living in areas with differing amenity levels. Specifically, in a location with higher amenities, Roback inferred that consumers' willingness to pay for those amenities can be obtained by taking the sum of the higher housing costs and the lower wages in that location. While more recent work (see, e.g., Roback 1988, Gyourko and Tracy 1991, Moretti 2011) loosens some of the assumptions in Roback's initial model (allowing labor to be heterogeneous rather than homogeneous, accommodating non-tradable consumption goods in addition to tradable goods, etc.), Roback's original finding remains largely intact: nominal wages and prices adjust to take into account differences in amenities across localities.

In considering the applicability of this approach to the poor, one should recognize that the adjustment of prices to local amenities might be largely determined by the amenities for the larger group of non-poor individuals. For example, the adjustment of prices may not reflect wages in the low-wage labor market or the availability of support through welfare programs available to the least well off. In such a case, one cannot expect price adjustments to make all geographic areas equally desirable to the poor. Migration costs or restrictions on wage adjustments such as minimum wage laws could have a similar effect.

To develop an optimal adjustment for differences across localities using the Rosen-Haurin-Roback framework, one would need to account for wages and amenities as well as prices. But even if an approximate equilibrium does not occur, that framework clarifies the information needed to optimally adjust for geographic differences in prices – namely, all characteristics of geographic areas that affect the desirability of those areas and how they are valued by low-resource families. This requirement is demanding, if not unattainable. One would have to estimate the value of all relevant amenities, which then could be used to construct a price index (Blomquist et al. 1988, Gyourko and Tracy 1991). Empirical implementation has found overwhelming evidence of the presence of amenities, but great difficulty in pinning down their values.[77] Greenstone (2017) argues that the omitted variable problem is so overwhelming that reliable estimates of the marginal value of certain amenities can only be obtained under special circumstances. Furthermore, even if one can obtain the valuation for the marginal person, it is the average valuation that is desired for a price adjustment. Estimating the average is an even harder task (Kaplow 1995, Greenstone 2017). In summary, the canonical economic model of prices and amenities suggests that amenity values are needed to make an appropriate geographic price adjustment, but attempts to empirically implement the model indicate that the current data are inadequate to estimate these values.

Closely related to the geographic adjustment of poverty thresholds, Kaplow (1995) and Glaeser (1998) examine the circumstances under which equity and efficiency are improved by tax and transfer payments that differ across geography. Besides the static changes in well-being from such adjustments, the authors also consider how such geographic differences would be affected by the possible migration of those with few resources. Kaplow (1995) undertakes a conceptual investigation of spatial cost-of-living adjustments in the tax and transfer system. In a benchmark case that leads to equal utility between regions, Kaplow reasons that it would be efficient and equitable to adjust transfers for cost-of-living differences. However, there are certain factors that

[77] As Glaeser (2011) succinctly summarizes, "unobserved amenity differences bedevil local price measurement."

suggest that cost-of-living adjustments may be undesirable. Principally, when differences in nominal cost-of-living are systematically correlated with differences in amenities across regions, making adjustments using standard price indices may be counterproductive. He also points out that increasing transfers to low-cost areas would reduce government costs if it induces a pattern of migration from high-cost regions to low-cost regions.

Similar to Kaplow (1995), Glaeser (1998) shows that the indexing of transfer payments to local price levels might increase social welfare under the following assumptions: amenities are complements (rather than substitutes) with income, prices are not being offset by higher wages, higher transfer levels do not induce greater mobility to the high transfer areas, and individuals are risk averse. Using his model, Glaeser performs a calibration exercise to calculate the optimal amount of indexing under various parameter values corresponding to his assumptions that are relevant for the transfer recipient population.[78] Under these parameters, Glaeser finds that a one percent increase in prices should optimally lead to a 0.33 percent increase in transfers relative to total income. Notably, under every combination of parameter estimates used to calibrate the model, he finds that this elasticity should never optimally exceed one. In contrast, using data on AFDC (Aid to Families with Dependent Children, the precursor to Temporary Assistance for Needy Families or TANF), Glaeser finds evidence that the elasticity of transfer payments with respect to local prices exceeds 1.5. He uses this result to conclude that the current level of indexing by geography appears to be too strong to be optimal.[79]

[78] His preferred estimate assumes that amenities and income are independent (i.e., neither complements nor substitutes), wages are fixed, the elasticity of migration with respect to income is 1, and the coefficient of relative risk aversion is equal to 2 (with the literature suggesting estimates that range from 1 to 10).
[79] Whereas Glaeser's analysis thinks about an adjustment that determines all payments (including local payments), our paper focuses on measurement and marginal changes to targeting by the federal government.

2.2.3 Geographic Price Indices

A number of different price indices have been proposed to adjust poverty thresholds by geography. The Supplemental Poverty Measure (SPM) relies on the Median Rent Index (MRI) to geographically adjust its thresholds (Fox 2019). Using information from the 5-year American Community Survey (ACS) files, the MRI for a given geographic area is calculated as the ratio of its median gross rent for a two-bedroom unit with a complete kitchen and bathroom to the median gross rent for the same type of unit in the U.S. Closely related to the MRI is a rescaled version of the MRI proposed by Renwick (2018, 2019) that seeks to reflect amenities in geographic adjustments of poverty rates. Renwick argues that the MRI will over-adjust thresholds if places with higher median rents also have greater amenities (and vice-versa). In order to adjust for amenities in a rough way, Renwick (2018) cuts the variation in the MRI index in half. This rescaled index is admittedly an “arbitrary” adjustment because the literature has established no clear methodology for incorporating amenities (p. 5).

Another proposed method for adding geographic adjustments to poverty is to use Regional Price Parities or RPPs (Aten 2005, Aten and D’Souza 2008). Calculated by the Bureau of Economic Analysis (BEA), RPPs are spatial price indices that measure price differences in a broad set of two hundred individual items comprising eight broad categories: housing, transportation, food, education, recreation, medical, apparel, and other. Thus, RPPs rely on a wider set of goods and services than the MRI, which only accounts for differences in rents. A final method for incorporating geographic adjustments to poverty is to use food, apparel, and rent regional price parities or FAR RPPs (Renwick et al. 2014, Renwick et al. 2017)., FAR RPPs cover only the subset of goods in the RPPs that are also included in the SPM poverty threshold – namely food, apparel, and rent. The appendix contains additional details about each of these specific indices.

2.2.4 Evidence on Association with Deprivation

Several papers have analyzed the desirability of a geographic adjustment by comparing the material deprivation of those classified as poor with and without an adjustment. In a supplementary analysis to their main results, Meyer and Sullivan (2012a, see online appendix) analyze the impact of adjusting thresholds for geographic variation in prices on the characteristics of the SPM poor in the CPS. They find that geographically adjusting the thresholds leads to a poor population that is more likely to be covered by private health insurance and has higher levels of education, but the statistical significance of these changes is not examined.

Renwick (2018, 2019) investigates the correlation between poverty rates calculated using income adjusted various ways – using the MRI, RPP, FAR RPP, and rescaled MRI – and measures of material deprivation. Specifically, using 51 observations (50 states plus the District of Columbia), she analyzes the correlation between state poverty rates (averaged across the 2015-2017 reference years using the CPS and adjusted using each of the methods above) with state-level indicators of economic well-being. She focuses on the multi-dimensional deprivation index (MDDI), a state-level measure of deprivation developed by the Census Bureau that combines measures of several components of well-being – including health, poverty, education, economic security, housing, and neighborhood quality (Glassman 2019). Five of the six components are estimated from responses to the ACS, while neighborhood quality is measured using the County Health Rankings and Roadmaps dataset.

Renwick finds that three of the geographically adjusted poverty measures are less correlated with the multi-dimensional deprivation index than the measure that does not adjust for cost of living, although tests of significance are not reported for all comparisons. A fourth geographic adjustment is more highly correlated with the well-being indicator than the unadjusted SPM. The winning adjustment is the rescaled MRI based on an ad hoc multiplication of the SPM price adjustment by one-half. When examining the components of the multidimensional measure, most are more correlated with the unadjusted SPM measure than with any of the geographically adjusted measures, although not all comparisons are tested. The well-being indicators are broader and the statistical tests stronger than in Meyer and Sullivan (2012a). However, the analysis uses state averages rather than individual data. Thus, it is informative about state-level differences in poverty, but it may be less informative for differences by characteristics like family type, race, age, or other geographic levels. The approach also does not keep the poverty rate the same across poverty measures so the measures are not completely comparable, although examining correlations should reduce or eliminate the impact of this non-comparability. Finally, the analysis does not hold constant demographic differences across the states which might confound the comparisons.

In another paper, Baker et al. (2019) provide empirical evidence on the relationship between poverty and mortality in Canada. They analyze Canadian poverty using the Canadian low-income cutoff (LICO). While the official version of the LICO uses a geographic adjustment based on the size of the city in which a person lives, they also construct a fixed-cutoff LICO that does not have geographic adjustments. When the authors analyze the relationship between the fixed-cutoff LICO and mortality, they find that areas with more poor people have higher rates of mortality. However, when analyzing the relationship between the official LICO and mortality, they find that, among some age groups, areas with more poor people have lower mortality rates. This counterintuitive result suggests that the geographic adjustment to Canada’s LICO does not identify those who are the most deprived.

2.3 Data and Poverty Measures

In this section, we describe the survey and administrative data used in this paper and the methods we use to link the data sources. We then discuss the features of our two core poverty measures – the Supplemental Poverty Measure (SPM) and the Comprehensive Income Poverty Measure (CIPM). In discussing each poverty measure, we focus on defining its key ingredients – namely, the resource measure, the resource-sharing unit, the poverty threshold, and the equivalence scale used to set poverty thresholds for families that differ in size or composition. We focus on reference year 2010, since this is a year for which we have a relatively complete set of administrative records covering all income sources.

2.3.1 Data

Survey Data

Our survey data pertain to calendar year 2010 and come from the 2011 Current Population Survey Annual Social and Economic Supplement (CPS) and the 2008 Panel of the Survey of Income and Program Participation (SIPP). Both surveys are designed to be representative of the civilian non-institutional population of the United States. The 2011 CPS interviewed 75,000 households between February and April of 2011 about their incomes in calendar year 2010. The 2008 SIPP was a longitudinal survey that followed 42,000 households for up to 16 four-month waves, though not all households were observed for all 16 waves due to survey attrition.

While the default reference period for the CPS is a calendar year, the reference period in the SIPP is four consecutive months. We therefore combine information across multiple interview waves in the SIPP to calculate annual incomes. Specifically, we take as our analysis sample all individuals who appear in reference month 4 of Wave 6 (which spans April-July 2010), incorporate information on survey incomes from other months in 2010 during which the individuals in the analysis sample appear, and proportionately scale up survey incomes for the 21% of individuals who are interviewed for only a portion of the year. This decision not only makes the CPS and SIPP income measures more comparable, but it also aligns the SIPP reference period with that of the linked calendar year tax data. In addition to collecting monthly data on a rich set of income sources, the SIPP collects measures of material well-being, certain expenses, and household structure in topical modules administered in the final month of various interview waves.

Administrative Data

We also employ a number of administrative data sources. We obtain earnings records from multiple sources of tax records (including Internal Revenue Service (IRS) W-2 Forms, the Detailed Earnings Record (DER) database of the Social Security Administration (SSA), and IRS 1040 Forms), asset income (namely interest and dividends) from IRS 1040 Forms, and retirement distributions from IRS 1099-R forms. We also simulate tax liabilities and credits from line items available on IRS 1040 Forms. We have a number of administrative program participation records from government agencies, covering Social Security, Supplemental Security Income (SSI), Service-Connected Disability payments to veterans, and HUD housing assistance. For this paper, we do not bring in administrative data from state agencies on SNAP or TANF, because these program data are only available for a subset of states and we want our analysis sample to cover every state in the nation. We also use administrative records to construct additional measures of well-being, including using IRS tax records from tax years 2008, 2009, 2011, and 2012 to construct a measure of permanent income and the Social Security Administration's Numident file to calculate mortality rates. The appendix contains additional details about the survey and administrative data sources.

Linking Data Sources

We link the survey and administrative data sources using individual identifiers called Protected Identification Keys (PIKs). PIKs are created by the U.S. Census Bureau's Person Identification Validation System (PVS), which is based on a reference file containing Social Security Numbers linked to names, addresses, and dates of birth (Wagner and Layne 2014). Our survey-based analyses use the full CPS or SIPP sample and original survey weights. For most of our analyses that analyze outcomes from the administrative data or that use CID income as the income base, we restrict our sample to individuals whose sharing units have at least one member with a PIK (and, in the CPS, no member that is whole imputed).[80] To account for the bias arising from non-random missing PIKs (and whole imputations in the CPS), we divide survey weights by the predicted probability that at least one member of the sharing unit has a PIK (and no member is whole imputed in the CPS), conditional on observable characteristics in the survey.

2.3.2 Poverty Measures

Supplemental Poverty Measure (SPM)

The Supplemental Poverty Measure (SPM) differs from the Official Poverty Measure (OPM) in several ways. Unlike the OPM, which uses pre-tax money income as its resource measure, the SPM resource measure covers a fuller set of resources available for consumption –

[80] In the specific case of analyzing patterns in individual-level mortality (from the SSA Numident) while relying on survey income, we restrict our sample to individuals who link to a PIK and adjust for non-PIKing at the individual level.

namely, pre-tax money income plus non-cash transfers net of certain expenses and taxes. Specifically, the SPM resource measure adds to pre-tax money income the estimated value of benefits received through SNAP, housing assistance, school meals, the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC), and the Low Income Home Energy Assistance Program (LIHEAP). It then subtracts federal and state income tax liabilities net of credits and payroll taxes. Finally, it subtracts estimated expenses for work, childcare, child support, and health care. We calculate the SPM resource measure using survey information only, following the methodologies in Fox (2019) and Short (2014) for the CPS and SIPP, respectively.

Next, while the OPM uses a family (defined as all people living together related by birth, marriage, or adoption) as its resource unit, the SPM resource unit additionally includes cohabiting partners, unrelated children under the age of 15, and foster children between the ages of 15 and 22. The SPM also uses different poverty thresholds than the OPM does and adjusts the thresholds in a different way. While the OPM thresholds reflect only economies of scale in food and do not adjust for geographic price differences, the SPM thresholds are based on out-of-pocket spending on a broader set of goods and adjust for geographic differences in rental prices. Formally, the threshold for an SPM unit can be written as the product of the following terms:

$$SPM\ Threshold_{t,ac,sm} = (Base\ Threshold)_t \times \frac{(Equivalence\ Scale\ Factor)_{ac}}{E} \times [(Housing\ Share_t \times MRI_{sm}) + (1 - Housing\ Share_t)], \quad (2.4)$$

where t is the unit's type of housing tenure, a and c represent the number of adults and children in the unit, and s and m denote the unit's state and metropolitan statistical area (MSA).

The first term in equation (2.4) is the base threshold, which is calculated as 1.2 times the average spending on food, shelter, clothing, and utilities of those in the 30th-36th percentiles of spending on these expenses, computed using five years of Consumer Expenditure (CE) Survey data. However, the SPM calculates spending separately for three housing tenure groups using the same percentiles: homeowners with mortgages, homeowners without mortgages, and renters (implicitly assuming they are otherwise the same).[81] The second term is a three-parameter equivalence scale that adjusts the threshold based on the number of adults and children in the reference unit, and we divide it by the equivalence scale for a two-adult, two-child unit (denoted by the constant E).[82] The final term is the geographic adjustment factor, which adjusts the threshold for geographic differences in rental prices using the Median Rent Index (MRI). The MRI is separately calculated for 358 geographic areas, including 264 publicly-identified MSAs, non-metropolitan areas in 48 states, and "other" metropolitan areas in 46 states.[83] The MRI is scaled by the share of expenditures taken up by housing costs, which again varies by housing tenure.[84] For a more detailed discussion of the methods used to construct the SPM, see the appendix.

Comprehensive Income Poverty Measure (CIPM)

For the second poverty measure that we analyze (the CIPM), we use the CID to construct an alternative resource measure that differs from the SPM resource measure conceptually and brings in administrative data to measure incomes more accurately. First, we replace survey reports

[81] In 2010, these spending amounts are $25,018 for homeowners with mortgages, $20,590 for homeowners without mortgages, and $24,391 for renters. See https://www.census.gov/prod/2011pubs/p60-241.pdf.

[82] The three-parameter equivalence scale factor is given by $(adults)^{0.5}$ for units without children, $(adults + 0.8 + 0.5 \times (children - 1))^{0.7}$ for single-parent units, and $(adults + 0.5 \times \text{children})^{0.7}$ for all other units.

[83] Not all states have observations in non-metropolitan areas or other metropolitan areas, leading to less than 50 such adjustment factors for non-metropolitan areas and for other metropolitan areas.

[84] These housing shares are 0.510 for homeowners with mortgages, 0.404 for homeowners without mortgages, and 0.497 for renters. See https://www.census.gov/prod/2011pubs/p60-241.pdf.

or imputations of asset income (namely interest and dividends), retirement income, Social Security, SSI, veterans' benefits, and tax liabilities and credits with their counterparts from the administrative data, and we combine survey and administrative sources to construct improved measures of earnings and housing assistance. Second, in line with the OPM and in contrast with the SPM, the CIPM resource measure does not subtract expenses for work, childcare, child support, and health care. While subtracting these expenses may theoretically yield a resource measure that better approximates the resources available for consumption, prior research has also shown that subtracting certain expenses (e.g., medical costs) identifies a poor population that appears less materially deprived (Meyer and Sullivan 2012a). Finally, the CIPM resource measure estimates a flow value of services from home and car ownership as well as an annuity value of other net assets. Because the CPS does not ask about assets in detail, we are only able to estimate these asset flows in the SIPP.

The CIPM uses the same resource unit as the SPM and also uses nearly the same poverty thresholds as the SPM, with the key exception being that the base threshold and housing share (in the geographic adjustment factor) in the CIPM no longer vary by housing tenure. Because the CIPM resource measure explicitly accounts for the flow value of home ownership, there is no longer a reason to set distinct thresholds (implicitly accounting for differences in available resources) for distinct housing status groups.[85]

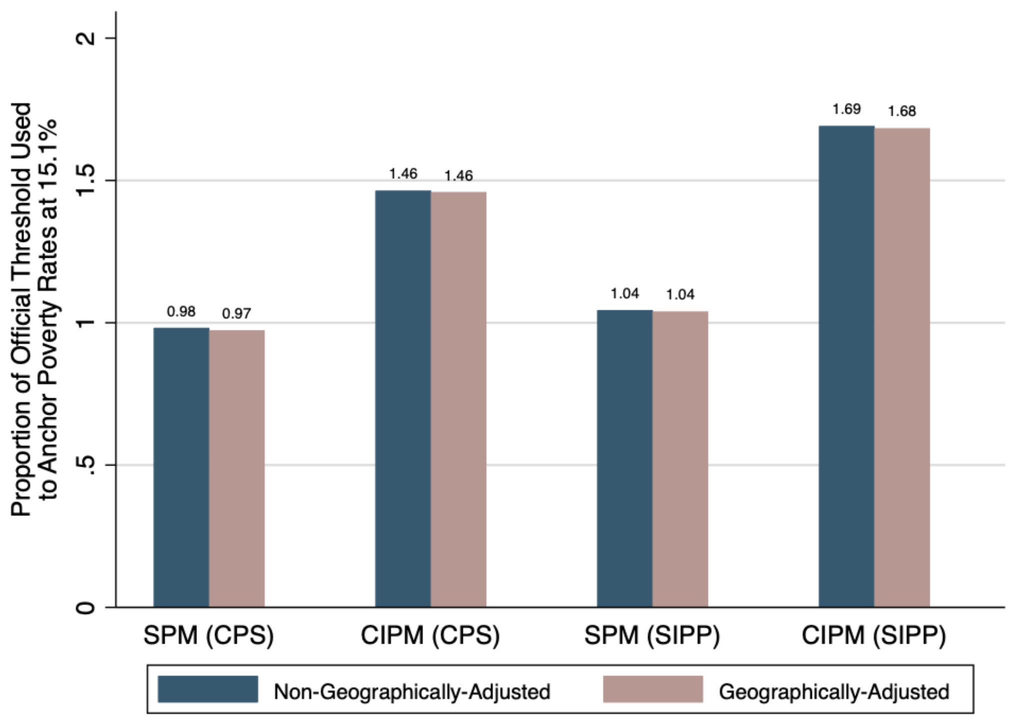
Additional Methods

To construct the SPM and CIPM without a geographic adjustment, we simply remove the geographic adjustment factor from the poverty threshold, meaning the overall threshold is now

[85] For the CIPM the housing share is set to 0.382, which is the share of overall consumption dedicated to housing in 2010 Consumer Expenditure (CE) Survey data.

just the product of the base threshold and equivalence scale. For both the SPM and CIPM (with and without a geographic adjustment), we proportionately adjust the thresholds so that the poverty rate is always fixed at 15.1%, which was the official poverty rate for 2010 found using the CPS. In other words, switching between poverty measures merely changes the “ranking” of individuals, not the absolute number of individuals in poverty. Anchoring the rates precludes us from concluding that one measure yields a more deprived population simply because it selects a smaller and thus more targeted segment of the poor.

Figure 2.1. Proportions of OPM Threshold Used to Anchor Poverty Rates at Official Levels



Data: 2011 CPS ASEC (public-use)
Notes: This figure shows the fixed proportions of the OPM (Official Poverty Measure) threshold used to adjust the SPM and CIPM thresholds in both the CPS and SIPP so that the poverty rates are always anchored at 15.1%, which was the official poverty rate in 2010. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization number CBDRB-FY21-ERD002-002.

Figure 2.1 shows the fixed proportions of the OPM threshold that are used to anchor the SPM and CIPM (in both the CPS and SIPP) at 15.1%. While the proportions applied to the SPM thresholds are between 0.97 and 1.04 (depending on the survey used and whether or not the poverty measure incorporates a geographic adjustment), the proportions are approximately 1.46 for the CIPM in the CPS and between 1.68 and 1.69 for the CIPM in the SIPP. These patterns can be explained by the CIPM bringing in administrative data to correct for underreported incomes in the survey, no longer subtracting expenses from the resource measure, and – in the case of the SIPP – adding asset flows to the resource measure.

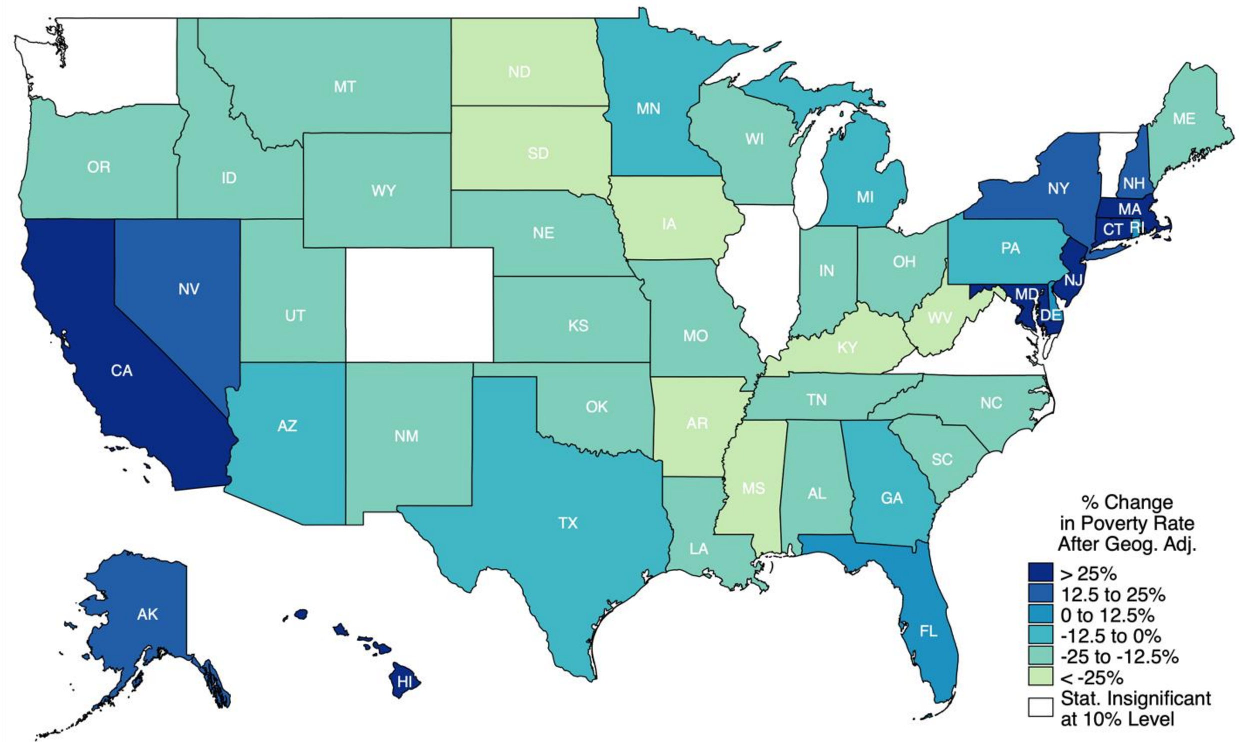
2.4 Summary Statistics

In this section, we show how poverty rates change with a geographic adjustment by state and sub-state region, race/ethnicity, and family type.

2.4.1 Changes in Poverty Rates by Geography

Figures 2.2a and 2.2b show how poverty rates change with a geographic adjustment (where the national poverty rate is always anchored to 15.1%) by state and sub-state region. Following Census Bureau standards for state-level and sub-state estimates, these differences are averaged over three years of the CPS (reference years 2009 through 2011). We focus on the SPM in the public-use CPS, as many of these estimates are based on small sample sizes that create disclosure concerns in restricted-use data. Furthermore, geographic adjustment factors can be publicly obtained for the CPS but not for the SIPP. Figure 2.2a shows state-level changes in poverty rates after applying a geographic adjustment. States with darker shading see increased poverty rates after adjusting for geographic differences in rental prices, states with lighter shading see decreased poverty rates, and states in white see statistically insignificant changes (at the 10% significance level) in poverty rates. Figure 2.2b is formatted similarly as Figure 2.2a, but it shows differences in poverty rates at the more granular CBSA (core-based statistical area) level. A CBSA is the finest level of geography that is identifiable in the public-use CPS, and it consists of one or more counties anchored by an urban center of at least 10,000 people (plus adjacent counties with high commuting ties to the urban center). We then classify counties that do not fall within a publicly identified CBSA into one of two groups within a state: “other metropolitan” or “non-metropolitan”.[86]

Figure 2.2a. Percent Change in Poverty Rates After Geographic Adjustments (State)

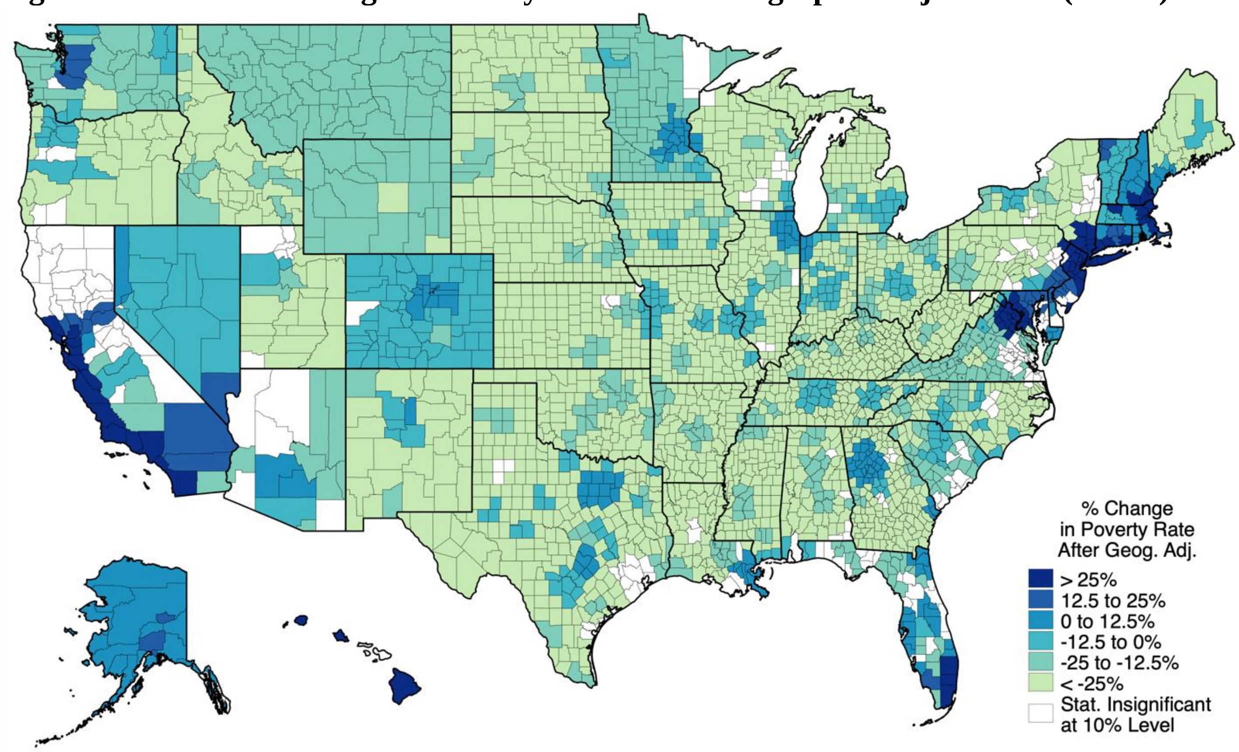


Data: 2010-2012 CPS ASEC (public-use)
Notes: This map shows the difference in poverty rates by state before and after geographic adjustments (where the national poverty rates are always anchored to 15.1%). The percentage change is calculated relative to a base poverty rate without geographic adjustments. Following Census Bureau standards, these state-level estimates are averaged over three years of the CPS ASEC (covering reference years 2009-2011).

[86] Note that Figure 2.2b uses a county-level template, even though the rates are calculated at the CBSA level and then assigned to all counties within that CBSA. If a county falls into one of the “other metropolitan” or “non-metropolitan” groups and no unit in that group is interviewed in the CPS, then that county is designated as having missing information (and shaded in white, like the counties with statistically insignificant differences in poverty rates).

Looking first at Figure 2.2a, we see that the states whose poverty rates increase the most with geographic adjustments are concentrated in coastal areas – namely New England, the mid-Atlantic region, and the West Coast – that typically have high rents. Specifically, the poverty rates for Connecticut, California, Hawaii, Maryland, Massachusetts, and New Jersey each increase by 25% or more after incorporating the geographic adjustment. On the other hand, states in the Deep South, Appalachia, and the Midwest – which typically have low housing rents – see lower poverty rates after the geographic adjustment. These patterns are starkest for Arkansas, Iowa, Kentucky, Mississippi, North Dakota, South Dakota, and West Virginia, with each seeing decreases in poverty rates of at least 25% after incorporating geographic adjustments.

Figure 2.2b. Percent Change in Poverty Rates After Geographic Adjustments (CBSA)



Data: 2010-2012 CPS ASEC (public-use)
Notes: This map shows the difference in poverty rates by CBSA before and after geographic adjustments (where the national poverty rates are always anchored to 15.1%). The percentage change is calculated relative to a base poverty rate without geographic adjustments. Rates are calculated at the CBSA level and applied to all counties in that CBSA. For areas outside of publicly identified CBSAs, we calculate rates for two general areas within a state – "non-metro" and "other metro". Rates are missing for some geographic areas because no individuals in those areas were interviewed in the survey. Following Census Bureau standards, these sub-state estimates are averaged over three years of the CPS ASEC (covering reference years 2009-2011).

One might be tempted to conclude from Figure 2.2a that state-level differences in rental costs drive much of the changes in poverty rate changes due to geographic adjustments. Yet, Figure 2.2b shows that there is substantial variation even within states – specifically between urban and rural areas. For example, California sees an increase in poverty of 37% at the state level after the geographic adjustment (driven by increases in the urban areas of Los Angeles, San Diego, and San Francisco), while rural areas in the San Joaquin Valley actually see decreases in poverty after a geographic adjustment. Conversely, while Mississippi experiences a decrease in poverty of 26% at the state level after a geographic adjustment (driven by decreases in its more rural areas), the urban areas of Jackson and Hattiesburg see decreases in poverty of less than 12.5% after geographic adjustments. More generally, the CBSA-level analysis in Figure 2.2b shows that incorporating a geographic adjustment appears to increase poverty rates in urban clusters and decrease poverty rates in rural areas.

The top halves of Tables 2.1a and 2.1b shed further light on the changes in poverty rates by geography. These tables show poverty rates before and after incorporating a geographic adjustment (along with the differences in those poverty rates) conditional on a set of characteristics. Table 2.1a uses the SPM and Table 2.1b uses the CIPM; within each table, Columns 1-3 pertain to estimates in the CPS and Columns 4-6 pertain to estimates in the SIPP. For now, we focus on estimates for individuals living in rural areas and in each of the nine Census Divisions. Focusing first on estimates using the SPM in the CPS (Table 2.1a), we find that the poverty rate for rural areas decreases from 14.0% to 11.4% after incorporating a geographic adjustment. This large and statistically significant decline corroborates a core finding from Figure 2.2b that a geographic adjustment indicates less poverty in rural areas. We observe similar patterns using the CIPM and looking at the SIPP, with the change in rural poverty rates being greater in the SIPP.

Using the SPM in the CPS, we also find that poverty rates increase following a geographic adjustment in the following Census divisions: New England (9.5% to 11.7%), Mid-Atlantic (13.1% to 14.4%), and Pacific (15.1% to 19.4%).[87] The difference in these rates is statistically significant at the 1% level for every region. Conversely, poverty rates decrease following a geographic adjustment in the following Census divisions: East North Central (14.6% to 13.4%), West North Central (12.3% to 10.0%), South Atlantic (16.3% to 15.9%), East South Central (18.7% to 14.2%), West South Central (17.6% to 15.6%), and Mountain (15.5% to 14.9%). The difference in these shares is statistically significant at the 1% level for every region except for South Atlantic (for which the difference is statistically significant at the 5% level). These patterns are once again consistent with the findings in Figures 2.2a and 2.2b, and they continue to largely hold when using the CIPM and looking at the SIPP.

2.4.2 Changes in Poverty Rates for Other Groups

Tables 2.1a and 2.1b also report poverty rates with and without a geographic adjustment for groups defined by the race/ethnicity of the sharing unit head and family type. Focusing again on the SPM in the CPS (Columns 1-3 of Table 2.1a), we see that poverty rates increase following a geographic adjustment for units with an Asian or Hispanic head, while the differences are statistically insignificant for units with a white, black, or other race head. The differences in Hispanic and Asian percentages may be driven by the substantial increases in poverty rates after a geographic adjustment in the Pacific region, which has large Hispanic and Asian populations. We

[87] See https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf for a map of the Census divisions.

Table 2.1a. Poverty Rates with and without Geographic Adjustments (SPM)

Characteristic	CPS Poor (No Geog. Adj.) (1)	CPS Poor (Geog. Adj.) (2)	CPS (2) minus (1) (3)	SIPP Poor (No Geog. Adj.) (4)	SIPP Poor (Geog. Adj.) (5)	SIPP (5) minus (4) (6)
Rural	13.96	11.39	-2.56***	19.96	14.31	-5.65***
Census Division						
New England	9.54	11.69	2.14***	10.75	12.00	1.25***
Mid-Atlantic	13.10	14.38	1.28***	11.62	13.41	1.79***
East North Central	14.63	13.42	-1.21***	14.72	12.09	-2.63***
West North Central	12.33	10.03	-2.30***	13.31	10.29	-3.03***
South Atlantic	16.31	15.87	-0.44**	16.84	16.18	-0.66***
East South Central	18.71	14.22	-4.49***	21.65	17.98	-3.68***
West South Central	17.63	15.62	-2.01***	17.17	14.67	-2.50***
Mountain	15.54	14.87	-0.67***	16.04	15.07	-0.97**
Pacific	15.11	19.44	4.33***	13.93	20.17	6.25***
Race/Ethnicity of Head						
White	13.62	13.54	-0.08	13.89	13.91	0.02
Black	24.22	23.76	-0.46	21.49	20.93	-0.56*
Asian	13.05	15.47	2.42***	14.44	17.22	2.78***
Other Race	20.04	20.48	0.44	21.18	19.69	-1.49*
Hispanic	24.43	26.87	2.44***	22.49	26.40	3.92***
Unit Type						
Elderly	15.77	15.40	-0.37*	14.05	13.44	-0.61**
Single Parent	32.67	31.43	-1.24***	34.68	32.81	-1.87**
Multiple Parents	12.98	13.37	0.40***	13.05	13.55	0.50*
Single Childless	25.89	25.63	-0.25	26.97	26.85	-0.11
Multiple Childless	10.02	9.94	-0.07	10.08	10.04	-0.04
Observations	205,000	205,000	205,000	88,000	88,000	88,000

Data: 2011 CPS ASEC and 2008 SIPP Panel

Notes: This table shows the poverty rate (weighted) of individuals who have a certain characteristic alongside the difference between the poverty rates. Sample consists of all individuals in each survey, and estimates are weighted using individual survey weights. *** p<0.01, ** p<0.05, * p<0.1. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization number CBDRB-FY21-ERD002-020.

Table 2.1b. Poverty Rates with and without Geographic Adjustments (CIPM)

Characteristic	CPS Poor (No Geog. Adj.) (1)	CPS Poor (Geog. Adj.) (2)	CPS (2) minus (1) (3)	SIPP Poor (No Geog. Adj.) (4)	SIPP Poor (Geog. Adj.) (5)	SIPP (5) minus (4) (6)
Rural	14.41	11.83	-2.59***	18.98	14.42	-4.55***
Census Division						
New England	8.56	10.52	1.96***	9.14	11.09	1.95***
Mid-Atlantic	12.57	14.50	1.93***	12.66	13.47	0.81***
East North Central	14.30	12.92	-1.38***	15.54	13.68	-1.85***
West North Central	12.84	10.94	-1.90***	11.61	9.63	-1.98***
South Atlantic	15.76	15.60	-0.16	14.98	15.08	0.10
East South Central	20.20	14.81	-5.39***	19.77	16.00	-3.77***
West South Central	18.02	16.07	-1.94***	17.47	15.63	-1.84***
Mountain	16.43	15.71	-0.72***	16.99	16.55	-0.45
Pacific	15.18	19.07	3.89***	15.73	19.90	4.17***
Race/Ethnicity of Head						
White	13.83	13.62	-0.21**	13.60	13.42	-0.18
Black	22.59	22.99	0.40	23.66	24.25	0.59
Asian	14.01	16.52	2.51***	13.48	15.82	2.34***
Other Race	19.90	20.02	0.12	21.41	20.70	-0.70
Hispanic	24.15	26.42	2.27***	27.28	30.68	3.40***
Unit Type						
Elderly	11.98	11.40	-0.58***	10.27	9.73	-0.55**
Single Parent	37.26	36.57	-0.69*	41.19	40.91	-0.28
Multiple Parents	14.26	14.61	0.34**	14.85	15.25	0.40
Single Childless	25.66	25.58	-0.08	27.02	26.61	-0.41**
Multiple Childless	8.81	8.73	-0.08	7.44	7.25	-0.19
Observations	170,000	170,000	170,000	85,000	85,000	85,000

Data: 2011 CPS ASEC and 2008 SIPP Panel linked to various administrative data.
Notes: This table shows the poverty rate (weighted) of individuals who have a certain characteristic alongside the difference between the poverty rates. Sample consists of individuals in PIKed sharing units (and, additionally in the CPS, no whole imputes), and estimates are weighted using individual survey weights adjusted for non-PIKing at the sharing unit level (and additionally for whole imputes in the CPS). Standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization number CBDRB-FY21-ERD002-020.

also find that poverty rates decrease among elderly and single parent families and increase among multiple parent families following a geographic adjustment, while the differences in poverty rates for elderly, single childless, and multiple childless units are statistically insignificant.

Turning next to estimates using the SPM in the SIPP (Columns 4-6 of Table 2.1a), we observe that many of the same patterns hold – although we now see slight decreases in poverty rates for black and other race units following a geographic adjustment. Many of these patterns also persist when using the CIPM. Concentrating first on the estimates in the CPS (Columns 1-3 of Table 2.1b), we find that poverty rates significantly decrease for white and elderly units and increase for Asian, Hispanic, and multiple parent units following a geographic adjustment. Similarly, in the SIPP (Columns 4-6 of Table 2.1b), poverty rates significantly decrease for elderly and single childless units and increase for Asian and Hispanic units following a geographic adjustment.

2.5 Methods: Deprivation Measures

In this section, we describe the empirical methods used to assess whether a poverty measure with or without a geographic adjustment does a better job of identifying material deprivation. We start by examining the measures of well-being analyzed and then discuss the methods used to compare deprivation across poverty measures.

2.5.1 Measures of Material Well-Being

We analyze ten broad domains of material well-being indicators. Four of these domains can be examined using both the CPS and SIPP (either because they are derived from administrative data or because they are asked about in both surveys), while six of these domains are specific to the SIPP. We chose these outcomes ex ante, and they constitute a broader set of outcomes than those used in Meyer, Wu, Mooers et al. (2021). The material well-being indicators in this paper also overlap with those used in Fox and Warren (2018) and Iceland et al. (2021).

Measures Available in Both Surveys (CPS and SIPP)

In both the CPS and SIPP, we examine permanent income, mortality, education, and health. Our measure of permanent income comes from tax records and is defined as the sum of income from tax records for 2008, 2009, 2011, and 2012. We use the PCE deflator to convert all amounts to 2010 dollars. If a person filed a Form 1040, then we use the AGI reported on this form as his/her income. If a person did not file a Form 1040, then we use the sum of incomes reported on Forms W-2 and 1099-R. In our analysis, we analyze permanent income at the SPM unit level and adjust income according to the equivalence scale recommended in Citro and Michael (1995).[88] Next, we use the SSA's Numident file to construct two measures of mortality: having died by December 31, 2015 and having died by March 1, 2019. For each of these measures, we assess mortality at both the individual and SPM unit head level.

We use years of education for an SPM unit's head as our education outcome in the CPS and SIPP. We compute years of education from an individual's survey-reported highest grade completed.[89] Lastly, we examine self-reported health in both the CPS and the SIPP. In both surveys, we report the presence of fair or poor health quality at both the head and individual levels.

[88] The equivalence scale is of the form $(A + PK)^F$, where A and K respectively designate the number of adults and children in the SPM unit. Following Meyer and Sullivan (2012b), we set $P = F = 0.7$.

[89] In cases where an individual reports a range of values for highest grade completed, (e.g., 1st to 4th grade, 5th to 6 grade, 7th to 8th grade), we take the midpoint. We set years of education to 14 for associate's degree, 16 for bachelor's degree, 18 for master's degree, 20 for professional school degree, and 21 for doctorate degree.

In the SIPP, we look at two other binary health indicators: having a condition that limits the kind or amount of work you can do and having a condition that prevents work.

Measures Available in SIPP Only

In the SIPP, we additionally analyze six broad domains of well-being: material hardships, home quality problems, appliances owned, assets owned, food security, and public services/safety. Our data for material hardships, home quality problems, and appliances come from the Wave 6 topical module. For each of these three domains, we use the same variables as in Meyer, Wu, Mooers et al. (2021). For material hardships, we examine the following eight binary indicators: not meeting all essential expenses, not paying full rent, being evicted because of rent, not paying full energy bill, having energy service disrupted, having telephone service disconnected, needing to see doctor but being unable to go, needing to see dentist but being unable to go, and not having enough food. The seven binary home quality indicators that we analyze cover the presence of pests, leaking roof, broken windows, electrical problems, plumbing problems, cracks in walls, and holes in floor. The eight appliances owned we consider are microwaves, dishwashers, air conditioning, televisions, personal computers (PCs), washing machines, dryers, and cell phones.

Next, we use the topical modules for Waves 4 and 7 on assets and liabilities to define measures of assets owned. We consider five different asset measures: total assets, home equity, vehicle equity, other assets, and net worth. Total assets cover the sum of home equity, vehicle equity, and other assets. Other assets consist of interest-earning assets, stocks and bonds, IRA and KEOGH accounts, 401(k) and Thrift accounts, business equity, and SIPP's blanket variable for other assets. We then calculate net worth as total assets minus total debt (secured and unsecured).[90]

[90] This differs from SIPP's definition of net worth, which consists of total assets minus total unsecured debt.

Next, we again use the Wave 6 topical module to define our measures of food security. The eight binary food security indicators that we consider are: not eating sufficient food in household, not having enough to eat in house, buying food that did not last, not being able to afford balanced meals, children not eating enough, cutting size or skipping meals, eating less than you feel you should, and not eating for a whole day.

Finally, we use the Wave 6 topical module to define our public services and safety measures. The twelve binary indicators that we consider are: having inadequate public transportation, being afraid to walk alone at night, carrying anything for safety when going out, having undesirable public services, being unsatisfied with fire department, being unsatisfied with the area's hospitals, being unsatisfied with the area's police, being unsatisfied with the area's public schools, being unsatisfied with the area's public services, staying at home for safety reasons, taking someone with you when going out for safety reasons, and having the threat of crime be enough that you would move.

2.5.2 Analytical Methods for Comparing Well-Being Across Poverty Measures

To empirically assess whether a poverty measure with or without a geographic adjustment better identifies a materially deprived population, we regress a measure of material well-being on 1) indicators for poverty status with and without a geographic adjustment and 2) covariates reflecting characteristics of the sharing unit or the head of the unit. Formally, we estimate the following regression using the sample of all sharing unit heads (and, for some outcomes related to mortality and health, all individuals):

$$\textit{Well-Being} = \alpha + \beta_1\ \textit{Geographic-Only Poor} + \beta_2\ \textit{Always Poor} + \beta_3\ \textit{Never Poor} + \boldsymbol{\lambda}'\boldsymbol{X} + \varepsilon, \qquad (2.5)$$

where *Geographic-Only Poor* is an indicator for being classified as poor with a geographic adjustment but not without, *Always Poor* is an indicator for being poor both with and without a geographic adjustment, and *Never Poor* is an indicator for being non-poor whether or not a geographic adjustment is used. The reference group is therefore those who are poor without a geographic adjustment but not poor with a geographic adjustment (i.e., non-geographic-only poor). We compare the non-overlapping groups that are affected by geographic adjustments (i.e., those added to and removed from poverty) to make our analyses clearer, as most individuals in poverty are not affected by geographic adjustments. ***X*** is a vector of characteristics of the sharing unit or its head that includes age, age-squared, an indicator for being female, an indicator for being married, an indicator for having a cohabiting partner, the number of adults in the sharing unit, the number of children in the sharing unit, indicators for race (i.e., dummies for White, Black, Asian, and other), a binary indicator for being Hispanic, and each of the race/ethnicity dummies interacted with being female. We discuss alternative estimates with a more limited set of covariates below.

For binary outcomes, we estimate equation (2.5) using a probit model and calculate average partial effects (APEs) of β_1 that are averaged over the geographic- and non-geographic-only poor subgroups.[91] For non-binary outcomes, we estimate equation (2.5) using a linear model. We report heteroskedastic-robust standard errors, and we use replicate weights to obtain standard errors accommodating complexities in the surveys' designs. We weight using individual survey weights corresponding to the sharing unit head multiplied by the number of individuals in the sharing unit.[92] In evaluating a geographic adjustment, we want to compare individuals whose poverty

[91] Note that it is crucial here to calculate APEs over only the part of the overall sample that provides the relevant identifying variation. Otherwise, calculating APEs over the entire sample (i.e., including the always and never poor subgroups) may lead to unrepresentative estimates of APEs from the probit models.
[92] For regressions at the individual level with mortality or health outcomes, we weight using individual survey weights.

status changes when switching between measures with and without a geographic adjustment. In the framework of equation (2.5), the relevant coefficient is β_1. If β_1 is less than 0, then the geographic-only poor are more likely to be disadvantaged than the non-geographic-only poor (assuming a higher value for the dependent variable signifies greater well-being). This would imply that incorporating a geographic adjustment helps to better identify a more deprived population.

2.5.3 Shares and Counts by Geographic Poverty Category

In Table 2.2, we report the weighted shares of individuals and un-weighted counts of individuals and sharing units for each of our four mutually exclusive and exhaustive poverty categories in the CPS and SIPP (using the SPM and CIPM). These four categories are: those who are not poor with or without a geographic adjustment ("Never Poor"), those who are poor without a geographic adjustment but not poor with a geographic adjustment ("Non-Geographic-Only Poor"), those who are poor with a geographic adjustment but not poor without a geographic adjustment ("Geographic-Only Poor"), and those who are poor with and without a geographic adjustment ("Always Poor").

Starting first with the SPM, Column 1 of Table 2.2 shows that 83.38% of population-weighted individuals in the CPS are never poor, 13.58% of individuals are always poor, and 1.52% of individuals each are non-geographic-only and geographic-only poor. The most important groups for our analysis are the non-geographic-only and geographic-only poor, as these two "switcher" groups are what our subsequent regressions rely upon for identification. By construction, the weighted shares for these two groups should be the same since poverty rates with and without geographic adjustments are always anchored to 15.1%. For the SPM in the SIPP, Column (4) shows that the individuals in each of these “switcher” groups constitute 1.83% of the population. Turning next to the CIPM, we find that the non-geographic-only and geographic-only poor each make up 1.43% of the population in the CPS and 1.46% of the population in the SIPP.

Table 2.2. Shares and Counts by Geographic Poverty Category

	CPS			SIPP		
Poverty Category	**Weighted Share of Individuals**	**Sample # of Individuals**	**Sample # of Sharing Units**	**Weighted Share of Individuals**	**Sample # of Individuals**	**Sample # of Sharing Units**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
			A. Supplemental Poverty Measure (SPM)			
Never Poor	0.8338	173,000	65,000	0.8307	73,000	29,000
Non-Geog.-Only Poor	0.0152	3,100	1,300	0.0183	1,800	750
Geog.-Only Poor	0.0152	3,300	1,200	0.0183	1,500	500
Always Poor	0.1358	26,000	11,500	0.1327	12,000	5,500
			B. Comprehensive Income Poverty Measure (CIPM)			
Never Poor	0.8346	163,000	60,500	0.8344	71,500	28,500
Non-Geog.-Only Poor	0.0143	2,800	1,100	0.0146	1,400	550
Geog.-Only Poor	0.0143	2,800	950	0.0146	1,100	400
Always Poor	0.1367	24,000	9,800	0.1365	11,500	4,800

Data: 2011 CPS ASEC and 2008 SIPP Panel linked to various administrative data
Notes: This table shows the number of individuals (both weighted share and unweighted count) and unweighted number of sharing units in each of our four geographic poverty categories for the SPM and CIPM in both the CPS and SIPP. “Never Poor” refers to being not poor under either geographic adjustments or no geographic adjustments “Non-Geographic-Only Poor” refers to being poor under no geographic adjustments and not poor under geographic adjustments, “Geographic-Only Poor” refers to being poor under geographic adjustments and not poor under no geographic adjustments, and “Always Poor” refers to being poor under both geographic adjustments and no geographic adjustments. Poverty rates are always anchored to 15.1%, which is the official rate in the CPS. The sample for the SPM estimates consists of all observations in the surveys, and estimates are weighted using original survey weights in Columns (1) and (4). The sample for the CIPM estimates in the CPS consists of all individuals in sharing units where at least one member has a PIK and no member is whole imputed, and estimates are weighted using individual survey weights adjusted for non-PIKing and whole imputes (at the sharing unit level). The sample for the CIPM estimates in the SIPP consists of all individuals in sharing units where at least one member has a PIK, and estimates are weighted using individual survey weights adjusted for non-PIKing (at the sharing unit level). The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization number CBDRB-FY21-ERD002-020.

Columns 2 and 5 of Table 2.2 report the un-weighted counts of individuals by geographic poverty category, while Columns 3 and 6 report counts of sharing units. Even though poverty rates are calculated across individuals, the numbers of sharing units in Columns 3 and 6 are particularly relevant for this analysis because most of our main regressions are at the sharing unit level. The number of sharing units in each of the "switcher" groups is lower in the CIPM for a given survey than in the SPM for that same survey. There are also fewer sharing units in the SIPP than in the CPS, because of the SIPP's smaller sample size. However, even our smallest geography category (those who are geographic-only poor using the CIPM in the SIPP) includes 400 sharing units.

2.6 Main Regression Results

We now describe our main regression estimates that compare differences in a wide variety of well-being indicators between the SPM and CIPM with and without a geographic adjustment. We start by describing the results for the SPM before moving to discuss the results for the CIPM.[93]

2.6.1 Estimates for the Supplemental Poverty Measure

Table 2.3a shows regression estimates of a wide range of well-being outcomes – encompassing 71 measures spanning ten domains and two surveys – on an indicator for being geographic-only SPM poor (relative to being non-geographic-only SPM poor). The point estimates in Column 1 correspond to β_1 in equation (2.5) for linear outcomes (and the APEs of β_1 for binary outcomes), and Column 2 displays the heteroskedastic-robust standard errors associated with the point estimates. For every outcome, Columns 3 and 4 show the mean value for the non-geographic-only poor (i.e., the reference group against which to evaluate the regression coefficient for being geographic-only poor) and the overall mean. Finally, Column 5 displays an indicator for whether the signs of the point estimates in Column 1 signify that geographic adjustments identify a more deprived population (indicated by "+") or a less deprived population (indicated by "–").

[93] Even though the SPM analyses rely on survey information only, most cannot be produced using the public-use data. This is because the vast majority of our well-being indicators are found only in the SIPP, and geographic adjustment factors are at the CBSA level (which is identifiable only in the restricted-use SIPP and not the public-use SIPP).

Table 2.3a. Regression Estimates of Well-Being on Geographic-Only Poor (SPM)

Well-Being Measures	**Point Estimate** (1)	**Standard Error** (2)	**Mean for Non-Geog Poor** (3)	**Overall Mean** (4)	**Supports Geog Adj? (+/-)** (5)
<u>Permanent Income (CPS & SIPP)</u>					
CPS	28,630***	(6,278)	26,980	98,620	–
SIPP	17,150***	(3,819)	33,870	92,930	–
<u>Years of Education (CPS & SIPP)</u>					
CPS	0.4000***	(0.1410)	11.840	13.640	–
SIPP	0.5630***	(0.2270)	12.060	13.700	–
<u>Mortality (CPS & SIPP)</u>					
Died by 2015 (ind.) – *CPS*	-0.0038	(0.0071)	0.051	0.036	–
Died by 2019 (ind.) – *CPS*	-0.0149*	(0.0090)	0.098	0.064	–
Died by 2015 (head) – *CPS*	-0.0079	(0.0120)	0.071	0.040	–
Died by 2019 (head) – *CPS*	-0.0226	(0.0140)	0.124	0.071	–
Died by 2015 (ind.) – *SIPP*	-0.0047	(0.0073)	0.054	0.041	–
Died by 2019 (ind.) – *SIPP*	-0.0132	(0.0087)	0.089	0.069	–
Died by 2015 (head) – *SIPP*	-0.0068	(0.0130)	0.064	0.047	–
Died by 2019 (head) – *SIPP*	-0.0259	(0.0165)	0.110	0.080	–
<u>Health Problems (CPS & SIPP)</u>					
Poor/Fair Health (ind.) – *CPS*	-0.0603***	(0.0137)	0.226	0.118	–
Poor/Fair Health (head) – *CPS*	-0.0597***	(0.0219)	0.344	0.166	–
Poor/Fair Health (ind.) – *SIPP*	-0.0711***	(0.0148)	0.182	0.103	–
Poor/Fair Health (head) – *SIPP*	-0.1430***	(0.0333)	0.304	0.136	–
Health Limits Work – *SIPP*	-0.0889***	(0.0133)	0.165	0.089	–
Health Prevents Work – *SIPP*	-0.0675***	(0.0105)	0.114	0.056	–
<u>Material Hardships (SIPP)</u>					
Total Number	-0.1550	(0.1370)	1.164	0.646	–
Did Not Meet All Essential Expenses	-0.0386	(0.0377)	0.317	0.180	–
Did Not Pay Full Rent	-0.0208	(0.0257)	0.172	0.092	–
Evicted Because of Rent	0.0015	(0.0076)	0.007	0.005	+
Did Not Pay Full Energy Bill	-0.0338	(0.0326)	0.215	0.120	–
Had Energy Cut Off	-0.0060	(0.0130)	0.038	0.020	–
Had Telephone Service Cut Off	-0.0019	(0.0229)	0.084	0.043	–
Needed to See Doctor but Could Not	-0.0351	(0.0270)	0.162	0.084	–
Needed to See Dentist but Could Not	-0.0314	(0.0259)	0.168	0.102	–

*** p<0.01, ** p<0.05, * p<0.1

Table 2.3a. Regression Estimates of Well-Being on Geographic-Only Poor (SPM) – cont.

Well-Being Measures	Point Estimate (1)	Standard Error (2)	Mean for Non-Geog Poor (3)	Overall Mean (4)	Supports Geog Adj? (+/-) (5)
Home Quality Problems (SIPP)					
Total Number	-0.1230*	(0.0654)	0.370	0.224	–
Pests	-0.0206	(0.0250)	0.112	0.080	–
Leaking Roof	-0.0391*	(0.0205)	0.091	0.050	–
Broken Windows	-0.0556***	(0.0165)	0.082	0.031	–
Electrical Problems	0.0010	(0.0067)	0.012	0.007	+
Plumbing Problems	0.0004	(0.0109)	0.022	0.020	+
Holes or Cracks in Wall	0.0014	(0.0163)	0.037	0.029	+
Holes in Floor	-0.0113*	(0.0066)	0.014	0.007	–
Appliances (SIPP)					
Total Number	-0.1900	(0.1660)	6.207	6.988	+
Microwave	-0.0070	(0.0210)	0.942	0.975	+
Dishwasher	0.0584	(0.0371)	0.476	0.711	–
Air Conditioning	-0.1750***	(0.0327)	0.891	0.886	+
Television	0.0023	(0.0173)	0.961	0.985	–
Personal Computer	0.1440***	(0.0404)	0.550	0.793	–
Washing Machine	-0.1560***	(0.0317)	0.845	0.879	+
Dryer	-0.1160***	(0.0390)	0.784	0.858	+
Cell Phone	0.1110***	(0.0313)	0.758	0.900	–
Assets (SIPP)					
Net Worth	59,580	(45,130)	42,350	272,200	–
Total Wealth	96,560**	(44,810)	82,930	384,900	–
Total Debt	36,980***	(10,130)	40,590	112,700	+
Home Equity	31,920***	(11,880)	40,200	114,400	–
Vehicle Equity	1,279***	(483)	3,743	7,324	–
Other Assets	63,370	(43,160)	38,990	263,200	–
Food Security Problems (SIPP)					
Total Number	-0.1750	(0.1220)	1.022	0.460	–
Not Enough Food	-0.0383**	(0.0161)	0.080	0.026	–
Food Bought Did Not Last	-0.0430	(0.0329)	0.289	0.147	–
Could Not Afford Balanced Meals	-0.0524*	(0.0288)	0.253	0.130	–
Children Not Eating Enough	0.0118	(0.0223)	0.069	0.032	+
Cut Size or Skipped Meals	-0.0330	(0.0242)	0.136	0.053	–
Ate Less Than Felt One Should	-0.0344	(0.0250)	0.144	0.058	–
Did Not Eat for Whole Day	-0.0044	(0.0138)	0.051	0.015	–

*** p<0.01, ** p<0.05, * p<0.1

Table 2.3a. Regression Estimates of Well-Being on Geographic-Only Poor (SPM) – cont.

Well-Being Measures	Point Estimate (1)	Standard Error (2)	Mean for Non-Geog Poor (3)	Overall Mean (4)	Supports Geog Adj? (+/-) (5)
Public Services and Safety (SIPP)					
Total Num. Public Services Problems	-0.0298	(0.1300)	1.165	0.949	–
Inadequate Public Transportation	-0.1220***	(0.0384)	0.263	0.206	–
Afraid to Walk Alone at Night	0.0286	(0.0274)	0.216	0.201	+
Carry Anything When Going Out	-0.0049	(0.0136)	0.054	0.060	–
Public Services Undesirable	0.0086	(0.0133)	0.024	0.019	+
Unsatisfied with Fire Department	-0.0021	(0.0134)	0.024	0.014	–
Unsatisfied with Hospitals	-0.0150	(0.0171)	0.091	0.064	–
Unsatisfied with Police	-0.0419**	(0.0206)	0.104	0.051	–
Unsatisfied with Public Schools	0.0159	(0.0191)	0.052	0.050	+
Unsatisfied with Public Services	-0.0191	(0.0166)	0.059	0.047	–
Stayed at Home for Safety Reasons	0.0571**	(0.0237)	0.118	0.102	+
Take Someone with You for Safety	0.0657***	(0.0215)	0.092	0.089	+
Threat of Crime Enough to Move	-0.0001	(0.0141)	0.069	0.045	–

*** p<0.01, ** p<0.05, * p<0.1

Data: 2011 CPS ASEC and 2008 SIPP Panel linked to IRS Forms 1040/W-2/1099-R and SSA Numident

Notes: This table shows the coefficient on an indicator for being geographic-only poor (vs. non-geographic-only poor) for regressions of a wide variety of well-being measures on indicators for being in one of three geographic SPM poverty categories (omitting the non-geographic-only poor) and a vector of covariates (for the sharing unit or its head) including age, gender, marital status, unit type, and race/ethnicity. Sample consists of all sharing unit heads for most outcomes, except for some of the mortality and health outcomes (which are at the individual level) and mortality and permanent income outcomes (where we restrict to PIKed units in both surveys and non-whole-imputed units in CPS). Probit APEs are reported for binary outcomes. Robust standard errors are in parentheses and are calculated using replicate weights. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization numbers CBDRB-FY20-ERD002-020 and CBDRB-FY2021-CES005-01.

We start by discussing the results for permanent income, where higher amounts signify greater well-being. We find that the geographic-only poor have higher permanent income by $28,630 in the CPS and $17,150 in the SIPP, both of which are statistically significantly different from zero at the 1% level. When evaluated against the means for the non-geographic-only poor in Column (3), these effects suggest that the geographic-only poor have approximately 100% higher permanent income in the CPS and 50% higher permanent income in the SIPP than the non-geographic-only poor. In other words, low incomes appear less permanent in high-cost areas, which is a noteworthy result given the damaging effects of persistent poverty found in prior studies (see, e.g., Duncan and Rodgers 1991). Looking next at years of education for the SPM unit head (where higher values again signal greater well-being), we find that the geographic-only poor have 0.40 and 0.56 more years of education than the non-geographic-only poor in the CPS and SIPP, respectively. Both of these effects are again statistically significant at the 1% level and translate into the geographic-only poor having 3% and 5% more years of education in the surveys than the non-geographic-only poor.

We subsequently turn to mortality, where a higher probability is now associated with lower well-being. We find that the geographic-only poor are associated with lower mortality rates than the non-geographic-only poor for every measure analyzed, although none of the estimates are statistically significant. The statistical imprecision of these estimates may stem from mortality being a relatively infrequent outcome, with the estimates associated with dying by 2015 generally being more imprecise than the estimates associated with dying by 2019. The estimates in the CPS also tend to be slightly less noisy than those in the SIPP, with the CPS having roughly double the sample size of the SIPP. Indeed, the estimate associated with the SPM unit head dying by 2019 in the CPS is only marginally insignificant at the 10% level. Turning from mortality to health problems more generally, we find in the CPS that the geographic-only poor are 27% and 17% less likely than the non-geographic-only poor to have poor or fair health quality at the individual and head levels, respectively. In the SIPP, the geographic-only poor are 39% and 47% less likely to have poor or fair health quality at the individual and head levels (respectively), 54% less likely to have a health condition that limits work, and 59% less likely to have a health condition that prevents work. All of these effects are statistically significant at the 1% level.

We next examine eight different material hardships in the SIPP as well as a summary measure of the total number of hardships (with more hardships being associated with lower well-being). The geographic-only poor have 0.16 (13%) fewer total hardships than the non-geographic-only poor, but this estimate is statistically insignificant. Out of the eight individual hardship measures, seven are associated with lower deprivation after geographic adjustments. However, none of the estimates for the individual hardship outcomes are statistically significant. We also examine seven different home quality problems in the SIPP along with a summary measure of the total number of home quality problems (with more home quality problems being associated with lower well-being). The geographic-only poor have 0.12 (33%) fewer total home quality problems than the non-geographic-only poor, with this estimate being statistically significant at the 10% level. Four out of the seven individual problems are associated with lower deprivation after geographic adjustments (with the estimates for leaking roof, broken windows, and holes in floor being statistically significant at the 10% level), while the three individual problems that suggest greater deprivation have statistically insignificant estimates.

We now turn to analyzing the ownership of eight different appliances in the SIPP and a summary measure of the total number of appliances (with more appliances being associated with greater well-being). The geographic-only poor have 0.19 (3%) fewer total appliances than the non-geographic-only poor. While this estimate is small and statistically imprecise, it is also the only summary measure estimate to suggest that geographic adjustments may identify a more deprived population in poverty. Breaking down the results by individual appliances shows that the geographic-only poor have greater ownership of four appliances (with ownership of computer and cell phone statistically significant at the 1% level) and lower ownership of four other appliances (with ownership of air conditioning, washer, and dryer statistically significant at the 1% level).

However, the differences in ownership of air conditioning and washers/dryers may in part reflect the characteristics of the locations in which the groups reside. Specifically, the geographic-only poor tend to be located in California and the Northeast while the non-geographic-only poor tend to be located in the Deep South; the former regions are likely to be cooler (with less of a need for air conditioning) and denser (leading to fewer in-unit washers and dryers) than the latter. These hypotheses hold up after controlling for average monthly temperatures at the county level and the proportion of single-family homes by county. After including these covariates, the geographic-only poor have (statistically insignificantly) more total appliances than the non-geographic-only poor, with the estimates for air conditioning and dryer now statistically insignificant.

We next examine the ownership of assets (with higher amounts associated with greater well-being) and find that the geographic-only poor have $59,580 (141%) more net worth than the non-geographic-only poor, although this difference is not statistically significant. This result is due to the geographic-only poor having significantly higher amounts of both total wealth and total debt than the non-geographic-only poor. It is worth noting that positive levels of debt may not necessarily reflect increased disadvantage (even though we treat it as such), as debt indicates the ability to borrow and allows one to consume. Breaking down the estimate for total wealth by its components, we find that the geographic-only poor have $31,920 (79%) more in home equity, $1,279 (34%) more in vehicle equity, and $63,370 (163%) more in other assets (which include checking and savings accounts, retirement accounts, stocks and bonds, etc.). The estimates for home and vehicle equity are statistically significant at the 1% level, while the estimate for other assets is not statistically different from zero. Note that it is not the case that unconditional home values are mechanically correlated with median rents – while it is true that home values are positively correlated with median rents conditional on owning a home, home ownership itself is likely negatively correlated with median rents.

Next, we examine the presence of seven different food security problems in the SIPP and a summary measure of the total number of food security problems (with more food security problems being associated with lower well-being). The geographic-only poor have 0.18 (17%) fewer total food security problems than the non-geographic-only poor, although this estimate is statistically insignificant. Six out of the seven individual food security problems are associated with lower deprivation after a geographic adjustment, although only the estimates corresponding to not having enough food to eat and not being able to afford balanced meals are statistically significant at conventional levels.

For the final domain of outcomes that we assess, we examine the presence of twelve public services/safety problems in the SIPP and a summary measure of the total number of problems (with more public services problems being associated with lower well-being). The geographic-only poor have 0.03 (3%) fewer public services/safety problems than the non-geographic-only poor, although this estimate is statistically insignificant. Out of the twelve individual public services problems, seven are associated with lower deprivation after geographic adjustments (having inadequate public transportation and being unsatisfied with the police are statistically significant at the 1% and 5% level, respectively). Of the five measures that are associated with higher deprivation after geographic adjustments, two are statistically significant (having stayed at home for safety and having taken someone with you for safety reasons) at conventional levels.

In summary, we find strong evidence that incorporating a geographic adjustment to SPM poverty thresholds identifies a less deprived poor population than those who are otherwise poor without a geographic adjustment. For 55 of the 71 total well-being indicators that we analyze, we find that the geographic-only poor appear to be less deprived than the non-geographic-only poor (with estimates being statistically significant for 24 outcomes). Note that many of the statistics within a given survey (e.g., SIPP) are not independent if the outcomes are correlated, while statistics are independent across surveys even if they correspond to a similar outcome. A caveat of these results, however, is that the SPM may classify some individuals with high levels of well-being as being in poverty, as it relies on survey-reported incomes that are subject to underreporting and does not explicitly account for the flow value of assets.[94] Partly as a result of these issues, we also examine results using an alternative poverty measure (the CIPM) that corrects for misreporting using administrative data and explicitly incorporates the income flow from assets.

2.6.2 Estimates for the Comprehensive Income Poverty Measure

Table 2.3b presents the analog to the estimates in Table 2.3a using the CIPM. Once again, our results broadly show that a geographic adjustment appears to identify a less deprived population in poverty. We find that the geographic-only poor have $24,140 (99%) and $17,800 (72%) more in permanent income in the CPS and SIPP, respectively, and 0.57 (5%) and 0.70 (6%) more years of education in the CPS and SIPP than the non-geographic-only poor. Each of these estimates is statistically significant at the 5% level. We also find that the geographic-only poor have lower rates of mortality than the non-geographic-only poor for five of the eight mortality measures analyzed; for one of these measures (death by 2019 in the CPS for the unit head), our estimates are statistically significant at the 5% level. In contrast, the estimates for the three measures (death by 2015 for both the individual and unit head as well as death by 2019 for the unit head, all in the SIPP) that point to higher rates of mortality among the geographic-only poor are statistically insignificant. Turning to health problems more generally in the SIPP, we find that the geographic-only poor are 21% and 31% less likely to have poor or fair health quality at the

[94] As an implication of this caveat, we find that those who are poor both with and without a geographic adjustment ("Always Poor") have lower levels of deprivation than the non-geographic-only poor across a number of domains, including permanent income, years of education, health problems, total wealth, and food security problems.

Table 2.3b. Regression Estimates of Well-Being on Geographic-Only Poor (CIPM)

Well-Being Measures	Point Estimate (1)	Standard Error (2)	Mean for Non-Geog Poor (3)	Overall Mean (4)	Supports Geog Adj? (+/-) (5)
Permanent Income (CPS & SIPP)					
CPS	24,140***	(5,377)	24,430	98,090	–
SIPP	17,800***	(3,461)	24,850	92,930	–
Years of Education (CPS & SIPP)					
CPS	0.5730***	(0.1610)	11.750	13.630	–
SIPP	0.6980**	(0.2850)	11.930	13.700	–
Mortality (CPS & SIPP)					
Died by 2015 (ind.) – *CPS*	-0.0046	(0.0055)	0.048	0.036	–
Died by 2019 (ind.) – *CPS*	-0.0110	(0.0069)	0.088	0.064	–
Died by 2015 (head) – *CPS*	-0.0069	(0.0082)	0.058	0.040	–
Died by 2019 (head) – *CPS*	-0.0247**	(0.0116)	0.109	0.071	–
Died by 2015 (ind.) – *SIPP*	0.0035	(0.0114)	0.052	0.041	+
Died by 2019 (ind.) – *SIPP*	-0.0037	(0.0115)	0.078	0.069	–
Died by 2015 (head) – *SIPP*	0.0235	(0.0215)	0.060	0.047	+
Died by 2019 (head) – *SIPP*	0.0209	(0.0220)	0.088	0.080	+
Health Problems (CPS & SIPP)					
Poor/Fair Health (ind.) – *CPS*	-0.0683***	(0.0147)	--	--	–
Poor/Fair Health (head) – *CPS*	-0.0863***	(0.0238)	--	--	–
Poor/Fair Health (ind.) – *SIPP*	-0.0372**	(0.0142)	0.181	0.104	–
Poor/Fair Health (head) – *SIPP*	-0.0793**	(0.0320)	0.254	0.136	–
Health Limits Work – *SIPP*	-0.0333**	(0.0146)	0.162	0.089	–
Health Prevents Work – *SIPP*	-0.0214*	(0.0121)	0.122	0.056	–
Material Hardships (SIPP)					
Total Number	-0.1830	(0.1510)	1.280	0.645	–
Did Not Meet All Essential Expenses	-0.0537	(0.0399)	0.346	0.179	–
Did Not Pay Full Rent	-0.0065	(0.0342)	0.160	0.093	–
Evicted Because of Rent	-0.0054	(0.0109)	0.017	0.005	–
Did Not Pay Full Energy Bill	-0.0387	(0.0380)	0.242	0.121	–
Had Energy Cut Off	-0.0259	(0.0180)	0.053	0.020	–
Had Telephone Service Cut Off	-0.0292	(0.0260)	0.095	0.042	–
Needed to See Doctor but Could Not	-0.0339	(0.0318)	0.166	0.084	–
Needed to See Dentist but Could Not	-0.0392	(0.0315)	0.202	0.102	–

*** p<0.01, ** p<0.05, * p<0.1

Table 2.3b. Regression Estimates of Well-Being on Geographic-Only Poor (CIPM) – cont.

Well-Being Measures	Point Estimate (1)	Standard Error (2)	Mean for Non-Geog Poor (3)	Overall Mean (4)	Supports Geog Adj? (+/-) (5)
<u>Home Quality Problems (SIPP)</u>					
Total Number	-0.1560**	(0.0747)	0.410	0.225	–
Pests	-0.0627**	(0.0243)	0.157	0.080	–
Leaking Roof	-0.0497**	(0.0250)	0.091	0.051	–
Broken Windows	-0.0065	(0.0186)	0.050	0.031	–
Electrical Problems	0.0041	(0.0085)	0.009	0.007	+
Plumbing Problems	-0.0193	(0.0181)	0.049	0.019	–
Holes or Cracks in Wall	-0.0144	(0.0172)	0.046	0.029	–
Holes in Floor	-0.0096**	(0.0047)	0.008	0.007	–
<u>Appliances (SIPP)</u>					
Total Number	-0.2300	(0.1620)	6.273	6.992	+
Microwave	0.0230	(0.0226)	0.940	0.976	–
Dishwasher	0.0930**	(0.0381)	0.429	0.711	–
Air Conditioning	-0.0932**	(0.0364)	0.858	0.887	+
Television	0.0195	(0.0159)	0.973	0.985	–
Personal Computer	0.1030**	(0.0419)	0.624	0.795	–
Washing Machine	-0.2250***	(0.0355)	0.858	0.880	+
Dryer	-0.1740***	(0.0423)	0.811	0.859	+
Cell Phone	0.0972***	(0.0290)	0.781	0.900	–
<u>Assets (SIPP)</u>					
Net Worth	26,070	(26,820)	6,785	273,500	–
Total Wealth	61,530**	(24,980)	29,120	386,300	–
Total Debt	35,450***	(8,693)	22,330	112,900	+
Home Equity	505	(6,785)	19,100	114,000	–
Vehicle Equity	867	(581)	3,062	7,324	–
Other Assets	60,150***	(21,320)	6,957	265,000	–
<u>Food Security Problems (SIPP)</u>					
Total Number	-0.2110*	(0.1280)	0.948	0.459	–
Not Enough Food	-0.0312	(0.0189)	0.061	0.026	–
Food Bought Did Not Last	-0.0624	(0.0393)	0.284	0.146	–
Could Not Afford Balanced Meals	-0.0588	(0.0361)	0.258	0.129	–
Children Not Eating Enough	0.0101	(0.0188)	0.040	0.032	+
Cut Size or Skipped Meals	-0.0529**	(0.0241)	0.128	0.053	–
Ate Less Than Felt One Should	-0.0451*	(0.0260)	0.129	0.058	–
Did Not Eat for Whole Day	-0.0252	(0.0165)	0.048	0.015	–

*** p<0.01, ** p<0.05, * p<0.1

Table 2.3b. Regression Estimates of Well-Being on Geographic-Only Poor (CIPM) – cont.

Well-Being Measures	Point Estimate (1)	Standard Error (2)	Mean for Non-Geog Poor (3)	Overall Mean (4)	Supports Geog Adj? (+/-) (5)
Public Services and Safety (SIPP)					
Total Num. Public Services Problems	-0.0694	(0.1390)	1.332	0.950	–
Inadequate Public Transportation	-0.1270***	(0.0356)	0.299	0.207	–
Afraid to Walk Alone at Night	0.0362	(0.0337)	0.249	0.202	+
Carry Anything When Going Out	-0.0082	(0.0166)	0.067	0.061	–
Public Services Undesirable	0.0189	(0.0204)	--	0.019	+
Unsatisfied with Fire Department	--	--	0.040	0.014	–
Unsatisfied with Hospitals	-0.0365	(0.0228)	0.110	0.064	–
Unsatisfied with Police	-0.0148	(0.0205)	0.094	0.051	–
Unsatisfied with Public Schools	-0.0251	(0.0218)	0.081	0.050	–
Unsatisfied with Public Services	-0.0264	(0.0214)	0.073	0.047	–
Stayed at Home for Safety Reasons	0.0345	(0.0296)	0.141	0.102	+
Take Someone with You for Safety	0.0585**	(0.0293)	0.096	0.089	+
Threat of Crime Enough to Move	0.0182	(0.0202)	0.048	0.045	+

*** p<0.01, ** p<0.05, * p<0.1

Data: 2011 CPS ASEC and 2008 SIPP Panel linked to various administrative data

Notes: This table shows the coefficient on an indicator for being geographic-only poor (vs. non-geographic-only poor) for regressions of a wide variety of well-being measures on indicators for being in one of three geographic CIPM poverty categories (omitting the non-geographic-only poor) and a vector of covariates (for the sharing unit or its head) including age, gender, marital status, unit type, and race/ethnicity. Sample consists of all heads in PIKed sharing units (and no whole imputes in the CPS) for most outcomes, except for some of the mortality and health outcomes (which are at the individual level). Probit APEs are reported for binary outcomes. Robust standard errors are in parentheses and are calculated using replicate weights. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization numbers CBDRB-FY20-ERD002-020 and CBDRB-FY2021-CES005-01.

individual and head level (respectively), 21% less likely to have a health condition that limits work, and 18% less likely to have a health condition that prevents work than the non-geographic-only poor. The health estimates are of the same sign and slightly larger magnitudes in the CPS, with all of these estimates being statistically significant at the conventional levels.

We also find that the geographic-only poor have 0.18 (14%) fewer total material hardships than the non-geographic-only poor, with all eight individual hardship measures associated with lower deprivation after a geographic adjustment. However, none of these hardship estimates are statistically significant. Additionally, the geographic-only poor have 0.16 (38%) fewer total home quality problems than the non-geographic-only poor, with this estimate being statistically significant at the 5% level. Six of the seven individual home quality problems are associated with lower deprivation after a geographic adjustment, with only the estimates for pests, leaking roof, and holes in the floor being statistically significant at conventional levels. The geographic-only poor have 0.23 (4%) fewer total appliances than the non-geographic-only poor, but this summary estimate is statistically insignificant and stems from the lower ownership of air conditioning units, washers, and dryers outweighing the higher ownership of dishwashers, computers, and cell phones among the geographic-only poor. Yet, after controlling for average monthly temperatures and the proportion of single-family homes by county, we again find that the geographic-only poor have more total appliances and higher ownership of air conditioning units than the non-geographic-only poor (although both estimates are statistically insignificant).

Using the CIPM, we also find that the geographic-only poor have $26,070 (384%) more net worth than the non-geographic-only poor. However, this estimate for net worth is statistically insignificant due once more to the offsetting effects of the geographic-only poor having $61,530 (211%) more total wealth and $35,450 (159%) more total debt. Breaking down by the components of total wealth, we find that the geographic-only poor have substantially higher amounts for non-home or vehicle assets than the non-geographic-only poor. Next, the geographic-only poor have 0.211 (22%) fewer total food security problems than the non-geographic-only poor, with this estimate being statistically significant at the 10% level. Six of the seven individual food security problems are associated with lower deprivation after geographic adjustments, although the only statistically significant estimates are those corresponding to cutting/skipping meals and eating less than one should. Finally, the geographic-only poor have 0.069 (5%) fewer public services problems than the non-geographic-only poor, although this estimate is statistically insignificant. Six of the eleven individual public services problems are associated with lower deprivation after geographic adjustments, although only the estimate for inadequate public transportation is statistically significant.

Putting these results together, we again show that incorporating a geographic adjustment to poverty thresholds – this time using the CIPM rather than the SPM – continues to identify a less deprived poor population than those who are otherwise poor without geographic adjustments. This result holds for 55 of the 70 total well-being indicators that we analyze (with estimates being statistically significant for 24 of these outcomes). It is noteworthy that the results using the SPM and CIPM are similar despite the differences between the poverty measures.

2.7 Extensions and Robustness Checks

In this section, we discuss a series of extensions and robustness checks to our main results. We begin by showing the effects of scaling the geographic adjustment factor. We then show results using Regional Price Parities (rather than the Median Rent Index) as the price index. We also show the effects of geographic adjustments (using the original Median Rent Index) on the material well-being of those at other income cutoffs (corresponding to deep and near poverty). Furthermore, we calculate estimates based upon the Official Poverty Measure (OPM) to complement existing estimates using the SPM and CIPM. Finally, we discuss analyses that control for a more parsimonious set of covariates as well as control for either rural status or geographic region.

2.7.1 Scaling the Geographic Adjustment Factor

We first examine the effects of scaling the geographic adjustment factor by fractions from 0.1 to 1, where 0 corresponds to no adjustment and 1 corresponds to the full geographic adjustment underlying our main results. If areas with higher median rents also have higher amenities, then a full adjustment for geographic differences in rents will over-adjust thresholds for geographic differences in well-being more broadly. Because there is no commonly accepted methodology for moderating the geographic adjustment factor to account for amenities, we explore a range of weights. Renwick (2018) provides the closest analog to our analysis, although she only uses a scaling of 0.5 to reduce the geographic adjustment.

More formally, recall that the original geographic adjustment factor for the SPM threshold in equation (2.4) can be written as follows:

$$Adjustment\ Factor_{t,sm} = (Housing\ Share_t \times MRI_{sm}) + (1 - Housing\ Share_t), \qquad (2.6)$$

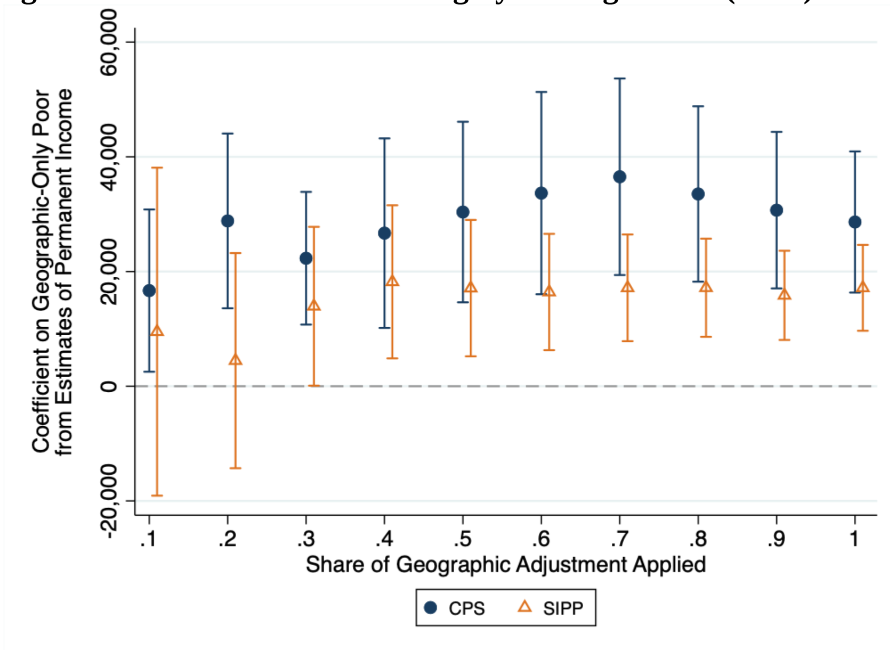
where t is the unit's housing tenure and s and m denote the unit's state and MSA (respectively). In contrast, the scaled adjustment factor for the SPM threshold can be written as:

$$\begin{aligned} Scaled\ Adjustment\ Factor_{t,sm} = {} & Fraction \times (Housing\ Share_t \times MRI_{sm}) \\ & + (1 - Fraction \times Housing\ Share_t), \end{aligned}$$

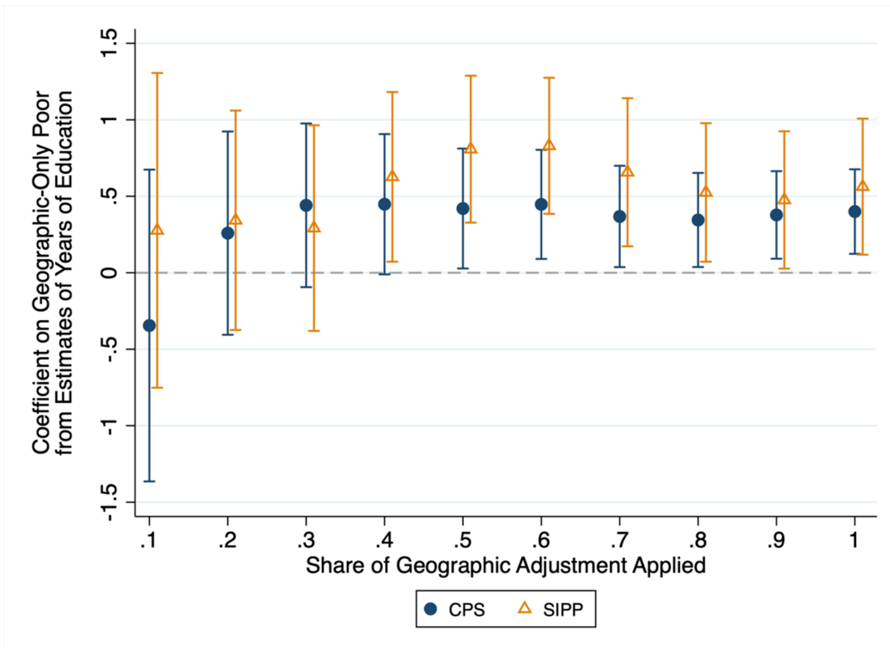
where *Fraction* ranges from 0.1 to 1 (in tenths).[95] In other words, we simply scale the full geographic adjustment factor towards 1. For example, suppose the full adjustment factors are 1.5 and 0.6 for two different observations. Applying a 0.5 scale factor changes the adjustment factors to1.25 and 0.8. We then multiply these scaled geographic adjustment factors by the base threshold and equivalence scale to obtain revised poverty thresholds corresponding to different scalings of the geographic adjustment. We continue to always anchor poverty rates at 15.1%. An implication

[95] For the CIPM, the housing share of expenditures no longer varies by housing tenure.

Figure 2.3. Regression Estimates of Well-Being by Scaling Factor (SPM)



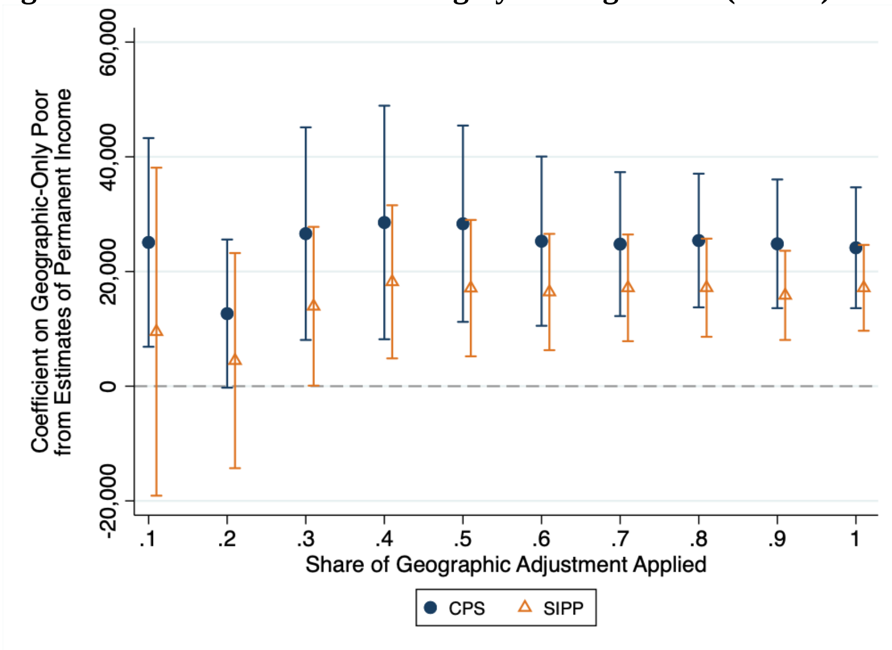
(a) Permanent Income



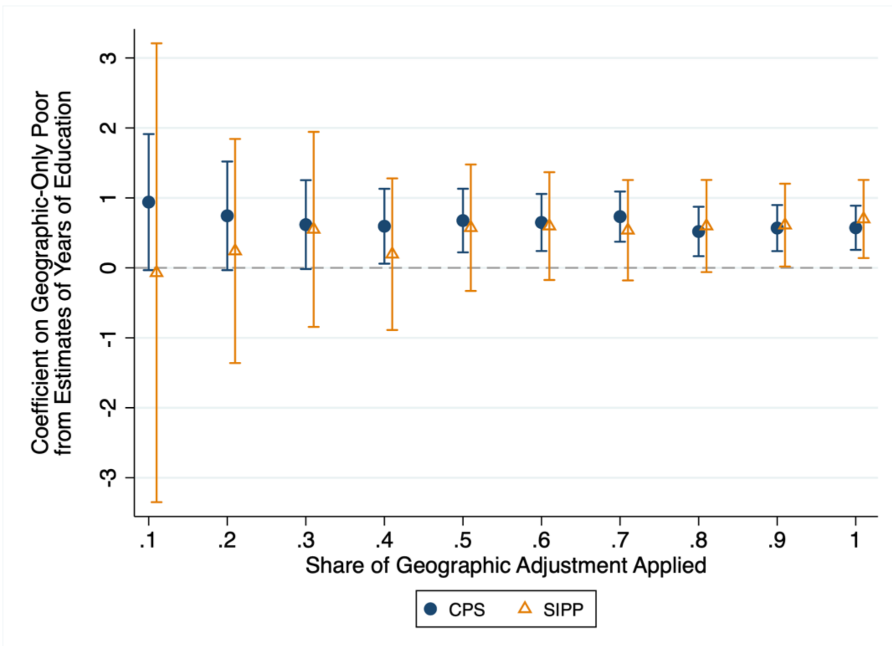
(b) Years of Education

Data: 2011 CPS ASEC and 2008 SIPP Panel linked to IRS Forms 1040/W-2/1099-R
Notes: These figures show the coefficients on an indicator for being geographic-only SPM poor (vs. non-geographic-only SPM poor) for regressions of permanent income and years of education on geographic poverty categories and covariates that vary the weight placed on the geographic adjustment factor. Sample consists of all heads in PIKed sharing units for permanent income and all sharing unit heads for years of education. Robust standard errors are in parentheses and are calculated using replicate weights. Confidence bands are at the 95% level. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization number CBDRB-FY20-ERD002-020.

Figure 2.4. Regression Estimates of Well-Being by Scaling Factor (CIPM)



(a) Permanent Income



(b) Years of Education

Data: 2011 CPS ASEC and 2008 SIPP Panel linked to various administrative data
Notes: These figures show the coefficients on an indicator for being geographic-only CIPM poor (vs. non-geographic-only CIPM poor) for regressions of permanent income and years of education on geographic poverty categories and covariates that vary the weight placed on the geographic adjustment factor. Sample consists of all heads in PIKed sharing units (and no whole imputes in the CPS). Robust standard errors are in parentheses and are calculated using replicate weights. Confidence bands are at the 95% level. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization number CBDRB-FY20-ERD002-020.

of a low fraction is that few observations will switch in or out of poverty as a result of applying a geographic adjustment. We therefore focus on permanent income and years of education in these analyses, as these outcomes – in addition to being available in both the CPS and SIPP – have substantial variation and thus allow for greater statistical power.

Figures 2.3a and 2.3b show regression estimates of permanent income and years of education – focusing on the coefficient for being geographic-only SPM poor – when varying the fraction of the MRI variation applied to the SPM poverty threshold. For each outcome, the solid circles reflect point estimates from the CPS and the hollow triangles reflect point estimates from the SIPP (surrounded by 95% confidence bands). Note that the confidence bands are wider (i.e., standard errors are larger) at lower fractions, where fewer observations are classified as geographic-only and non-geographic-only poor. A key result in these figures is that the point estimates for permanent income and education in both surveys are always positive across the entire distribution of scaling factors – with the exception of the statistically insignificant estimate for years of education in the CPS after applying 10% of the geographic adjustment. This result suggests that even a partial adjustment for geographic differences in rental costs is likely to identify a poor population that is less deprived. The point estimates are in a fairly tight range for fractions between 0.5 and 1.0 – for permanent income, the coefficients on being geographic-only poor range from \$28,630 to \$36,520 in the CPS and from \$15,840 to \$17,170 in the SIPP; for years of education, the coefficients range from 0.345 to 0.447 in the CPS and from 0.476 to 0.830 in the SIPP.

Figures 2.4a and 2.4b show the analogs of Figures 2.3a and 2.3b using the CIPM in place of the SPM. The patterns are similar. The point estimates for permanent income and education in both surveys are positive for every scaling factor, with the exception of the statistically insignificant estimate for years of education in the SIPP after applying 10% of the geographic adjustment. In addition, the point estimates are again in a fairly tight range for fractions between 0.5 and 1.0 – for permanent income, the coefficients on being geographic-only poor range from $24,140 to $28,330 in the CPS and from $14,600 to $18,420 in the SIPP; for years of education, the coefficients range from 0.52 to 0.732 in the CPS and from 0.538 to 0.698 in the SIPP. Thus, these figures show that partial and full geographic adjustments for housing costs each continue to identify a poor population that is better-off, regardless of the poverty measure or survey analyzed.

2.7.2 Using Regional Price Parities Rather Than the MRI Price Index

We next assess whether a geographic adjustment to poverty thresholds continues to identify a less deprived poor population if we use an alternative geographic price index. Specifically, in place of the MRI, we use Regional Price Parities (RPPs) that reflect the variation in prices across a broad set of goods covering housing, transportation, food, education, recreation, medical, apparel, and other items. We obtain RPP values from the BEA for calendar year 2010 by metropolitan area and – for areas that do not fall into a specified metropolitan area – for all other metropolitan areas and non-metropolitan areas within a state. We then assign the correct RPP to each individual based on his or her survey-identified place of living, first trying to match on specific metropolitan area before matching on state/metropolitan status. These RPPs are defined such that the national average is 1, so the RPP-adjusted poverty threshold is simply the base threshold multiplied by the equivalence scale and the RPP. In other words, we compute and anchor poverty in the same way as for the MRI, except we do not need to multiply by an expenditure share since the RPPs themselves are the full geographic adjustment factors.

Table 2.4. Shares and Counts by Geographic Poverty Category (using RPP Adjustments)

Poverty Category	CPS Weighted Share of Individuals (1)	CPS Sample # of Individuals (2)	CPS Sample # of Sharing Units (3)	SIPP Weighted Share of Individuals (4)	SIPP Sample # of Individuals (5)	SIPP Sample # of Sharing Units (6)
			A. Supplemental Poverty Measure (SPM)			
Never Poor	0.8368	174,000	65,000	0.8333	73,500	29,000
Non-Geog.-Only Poor	0.0122	2,400	1,000	0.0157	1,500	650
Geog.-Only Poor	0.0122	2,600	1,000	0.0157	1,300	450
Always Poor	0.1388	26,500	11,500	0.1353	12,000	5,600
			B, Comprehensive Income Poverty Measure (CIPM)			
Never Poor	0.8345	144,000	54,000	0.8338	71,000	28,000
Non-Geog.-Only Poor	0.0145	2,400	950	0.0151	1,500	600
Geog.-Only Poor	0.0145	2,500	900	0.0151	1,200	450
Always Poor	0.1365	21,500	8,800	0.1359	11,000	5,000

Data: 2011 CPS ASEC and 2008 SIPP Panel linked to various administrative data

Notes: This table shows the number of individuals (both weighted share and unweighted count) and unweighted number of sharing units in each of our four geographic poverty categories for the SPM and CIPM in both the CPS and SIPP, when using Regional Price Parities (RPPs) rather than the Median Rent Index (MRI) to adjust poverty thresholds for geographic variation in cost-of-living. Poverty rates are always anchored to 15.1%, which is the official rate in the CPS. The sample for the SPM estimates consists of all observations in the surveys, and estimates are weighted using original survey weights in Columns (1) and (4). The sample for the CIPM estimates in the CPS consists of all individuals in sharing units where at least one member has a PIK and no member is whole imputed, and estimates are weighted using individual survey weights adjusted for non-PIKing and whole imputes (at the sharing unit level). The sample for the CIPM estimates in the SIPP consists of all individuals in sharing units where at least one member has a PIK, and estimates are weighted using individual survey weights adjusted for non-PIKing (at the sharing unit level). The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization number CBDRB-FY20-ERD002-020.

Table 2.4 shows the weighted shares and un-weighted counts of observations falling into each of the geographic poverty categories defined using the RPP adjustment. For the SPM, note that the shares of individuals falling into the geographic-only and non-geographic-only poor groups under the RPP adjustment (1.22% each in the CPS and 1.57% each in the SIPP) are slightly lower than the corresponding shares under the original MRI adjustment (1.52% in the CPS and 1.83% in the SIPP). Conversely, the shares of individuals in each of the "switcher" groups for the CIPM under the RPP adjustment (1.45% in the CPS and 1.51 in the SIPP) are comparable to the corresponding shares under the MRI adjustment (1.43% in the CPS and 1.46 in the SIPP). Similar to the MRI, we find that the RPP adjustment decreases poverty rates in rural areas (particularly those in the Midwest and Deep South) and increases poverty rates in urban areas (particularly in New England, the mid-Atlantic, and California). While the MRI and RPP adjustments move poverty rates in the same direction in the vast majority of the country, there are also a few areas where they diverge.[96]

Table 2.5a. Regression Estimates of Well-Being on Geographic-Only Poor Using RPP Adjustments (SPM)

Well-Being Measures	Point Estimate	Standard Error	Mean for Non-Geog Poor	Overall Mean	Supports Geog Adj? (+/-)
	(1)	(2)	(3)	(4)	(5)
Permanent Income (*CPS*)	26,230***	(7,985)	27,380	98,090	–
Permanent Income (*SIPP*)	17,310***	(4,074)	34,050	92,930	–
Years of Education (*CPS*)	0.4720***	(0.1650)	11.900	13.640	–
Years of Education (*SIPP*)	0.6370***	(0.2360)	12.090	13.700	–
Head Died by 2019 (*CPS*)	-0.0266	(0.0164)	0.130	0.071	–
Head Died by 2019 (*SIPP*)	-0.0229	(0.0185)	0.121	0.080	–
Ind. Has Poor/Fair Health (*CPS*)	-0.0558***	(0.0153)	0.230	0.118	–
Ind. Has Poor/Fair Health (*SIPP*)	-0.0729***	(0.0149)	0.189	0.103	–
# of Material Hardships (*SIPP*)	-0.2410*	(0.1320)	1.157	0.646	–
# of Home Problems (*SIPP*)	-0.0908	(0.0641)	0.349	0.224	–
# of Appliances (*SIPP*)	-0.3240**	(0.1460)	6.271	6.988	+
Total Wealth (*SIPP*)	97,360**	(46,300)	74,900	384,900	–
# of Food Security Problems (*SIPP*)	-0.2430*	(0.1260)	0.992	0.460	–
# of Public Service Problems (*SIPP*)	-0.0078	(0.1430)	--	0.949	–

*** p<0.01, ** p<0.05, * p<0.1

Data: 2011 CPS ASEC and 2008 SIPP Panel linked to various administrative data

Notes: This table shows the coefficient on an indicator for being geographic-only poor (vs. non-geographic-only poor) for regressions of a wide variety of well-being measures on indicators for being in one of three geographic SPM poverty categories (omitting the non-geographic-only poor) and a vector of covariates (for the sharing unit or its head) including age, gender, marital status, unit type, and race/ethnicity. These estimates use Regional Price Parities (RPPs) rather than the Median Rent Index (MRI) to adjust poverty thresholds for geographic variation in cost-of-living. Sample consists of all sharing unit heads for most outcomes, except for some of the mortality and health outcomes (which are at the individual level) and mortality and permanent income outcomes (where we restrict to PIKed units in both surveys and non-whole-imputed units in CPS). For most outcomes, sample sizes are 78,500 and 36,000 in the CPS and SIPP, respectively; for permanent income, it is 72,500 and 34,000 in the CPS and SIPP; for head mortality, it is 71,500 and 33,000 in the CPS and SIPP; for individual heath, it is 88,000 in the SIPP; for assets, it is 35,500 in the SIPP. Probit APEs are reported for binary outcomes. Robust standard errors are in parentheses and are calculated using replicate weights. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization numbers CBDRB-FY20-ERD002-020 and CBDRB-FY2021-CES005-016.

[96] Specifically, using public-use data, we can determine that the RPP adjustment increases poverty while the MRI adjustment decreases poverty in more remote areas (e.g., rural California and Indiana) where housing costs are lower but the costs of other goods are higher. In contrast, the RPP adjustment decreases poverty while the MRI adjustment increases poverty primarily in coastal areas (e.g., parts of Florida, South Carolina, Delaware, etc.) where housing costs are higher but the costs of other goods are lower.

Table 2.5a shows regression estimates of selected well-being indicators on SPM poverty categories using the RPP adjustment, focusing on the coefficient for being geographic-only poor. We examine fourteen summary outcomes, encompassing ten unique measures (one from each domain of outcomes for each survey). Four of these measures are available in both the CPS and SIPP, and six of these measures are available only in the SIPP. For thirteen of the fourteen measures, the sign of the regression coefficient suggests that incorporating a geographic RPP adjustment identifies a less deprived population (with nine of these estimates statistically significant at the 10% level). The only well-being indicator for which the RPP adjustment identifies a more deprived population is the number of appliances, but – as we previously discussed – this effect is likely driven in part by the location-specific needs of the non-geographic-only poor. The magnitudes of the regression estimates using the RPP adjustment are strikingly similar to those of the regression estimates using the MRI adjustment in Table 2.3a, although there are some slight differences with regard to statistical significance. In particular, the estimates for mortality (in the CPS), the number of hardships, and the number of appliances are statistically significant under the RPP adjustment (but not the MRI adjustment), while the estimate for the number of home quality problems is statistically significant under the MRI adjustment (but not the RPP adjustment).

Table 2.5b shows the analog of Table 2.5a using the Comprehensive Income Poverty Measure (CIPM) instead of the SPM. Once again, these estimates suggest that geographic adjustments to poverty using the RPP lead to a less deprived population in poverty for the same thirteen outcomes as in Table 2.5a (with the estimates statistically significant for ten of these outcomes). The estimates in Table 2.5b are again comparable to their counterparts using the MRI

Table 2.5b. Regression Estimates of Well-Being on Geographic-Only Poor Using RPP Adjustments (CIPM)

Well-Being Measures	Point Estimate (1)	Standard Error (2)	Mean for Non-Geog Poor (3)	Overall Mean (4)	Supports Geog Adj? (+/-) (5)
Permanent Income (*CPS*)	22,160***	(5,632)	24,860	98,090	–
Permanent Income (*SIPP*)	15,030***	(3,426)	24,990	92,930	–
Years of Education (*CPS*)	0.7800***	(0.1430)	11.730	13.630	–
Years of Education (*SIPP*)	0.4950*	(0.2930)	11.930	13.700	–
Head Died by 2019 (*CPS*)	-0.0197*	(0.0112)	0.100	0.071	–
Head Died by 2019 (*SIPP*)	-0.0085	(0.0188)	0.097	0.080	–
Ind. Has Poor/Fair Health (*CPS*)	-0.0648***	(0.0149)	--	--	–
Ind. Has Poor/Fair Health (*SIPP*)	-0.0361**	(0.0151)	0.182	0.104	–
# of Material Hardships (*SIPP*)	-0.2930**	(0.1390)	1.270	0.645	–
# of Home Problems (*SIPP*)	-0.1090	(0.0718)	0.401	0.225	–
# of Appliances (*SIPP*)	-0.3590**	(0.1530)	6.256	6.992	+
Total Wealth (*SIPP*)	58,590***	(21,630)	26,340	386,300	–
# of Food Security Problems (*SIPP*)	-0.2910**	(0.1130)	0.975	0.459	–
# of Public Service Problems (*SIPP*)	-0.0810	(0.1540)	--	0.950	–

*** p<0.01, ** p<0.05, * p<0.1

Data: 2011 CPS ASEC and 2008 SIPP Panel linked to various administrative data

Notes: This table shows the coefficient on an indicator for being geographic-only poor (vs. non-geographic-only poor) for regressions of a wide variety of well-being measures on indicators for being in one of three geographic CIPM poverty categories (omitting the non-geographic-only poor) and a vector of covariates (for the sharing unit or its head) including age, gender, marital status, unit type, and race/ethnicity. These estimates use Regional Price Parities (RPPs) rather than the Median Rent Index (MRI) to adjust poverty thresholds for geographic variation in cost-of-living. Sample consists of all heads in PIKed sharing units (and no whole imputes in the CPS) for most outcomes, except for some of the mortality and health outcomes (which are at the individual level). For most outcomes, sample sizes are 64,500 and 34,000 in the CPS and SIPP, respectively; for head mortality, it is 63,500 and 33,000 in the CPS and SIPP; for individual heath, it is 85,000 in the SIPP; for assets, it is 33,500 in the SIPP. Probit APEs are reported for binary outcomes. Robust standard errors are in parentheses and are calculated using replicate weights. Results have been approved for release by the U.S. Census Bureau, authorization numbers CBDRB-FY20-ERD002-020 and CBDRB-FY2021-CES005-016.

in Table 2.3b, although the estimates for the numbers of hardships and appliances are statistically significant under the RPP adjustment (but not the MRI adjustment) while the estimate for the number of home quality problems is statistically significant under the MRI adjustment (but not the RPP adjustment). In sum, these results show that adjusting for price differences across a broader bundle of goods (beyond housing) does not change our central finding that a geographic adjustment to poverty identifies a less deprived poor population. This finding holds across the SPM and CIPM, the vast majority of well-being indicators, and the CPS and SIPP. Moreover, the consistency of the results using the RPP and MRI adjustments suggests that using an intermediate geographic adjustment index like the Food, Apparel, and Rent RPP (which covers a set of goods strictly between those covered in the MRI and RPP) is likely to yield similar results.

2.7.3 Deep and Near Poverty

In another set of analyses, we examine the effects of a geographic adjustment to poverty thresholds (using the MRI) on the deprivation of those classified as deep poor and near poor. Deep poverty is defined as having incomes below 50% of the poverty line, while near poverty is defined as having incomes below 150% of the poverty line. To calculate deep and near poverty with and without a geographic adjustment, we follow the same methodology as that used for regular poverty – with the only difference being that we anchor deep poverty rates to 6.7% and near poverty rates to 24.6%. These rates are based on the deep and near poverty rates calculated in the CPS (using survey-reported pre-tax money income and OPM thresholds) for reference year 2010. Table 2.6 shows the weighted shares and un-weighted counts of observations falling into each of the geographic deep and near poverty categories. The shares of individuals in each of the "switcher" groups for deep poverty range between 0.49% and 0.51% for the SPM and between 0.64% and 0.66% for the CIPM (depending on the survey analyzed). On the other hand, the shares of individuals in each of the "switcher" groups for near poverty range between 2.46% and 2.65% for the SPM and between 1.84% and 2.10% for the CIPM.

Table 2.7a presents regression estimates of selected well-being indicators – the same ones as those analyzed in Tables 2.5a and 2.5b – on deep and near poverty categories under the SPM, focusing on the coefficient for being geographic-only poor. Examining first the deep poverty estimates in Panel A, we find that a geographic adjustment identifies a less deprived population in

Table 2.6. Shares and Counts by Geographic Poverty Category (Deep and Near Poverty)

Poverty Category	CPS Weighted Share of Individuals (1)	CPS Sample # of Individuals (2)	CPS Sample # of Sharing Units (3)	SIPP Weighted Share of Individuals (4)	SIPP Sample # of Individuals (5)	SIPP Sample # of Sharing Units (6)
DEEP POVERTY						
A. Supplemental Poverty Measure (SPM)						
Never Deep Poor	0.9279	191,000	72,500	0.9280	82,000	32,500
Non-Geog-Only Deep Poor	0.0051	1,000	400	0.0049	500	200
Geog-Only Deep Poor	0.0051	1,000	400	0.0050	400	150
Always Deep Poor	0.0619	11,500	5,500	0.0621	5,500	2,800
B. Comprehensive Income Poverty Measure (CIPM)						
Never Deep Poor	0.9266	159,000	59,500	0.9265	79,000	31,500
Non-Geog-Only Deep Poor	0.0064	1,000	450	0.0065	600	250
Geog-Only Deep Poor	0.0064	1,000	350	0.0066	500	200
Always Deep Poor	0.0606	8,900	4,000	0.0604	4,800	2,200
NEAR POVERTY						
C. Supplemental Poverty Measure (SPM)						
Never Near Poor	0.7294	151,000	56,500	0.7275	63,500	25,500
Non-Geog-Only Near Poor	0.0246	5,300	2,000	0.0265	2,600	1,000
Geog-Only Near Poor	0.0246	5,300	1,900	0.0265	2,200	750
Always Near Poor	0.2214	43,500	18,000	0.2195	19,500	8,600
D. Comprehensive Income Poverty Measure (CIPM)						
Never Near Poor	0.7330	126,000	47,500	0.7356	62,500	25,000
Non-Geog-Only Near Poor	0.0210	4,000	1,400	0.0185	1,800	650
Geog-Only Near Poor	0.0210	3,700	1,300	0.0184	1,500	500
Always Near Poor	0.2250	36,500	14,000	0.2275	19,500	7,800

Data: 2011 CPS ASEC and 2008 SIPP Panel linked to various administrative data

Notes: This table shows the number of individuals (both weighted share and unweighted count) and unweighted number of sharing units in each of our four geographic deep and near poverty categories for the SPM and CIPM in both the CPS and SIPP. Rates for deep poverty (i.e., having incomes below 50% of the poverty line) are anchored to 6.7%, and rates for near poverty (i.e., having incomes below 150% of the poverty line) are anchored to 24.6%. Both rates correspond to what we obtain using pre-tax money income in the CPS for reference year 2010. The sample for the SPM estimates consists of all observations in the surveys, and estimates are weighted using original survey weights in Columns (1) and (4). The sample for the CIPM estimates in the CPS consists of all individuals in sharing units where at least one member has a PIK and no member is whole imputed, and estimates are weighted using individual survey weights adjusted for non-PIKing and whole imputes (at the sharing unit level). The sample for the CIPM estimates in the SIPP consists of all individuals in sharing units where at least one member has a PIK, and estimates are weighted using individual survey weights adjusted for non-PIKing (at the sharing unit level). The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization number CBDRB-FY20-ERD002-020.

Table 2.7a. Regression Estimates of Well-Being on Geographic-Only Deep & Near Poor (SPM)

Well-Being Measures	Point Estimate (1)	Standard Error (2)	Mean for Non-Geog Poor (3)	Overall Mean (4)	Supports Geog Adj? (+/-) (5)
			A. Deep Poverty		
Permanent Income (*CPS*)	19,270***	(6,748)	27,840	98,620	–
Permanent Income (*SIPP*)	7,228	(7,340)	29,770	92,930	–
Years of Education (*CPS*)	1.2130***	(0.2600)	11.420	13.640	–
Years of Education (*SIPP*)	0.7400**	(0.3700)	11.530	13.700	–
Head Died by 2019 (*CPS*)	0.0076	(0.0294)	0.108	0.071	+
Head Died by 2019 (*SIPP*)	-0.0058	(0.0399)	0.131	0.080	–
Ind. Has Poor/Fair Health (*CPS*)	-0.0527**	(0.0212)	0.218	0.118	–
Ind. Has Poor/Fair Health (*SIPP*)	-0.0793***	(0.0265)	0.194	0.103	–
# of Material Hardships (*SIPP*)	-0.4430*	(0.2330)	1.528	0.646	–
# of Home Problems (*SIPP*)	0.0338	(0.1440)	0.475	0.224	+
# of Appliances (*SIPP*)	-0.3480*	(0.2050)	6.053	6.988	+
Total Wealth (*SIPP*)	27,860	(48,670)	92,390	384,900	–
# of Food Security Problems (*SIPP*)	-0.3270	(0.2350)	1.027	0.460	–
# of Public Service Problems (*SIPP*)	-0.2920	(0.2230)	--	0.949	–
			B. Near Poverty		
Permanent Income (*CPS*)	20,610***	(3,398)	37,170	98,620	–
Permanent Income (*SIPP*)	10,880*	(6,327)	37,990	92,930	–
Years of Education (*CPS*)	0.3530***	(0.1080)	12.420	13.630	–
Years of Education (*SIPP*)	0.6350***	(0.1740)	12.450	13.700	–
Head Died by 2019 (*CPS*)	-0.0253***	(0.0092)	0.108	0.071	–
Head Died by 2019 (*SIPP*)	-0.0124	(0.0150)	0.098	0.080	–
Ind. Has Poor/Fair Health (*CPS*)	-0.0384***	(0.0098)	0.154	0.118	–
Ind. Has Poor/Fair Health (*SIPP*)	-0.0481***	(0.0123)	0.163	0.103	–
# of Material Hardships (*SIPP*)	-0.1710*	(0.1010)	1.000	0.646	–
# of Home Problems (*SIPP*)	-0.0412	(0.0389)	0.268	0.224	–
# of Appliances (*SIPP*)	-0.1950*	(0.1060)	6.526	6.988	+
Total Wealth (*SIPP*)	100,400***	(25,630)	85,080	384,900	–
# of Food Security Problems (*SIPP*)	-0.0401	(0.0732)	0.611	0.460	–
# of Public Service Problems (*SIPP*)	-0.0114	(0.1200)	--	0.949	–

*** p<0.01, ** p<0.05, * p<0.1

Data: 2011 CPS ASEC and 2008 SIPP Panel linked to IRS Forms 1040/W-2/1099-R and SSA Numident

Notes: This table shows the coefficient on an indicator for being geographic-only deep/near poor (vs. non-geographic-only deep/near poor) for regressions of a wide variety of well-being measures on indicators for being in one of three geographic SPM deep/near poverty categories (omitting the non-geographic-only deep/near poor) and a vector of covariates (for the sharing unit or its head) including age, gender, marital status, unit type, and race/ethnicity. Sample consists of all sharing unit heads for most outcomes, except for some of the mortality and health outcomes (which are at the individual level) and mortality and permanent income outcomes (where we restrict to PIKed units in both surveys and non-whole-imputed units in CPS). For most outcomes, sample sizes are 78,500 and 36,000 in the CPS and SIPP, respectively; for permanent income, it is 72,500 and 34,000 in the CPS and SIPP; for head mortality, it is 71,500 and 33,000 in the CPS and SIPP; for individual heath, it is 88,000 in the SIPP; for assets, it is 35,500 in the SIPP. Probit APEs are reported for binary outcomes. Robust standard errors are in parentheses and are calculated using replicate weights. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization numbers CBDRB-FY20-ERD002-020 and CBDRB-FY2021-CES005-016.

Table 2.7b. Regression Estimates of Well-Being on Geographic-Only Deep & Near Poor (CIPM)

Well-Being Measures	Point Estimate (1)	Standard Error (2)	Mean for Non-Geog Poor (3)	Overall Mean (4)	Supports Geog Adj? (+/-) (5)
			A. Deep Poverty		
Permanent Income (*CPS*)	22,260***	(6,690)	19,050	98,090	–
Permanent Income (*SIPP*)	21,230***	(6,977)	17,080	92,930	–
Years of Education (*CPS*)	0.4280*	(0.2490)	11.700	13.630	–
Years of Education (*SIPP*)	0.3410	(0.3760)	11.560	13.700	–
Head Died by 2019 (*CPS*)	-0.0491**	(0.0210)	0.123	0.071	–
Head Died by 2019 (*SIPP*)	-0.0912***	(0.0325)	0.163	0.080	–
Ind. Has Poor/Fair Health (*CPS*)	-0.0851***	(0.0229)	--	--	–
Ind. Has Poor/Fair Health (*SIPP*)	-0.1010***	(0.0224)	0.211	0.104	–
# of Material Hardships (*SIPP*)	0.0152	(0.2390)	1.393	0.645	+
# of Home Problems (*SIPP*)	-0.0610	(0.1090)	0.433	0.225	–
# of Appliances (*SIPP*)	0.1630	(0.2850)	5.709	6.992	–
Total Wealth (*SIPP*)	41,220	(33,480)	22,160	386,300	–
# of Food Security Problems (*SIPP*)	-0.0473	(0.2610)	1.152	0.459	–
# of Public Service Problems (*SIPP*)	-0.00463	(0.234)	--	0.950	–
			B. Near Poverty		
Permanent Income (*CPS*)	14,220***	(2,792)	33,340	98,090	–
Permanent Income (*SIPP*)	15,220***	(2,709)	35,940	92,930	–
Years of Education (*CPS*)	0.3830**	(0.1550)	12.390	13.630	–
Years of Education (*SIPP*)	0.2970	(0.2180)	12.520	13.700	–
Head Died by 2019 (*CPS*)	-0.0014	(0.0106)	0.088	0.071	–
Head Died by 2019 (*SIPP*)	-0.0164	(0.0138)	0.083	0.080	–
Ind. Has Poor/Fair Health (*CPS*)	-0.0164	(0.0112)	--	--	–
Ind. Has Poor/Fair Health (*SIPP*)	-0.0317**	(0.0121)	0.139	0.104	–
# of Material Hardships (*SIPP*)	-0.1930	(0.1180)	1.084	0.645	–
# of Home Problems (*SIPP*)	-0.0822	(0.0592)	0.326	0.225	–
# of Appliances (*SIPP*)	-0.3360***	(0.1050)	6.599	6.992	+
Total Wealth (*SIPP*)	77,120***	(23,380)	48,800	386,300	–
# of Food Security Problems (*SIPP*)	0.0656	(0.1230)	0.719	0.459	+
# of Public Service Problems (*SIPP*)	0.0546	(0.1400)	--	0.950	+

*** p<0.01, ** p<0.05, * p<0.1

Data: 2011 CPS ASEC and 2008 SIPP Panel linked to various administrative data

Notes: This table shows the coefficient on an indicator for being geographic-only deep/near poor (vs. non-geographic-only deep/near poor) for regressions of a wide variety of well-being measures on indicators for being in one of three geographic CIPM deep/near poverty categories (omitting the non-geographic-only deep/near poor) and a vector of covariates (for the sharing unit or its head) including age, gender, marital status, unit type, and race/ethnicity. Sample consists of all heads in PIKed sharing units (and no whole imputes in the CPS) for most outcomes, except for some of the mortality and health outcomes (which are at the individual level). For most outcomes, sample sizes are 64,500 and 34,000 in the CPS and SIPP, respectively; for head mortality, it is 63,500 and 33,000 in the CPS and SIPP; for individual heath, it is 85,000 in the SIPP; for assets, it is 33,500 in the SIPP. Probit APEs are reported for binary outcomes. Robust standard errors are in parentheses and are calculated using replicate weights. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization numbers CBDRB-FY20-ERD002-020 and CBDRB-FY2021-CES005-016.

deep poverty for eleven out of the fourteen outcomes (with the estimates being statistically significant at the 10% level for only six of these eleven outcomes). For the three outcomes for which a geographic adjustment identifies a more deprived population in deep poverty, only one (the number of appliances) is associated with a statistically significant difference at the 10% level. Many of the estimates for deep poverty are statistically imprecise because there are relatively few individuals switching in and out of deep poverty with a geographic adjustment.

Moving onto the near poverty estimates in Panel B of Table 2.7a, we find that a geographic adjustment identifies a less deprived population in near poverty for thirteen of the fourteen outcomes (with the estimates being statistically significant at the 10% level for nine outcomes). Again, the only well-being indicator for which a geographic adjustment identifies a more deprived population in near poverty is the number of appliances. For several outcomes (including permanent income, years of education, and material hardships), we find that the non-geographic-only near poor appear to be more well-off than the non-geographic-only deep poor (Column 3). This makes sense if incomes are negatively correlated with deprivation. However, there are also a few measures (namely, mortality in the CPS and total wealth) on which the non-geographic-only near poor do not appear better off than the non-geographic deep poor. This suggests that the SPM may misclassify individuals who are advantaged (e.g., who underreport incomes in the survey, have substantial assets not captured in the SPM resource measure, etc.) as being in deep poverty.

We therefore also analyze the regression estimates of selected well-being indicators on deep and near poverty categories under the CIPM, again focusing on the coefficient for being geographic-only poor (Table 2.7b). Concentrating first on deep poverty in Panel A, we find using the CIPM that geographic adjustments identify a less deprived population in deep poverty for thirteen out of the fourteen outcomes (with the estimates being statistically significant at the 10%

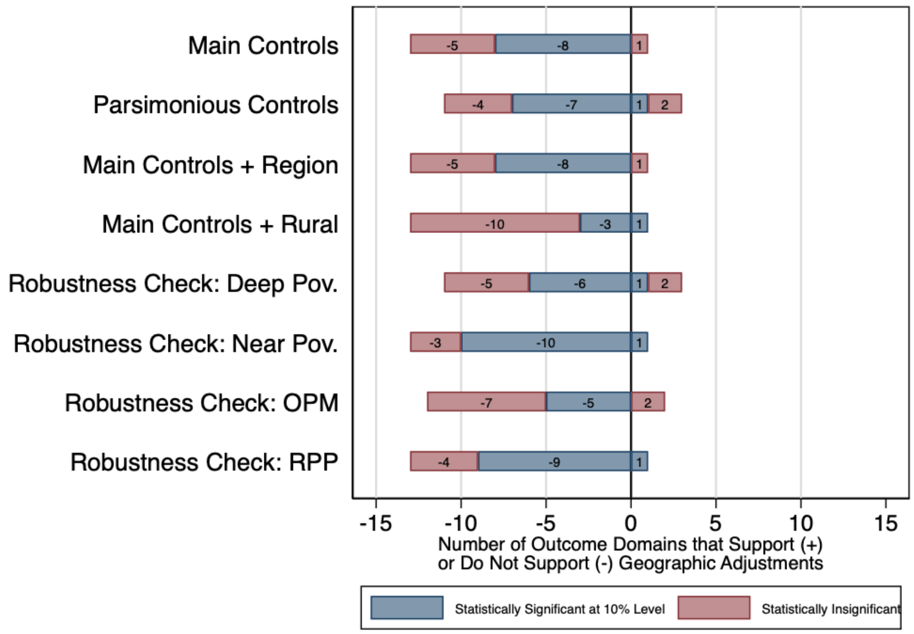
level for seven of these thirteen outcomes). Moreover, the only outcome (i.e., material hardships) for which the geographic-only poor appear more deprived is associated with a statistically insignificant estimate. Note that the point estimates for mortality are particularly notable for the CIPM, with the geographic-only deep poor being approximately 40% and 50% (in the CPS and SIPP, respectively) *less* likely to have a head die by 2019 than the non-geographic-only deep poor. The estimates are so striking potentially because mortality is a tail event and the non-geographic-only deep poor (particularly those living in the Deep South) may be especially prone to suffer from health issues resulting in higher mortality.

Turning finally to the near poverty estimates in Panel B of Table 2.7b, we find that geographic adjustments identify a less deprived population in near poverty for eleven of the fourteen outcomes (with the estimates being statistically significant at the 10% level for five of the eleven outcomes). Furthermore, the non-geographic-only near poor appear to be more well-off than the non-geographic-only deep poor on every outcome analyzed when using the CIPM (unlike the SPM). These patterns provide further evidence that the CIPM performs better than the SPM when validated against measures of well-being. Despite these differences, the findings in Tables 2.7a and 2.7b strongly and consistently show that a geographic adjustment to poverty thresholds continues to identify a less deprived population even after extending our analyses to other cutoffs.

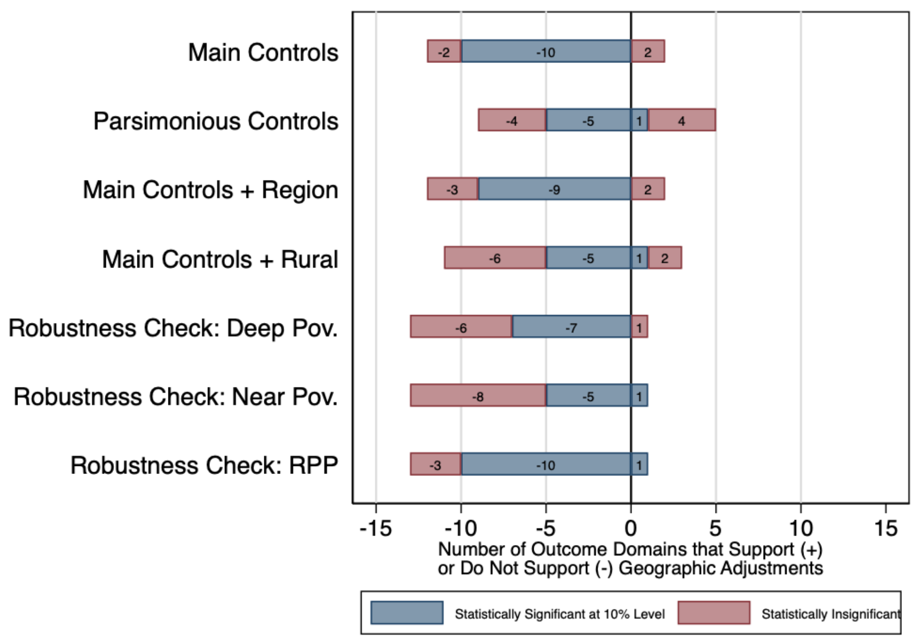
2.7.4 Additional Analyses

In this final subsection, we describe a series of additional extensions and robustness checks. Figures 2.5a and 2.5b summarize the results of these additional analyses, along with the results of the RPP and deep/near poverty analyses. Specifically, for a given model, Figure 2.5a shows the number of outcomes (out of the fourteen outcomes examined in Tables 2.5a-2.5b and 2.7a-2.7b)

Figure 2.5. Summary of Geographic Adjustment Effects on Well-Being by Model



(a) SPM



(b) CIPM

Data: 2011 CPS ASEC and 2008 SIPP Panel linked to various administrative data
Notes: These figures show the number of outcomes for which a geographic adjustment identifies a more deprived population, using either the SPM or CIPM. Outcomes are those in Tables 2.5a-2.5b and 2.7a-2.7b; outcome domains include mortality, permanent income, education, and health problems (in CPS and SIPP), and appliances, assets, food security problems, home quality problems, material hardships, and public services problems (in SIPP only). The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization numbers CBDRB-FY20-ERD002-020 and CBDRB-FY2021-CES005-016.

that support or do not support geographic adjustments when using survey-based poverty measures.[97] Figure 2.5b shows the analog of Figure 2.5a when using the CIPM.

First, we examine the effects of geographic adjustments on the well-being of those classified as poor under the Official Poverty Measure (OPM). Although the OPM has many well-known limitations, it is also the poverty measure that is conceptually most similar to what various government programs (e.g., SNAP, Head Start, etc.) use to determine benefit eligibility. Like the SPM, the OPM relies exclusively on survey-reported income. Unlike the SPM, the OPM uses pre-tax money income as its resource measure, uses the survey family (containing only related individuals) as its resource unit, and relies on poverty thresholds that vary by family size and the number of related children and do not feature geographic adjustments. To incorporate geographic adjustments, we multiply OPM thresholds by the geographic adjustment factor in equation (2.6).[98] Using the OPM, we find that geographic adjustments identify a less deprived population for twelve of the fourteen summary well-being outcomes (although estimates for only five of these twelve outcomes are statistically significant at the 10% level). Moreover, the estimates for the two summary outcomes (number of appliances and number of public services problems) that point in the opposite direction are statistically insignificant.

Next, we examine the sensitivity of our results to the covariates used in our regression specifications. First, we re-estimate our regressions using the most parsimonious set of covariates possible.[99] Upon doing so, our results continue to indicate that geographic adjustments identify a

---

[97] Outcomes that support (do not support) geographic adjustments are tabulated to the left (right) of zero, with different color shadings to distinguish outcomes that are statistically significant at the 10% level from outcomes that are statistically insignificant.

[98] This is the same adjustment factor used in the SPM and CIPM to adjust for differences in local housing rents.

[99] This entails controlling for age when examining permanent income, education, and mortality (as these outcomes are mechanically correlated with age), unit size when examining material hardships, home quality problems, appliances, food security problems, and public services (as these outcomes are asked of anyone in the household), and age and unit size when examining assets (as they are mechanically correlated with age and are not equivalized by unit size or composition). Importantly, we no longer control for gender, marital status, unit type, or race/ethnicity.

less deprived population in poverty (although the patterns are slightly weaker). Using the SPM and CIPM, we find that geographic adjustments identify a more well-off poor population for eleven and nine of the fourteen summary outcomes in Figures 2.5a and 2.5b, respectively, when controlling for as few demographics as possible (with seven and five of these estimates statistically significant at the 10% level, respectively). Years of education is a key domain for which geographic adjustments no longer identify a significantly less deprived population after reducing the number of covariates. One reason for this is that we no longer control for Hispanic status, and Hispanics tend to have fewer years of education and are over-represented among the geographic-only poor.

Finally, we analyze how our regression estimates change when adding either binary variables for geographic region (Northeast, Midwest, South, West) or a binary variable for rural status to our main set of controls. Our results barely budge after controlling for geographic region. Specifically, after controlling for geographic region, we find that geographic adjustments continue to identify a less deprived population in poverty for thirteen of the fourteen summary outcomes using the SPM (with eight significant) and twelve of the fourteen using the CIPM (with nine of these estimates statistically significant at the 10% level). In contrast, our results weaken substantially after controlling for rural status. After doing so, we find that geographic adjustments continue to identify a less deprived population in poverty for thirteen of the fourteen summary outcomes using the SPM (with only three statistically significant) and eleven using the CIPM (with five of these estimates statistically significant at the 10% level). In summary, the urban versus rural distinction between the geographic-only and non-geographic-only poor appears to be a key driver of the results we see, whereas our results remain largely unchanged whether we examine differences across geographic regions or within geographic regions.

2.8 Empirical Explanations for Results

In this section, we provide some empirical explanations for the overarching result that those who are poor in higher-cost areas appear to be less deprived than those who are poor in lower-cost areas. Intuitively, this result must arise out of the idea that higher prices in certain areas are correlated with greater provision of other things valued by the poor in those areas. We test this hypothesis by calculating the correlation of local hourly wages, non-labor income, and spending on various types of amenities with local prices. While these analyses are only meant to be exploratory, they provide further evidence that it is highly complicated – if not impossible – to perfectly calibrate how local amenities and incomes adjust with local prices.

Table 2.8a. Elasticities of Wage and Non-Wage Income with Respect to Price Indices

Outcome	Elasticity of Outcome With Respect to MRI (1)	Elasticity of Outcome With Respect to RPP (2)
Hourly Wage (per person 18-64 with HS or less)	0.874***	1.072***
Social Security Retirement Income (per person 62+)	0.160	0.199*
Social Security Disability Income (per capita)	0.296**	0.396***
Retirement Income (per person 60+)	-2.173***	-2.151***
SNAP (per capita)	1.369***	1.381***
Housing Assistance (per capita)	-2.461***	-2.972***
SSI (per capita)	3.643***	3.304***
Observations	341	341
Unit of Analysis	CBSA	CBSA

*** p<0.01, ** p<0.05, * p<0.1

Data: 2011 CPS ASEC (public-use), MRI and RPP values for 2010

Notes: This table shows the coefficients from regressions of the natural log of various income sources on the natural log of local prices calculated using either the MRI or the RPP. For wages, we use the 2011 CPS ASEC for individuals ages 18-64 with a high school degree or less and weight the average using survey weights. We calculate per capita outcomes as the weighted total of an outcome divided by the weighted population. Housing assistance is drawn from the Census Bureau's SPM Research File. Both the MRI and RPP are calculated for calendar year 2010. In Column (1), we use 0.618 + 0.382*MRI as the price index to make the results comparable where 0.382 is the housing share of consumption found using the Consumer Expenditure Survey (CE).

We first examine the elasticities of local hourly wages and non-labor income with respect to two price indices. Table 2.8a shows the coefficients from CBSA-level regressions of the natural log of various income sources on the natural log of local prices calculated using either the MRI or RPP. Average incomes are calculated using the CPS ASEC for reference year 2010, and the MRI and RPP are also calculated for 2010. To make the MRI and RPP elasticities comparable, we scale the MRI by 38.2% (which is the housing share of consumption found using the CE Survey). First, we find that hourly wages (calculated for non-elderly adults with a high school diploma or less) increase by 0.87% and 1.07% given a 1% increase in the MRI and RPP, respectively. The hourly wage elasticities for both price indices are statistically significant at the 1% level.

Turning next to various sources of non-labor income, we find that Social Security retirement income (per individual aged 62 and older) also increases by 0.30% and 0.40% given a 1% increase in the MRI and RPP, respectively, while Social Security disability income (per capita) actually decreases by 2.17% and 2.15% given a 1% increase in the MRI and RPP, respectively. Like Social Security retirement income, private pensions (per individual aged 60 and older) are also positively correlated with prices – but the elasticities are above unity and therefore much larger (ranging from 1.37 to 1.38). Focusing next on various means-tested transfers, we find a significantly positive correlation between housing assistance (per capita) and prices, with housing assistance increasing by 3.6% and 3.3% given a 1% increase in the MRI and RPP, respectively. The fact that these elasticities are much higher than one suggests that the relationship between housing assistance and median rents is not likely to be purely mechanical. There are also a couple of transfers (SNAP and SSI) whose per-capita amounts are negatively correlated with prices, although only the elasticities for SNAP are statistically significant.

Table 2.8b. Elasticities of Per-Capita State Spending with Respect to Price Indices

Outcome	Elasticity of Outcome With Respect to MRI (1)	Elasticity of Outcome With Respect to RPP (2)
Welfare	1.200**	1.256**
All Education	0.671**	0.840***
K-12 Education	1.206**	1.363***
Higher Education	-1.040	-0.852
Health and Hospitals	-0.591	-0.668
Police	1.800***	1.901***
Environment, Housing	1.773***	1.937***
Other Spending	3.716***	3.871***
Observations	51	51
Unit of Analysis	State	State

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Data: Gordon et al. (2016) for 2012 spending measures, MRI and RPP values for 2012

Notes: This table shows the coefficients from regressions of the natural log of per capita spending on the natural log of local prices, calculated using both the MRI and the RPP. We obtain per capita state-level spending for fiscal year 2012 from Gordon et al. (2016). Both the MRI and RPP are calculated for calendar year 2012. In column (1), we use 0.618 + 0.382*MRI as the price index in order to make the results comparable. 0.382 is the housing share of total expenditures in the Consumer Expenditure Survey (CE).

In Table 2.8b, we examine the elasticities of various categories of state and local spending (per capita) with respect to the MRI or RPP. Specifically, Table 2.8b shows coefficients from state-level regressions of the natural log of various spending amounts on the natural log of local prices calculated using either the MRI or RPP. Many of these categories (e.g., spending on education, police, environment, etc.) can be interpreted as spending on amenities. We calculate these elasticities for calendar year 2012 because the spending measures are derived from a report containing 2012 values (Gordon et al. 2016). As a result, we correspondingly use MRI and RPP values calculated for calendar year 2012. The MRI values in Table 2.8b are scaled similarly as those in Table 2.8a to make the MRI and RPP elasticities comparable.

For the majority of spending categories (including welfare, K-12 education, police, environment and housing, and other spending), the elasticities of spending with respect to the MRI and RPP are significantly positive and above unity. Conversely, the elasticities of higher education and health/hospital spending with respect to prices are negative (albeit statistically insignificant).

Thus, the elasticity of overall education spending (which combines K-12 and higher education spending) is still significantly positive but below unity. Taken together, the results in Tables 2.8a and 2.8b indicate that hourly wages, various sources of non-labor income, and the majority of state and local spending categories are positively correlated with prices. The results using the MRI and RPP are very similar to each other, and the elasticities in many cases are above one. These results therefore help to rationalize the finding that geographic adjustments for local price differences lead to the identification of a less deprived poor population – as places with higher prices also have higher incomes and more amenities.

2.9 Conclusion

In this paper, we assess the desirability of a geographic adjustment to poverty measures by examining whether or not it identifies a more deprived population. For nine of the ten domains of well-being indicators that we consider, the majority of outcomes suggest that those classified as poor with a geographic adjustment (many of whom live in urban areas) are less deprived than those classified as poor without a geographic adjustment (many of whom live in rural areas). Among eight of these nine domains, at least two measures suggest that geographic adjustments statistically significantly identify a less deprived poor population. This broad finding holds under three separate poverty measures (SPM, CIPM, and OPM) analyzed in two separate surveys (CPS and SIPP). It also persists after a variety of extensions and robustness checks, including partial adjustments that scale the geographic adjustment factor by different weights (to crudely account for amenities), using Regional Price Parities as an alternative geographic adjustment index, analyzing the effects on deep and near poverty, and varying the covariates used in the regression specification. In short, the preponderance of our evidence strongly suggest that incorporating a geographic adjustment runs counter to the central objective of a poverty measure: identifying the least advantaged population.

The results in this paper are directly relevant to efforts that seek to incorporate geographic cost-of-living differences into official poverty measures. Such efforts have been proposed by a wide variety of stakeholders and – over the past decade – have been experimentally implemented by the Census Bureau through its Supplemental Poverty Measure. Geographic adjustments to poverty thresholds would not only transform the face of poverty (by classifying fewer people as poor in lower-cost areas and more people as poor in higher-cost areas), but they would also have potentially enormous ramifications for the geographic allocation of anti-poverty funding that depends on poverty rates or an individual's poverty status. These settings range from individual eligibility for key transfer programs (such as SNAP and Medicaid) to school district eligibility for Title I funding from the federal government. Moreover, our results are relevant for analyzing broader efforts by governments and other entities to vary grants and subsidies to locations based on geographic differences in cost-of-living.

Future researchers might consider using more years of data to increase the statistical power of the estimates and examine if our results generalize to other time periods. However, the benefit of additional years for statistical power is limited in the SIPP, as the panel nature of the survey implies that observations are not independent over time within a panel. We also hope to use the fine geography that we have available in the surveys to bring in other indicators of well-being (such as mobility) at the Census Tract level. Finally, one of the key contributions of this paper is that it identifies and uses an extensive assortment of well-being outcomes in the survey and administrative data – building upon those used in Meyer, Wu, Mooers et al. (2021) – to evaluate the suitability of modifications to a poverty measure.[100] Going forward, these well-being indicators open the door for a variety of other analyses, including validating other changes to the poverty measure (e.g., incorporating in-kind transfers and asset flows to the resource measure) and measuring the targeting of government programs.

[100] While our assortment of well-being indicators provides a useful framework for understanding the circumstances of those in poverty, they may be less suitable for understanding the circumstances of those in the middle class or with higher incomes more broadly. This is because many of these indicators reflect "tail events" that may not be particularly relevant for those outside of poverty.

2.A Appendix

2.A.1 Detailed Description of Geographic Price Indices

A number of different price indices have been proposed to adjust poverty thresholds by geography. The Supplemental Poverty Measure (SPM) relies on the Median Rent Index (MRI) to geographically adjust its thresholds (Fox 2019). Using information from the 5-year American Community Survey (ACS) files, the MRI is calculated as the ratio of the median gross rent for a two-bedroom unit with a complete kitchen and bathroom in a specific metropolitan area or state to the median gross rent for the same type of unit in the U.S. The MRI is calculated for all Metropolitan Statistical Areas (MSAs) with populations of at least 100,000. Within each state, all metropolitan areas that have fewer than 100,000 people are grouped together into an "other metropolitan" category and all non-metropolitan areas are grouped together in a "non-metropolitan" category. In the SPM, one multiplies the share of the SPM poverty threshold that is attributable to housing and utility expenditures by the MRI to calculate the MRI-adjusted SPM poverty threshold.

Closely related to the MRI is a rescaled version of the MRI proposed by Renwick (2018, 2019) that seeks to reflect amenities in geographic adjustments of poverty rates. Renwick argues that the MRI will over-adjust thresholds if places with higher median rents also have greater amenities and will under-adjust thresholds if places with higher average rents have lower average amenities. In order to adjust for amenities, Renwick (2018) cuts the variation in the MRI index in half. This rescaled index is admittedly an "arbitrary" adjustment because the literature has established no clear methodology for incorporating amenities (p. 5).

Another proposed method for adding geographic adjustments to poverty is to use Regional Price Parities or RPPs (Aten 2005, Aten and D'Souza 2008). Calculated by the Bureau of

Economic Analysis (BEA), RPPs are spatial price indices that measure price differences in a broad set of goods across regions. They are constructed in two stages. First, price data and expenditure weights are derived from the Consumer Price Index (CPI) and the Consumer Expenditure (CE) survey.[101] These data come from over 200 individual items and comprise eight broad categories: housing, transportation, food, education, recreation, medical, apparel, and other. Thus, RPPs rely on a wider set of goods and services than the MRI, which only accounts for differences in rents. In all, this first stage yields information for 38 urban areas plus four rural regions. Second, these price levels and weights from these urban and rural regions are applied to all counties in the US. This stage incorporates data for housing from the ACS: rent levels are estimated annually for each state and across 3 years for metropolitan areas. These rent levels are quality-adjusted using a hedonic model that controls for characteristics such as type of structure, number of bedrooms, number of rooms, when it was built, and whether or not utilities are included in rent. One then multiplies the entire poverty threshold by the RPPs, as RPPs are designed to measure broad price levels (Bureau of Economic Analysis 2016).

A final method for incorporating geographic adjustments to poverty is to use food, apparel, and rent regional price parities or FAR RPPs (Renwick et al. 2014, Renwick et al. 2017). While RPPs cover many goods and services, many of these items are not included in the SPM thresholds. In contrast, FAR RPPs cover only the subset of goods in the RPPs that are also included in the SPM poverty threshold – namely food, apparel, and rent. In the case of the SPM, one multiplies the share of the SPM poverty threshold attributable to food, apparel, and rent expenditures by the FAR RPP to calculate the FAR-RPP-adjusted SPM poverty thresholds.

---

[101] Specifically, expenditure weights are found using CPI expenditure weights, BEA's personal consumption expenditures, and ACS rent expenditures.

Renwick et al. (2014) and Renwick et al. (2017) empirically compare the MRI, RPP, and FAR RPP by computing SPM poverty using each of these three measures. The authors compute various SPM poverty measures for reference years 2011 and 2015 and find significant differences between the measures.[102] The authors suggest that the differences between the indices are driven by the different weights given to housing expenditures. In the MRI, housing expenditures represent between 40 and 51 percent of the threshold while in the FAR RPP, housing expenditures represent between 28 and 65 percent of the threshold.[103] Dubay, Wheaton, and Zedlewski (2013) also assess how adjustments to poverty guidelines using the MRI and RPP would affect eligibility for several government transfer programs. The authors find that SNAP eligibility would decrease slightly under both indices, while Medicaid eligibility would increase under the MRI and decrease under the RPP.[104]

2.A.2 Detailed Description of Data Sources

In this section, we describe the survey and administrative data used in this paper and the methods we use to link the data sources. We focus on reference year 2010, since this is a year for which we have a relatively complete set of administrative records covering all income sources.

[102] For 2011, they find that the national poverty rate was 16.1% using the MRI, 15.6% using the RPP, and 16.4% using the FAR RPP. For 2015, the national poverty rate was 14.3% using the MRI, 13.8% using the RPP, and 13.7% using the FAR RPP. They also find that for 2009-2011, 19 states and Washington, D.C. had FAR RPP poverty rates higher than MRI poverty rates, 22 states had MRI poverty rates higher than FAR RPP poverty rates, and 9 states had a statistically insignificant difference. Lastly, they find that for 2014-2015, 5 states and Washington, D.C. had FAR RPP poverty rates higher than MRI poverty rates, 34 states had MRI poverty rates higher than FAR RPP poverty rates, and 11 states had a statistically insignificant difference.

[103] Variation in housing expenditure shares leads to a range of thresholds (rather than a single value) for both indices because the MRI's housing shares vary by housing tenure: the housing share for owners without a mortgage is 40 percent, while the housing share for owners with a mortgage and renters is about 50 percent. On the other hand, the FAR RPP's housing shares vary by metro area. The housing shares range from 28 percent in non-metropolitan West Virginia to 65 percent in other metropolitan Florida.

[104] While the authors find small effects at the national level, there are larger differences across states. In general, eligibility for government transfers increases in more urban states and decreases in more rural states.

Survey Data

Our survey data pertain to calendar year 2010 and come from the 2011 Current Population Survey Annual Social and Economic Supplement (CPS) and the 2008 Panel of the Survey of Income and Program Participation (SIPP). Both surveys are designed to be representative of the civilian non-institutional population of the United States. The 2011 CPS interviewed 75,000 households between February and April of 2011 about their incomes in calendar year 2010. The 2008 SIPP was a longitudinal survey that followed 42,000 households for up to 16 four-month waves, though not all households were observed for all 16 waves due to survey attrition.

For the CPS, we also bring in the publicly available SPM Research File produced by the Census Bureau. This dataset includes information necessary to calculate the SPM for all individuals in the CPS, including every individual's SPM reference unit structure, SPM income, SPM threshold, and geographic adjustment factor. Furthermore, we use the 5-year American Community Survey (ACS) to construct the Median Rent Index, which forms the geographic adjustment for the SPM and is therefore the primary index that we consider in this paper

While the default reference period for the CPS is a calendar year, the reference period in the SIPP is four consecutive months. We therefore combine information across multiple interview waves in the SIPP to calculate annual incomes. This decision not only makes the CPS and SIPP income measures more comparable, but it also facilitates the use of calendar year tax data. As our first step, we take as our analysis sample all individuals who appear in reference month 4 of Wave 6 (which spans April-July 2010). Then, we incorporate information on survey incomes from other months in 2010 during which the individuals in the analysis sample appear. Lastly, we proportionately scale up survey incomes for the 21% of individuals who are interviewed for only a portion of the year. This method will overstate incomes for individuals who leave the survey because of death, institutionalization, or unemployment, but understate incomes for those who do not respond because of increased employment, for example.

In addition to collecting monthly data on a rich set of income sources, the SIPP collects measures of material well-being, certain expenses, and household structure in topical modules administered in the final month of various interview waves. The topical module corresponding to Wave 6 assesses adult well-being and allows us to analyze five broad domaines of material well-being measures: material hardships, home quality problems, appliances owned, food security, and health. Because this topical module corresponds to reference month 4 of Wave 6, all members of our reference group appear in this module. As part of constructing the SPM using the SIPP, we must also incorporate certain expenses. The topical modules for Waves 4 and 7 cover many of these expenses, including medical and childcare expenses. However, due to survey attrition, 11% of individuals who appear in reference month 4 of Wave 6 do not appear in Wave 7's topical module, so we instead use Wave 4 information for these individuals. The topical modules for Waves 4 and 7 also contain information on assets owned, which constitute another measure of material well-being that we analyze. Lastly, the topical module for Wave 2 contains information on household structure necessary to construct SPM reference units.[105]

Administrative Data

We also bring in a number of administrative data sources. We first discuss the administrative records used to measure income. Earnings records come from Internal Revenue Service (IRS) W-2 Forms, the Detailed Earnings Record (DER) database of the Social Security Administration (SSA), and IRS 1040 Forms. IRS 1040 Forms also contain information on various

[105] This topical module allows us to find out which members of a household are cohabiting partners and which members are unrelated children. This information is then used to construct sharing units.

forms of asset income, including interest and dividends. The IRS 1040 extracts do not contain actual amounts corresponding to taxes paid and tax credits received, although they cover enough line items (e.g., Adjusted Gross Income (AGI), filing status, etc.) that we can simulate tax liabilities and credits relatively accurately. Data on retirement distributions come from IRS 1099-R Forms, which cover gross payments from employer-sponsored plans and Individual Retirement Account (IRA) withdrawals.

We also have a number of administrative program participation records from government agencies. Social Security benefit amounts come from the SSA's Payment History Update System (PHUS) file. Using the SSA's Master Beneficiary Record (MBR), we can distinguish between payments from Old-Age and Survivors Insurance (OASI) and payments from Social Security Disability Insurance (DI). Supplemental Security Income (SSI) benefits – covering both federal and federally-administered state payments – come from the SSA's Supplemental Security Record (SSR) file, and Service-Connected Disability payments to veterans come from the Department of Veterans Affairs. Finally, housing assistance records come from the Public and Indiana Housing Information Center (PIC) and Tenant Rental Assistance Certification System (TRACS) files, which cover all public and subsidized housing programs under the jurisdiction of the Department of Housing and Urban Development (HUD) but miss payments associated with housing programs administered by the Department of Agriculture (USDA), states, and localities.[106] For this paper, we do not bring in administrative data from state agencies on SNAP or TANF, because these program data are only available for a subset of states and we want our analysis sample to cover every state in the nation.

[106] We calculate a household's benefit amount as the difference between the gross rent and actual tenant payment. For public housing units, we impute gross rent amounts based on the average rent by zip code and household size in a given year. Olsen (2003, 2019) finds that the gross rents used to calculate housing assistance amounts are close to market rents and thus similar to the valuation placed by private renters on the units.

We also use administrative records to construct additional measures of well-being. First, we use IRS tax records from tax years 2008, 2009, 2011, and 2012 to construct a measure of permanent income, which includes AGI from Forms 1040 for individuals who filed a tax return and wages and retirement distributions from W-2s and 1099-Rs, respectively, for individuals who did not file a tax return. Second, we use the Social Security Administration's Numident file to calculate mortality rates, relying on its information on dates of death for all individuals with Social Security Numbers in the United States.

2.A.3 Methods for Constructing Poverty Measures

Description of Official Poverty Measure (OPM)

To provide some context for the SPM and CIPM, we start by briefly describing its predecessor – the Official Poverty Measure (OPM). The OPM uses pre-tax money income as its resource measure, calculated for a sharing unit consisting of a family or single unrelated individual. Poverty thresholds are based on the cost of a food plan consisting of a nutritionally balanced, low-cost diet for families of different sizes and composition. Because the cost of the plan varies by family size and composition, it implicitly embeds an equivalence scale that accounts for different food needs across different types of families.

However, the OPM features several well-documented shortcomings. Its resource definition does not incorporate in-kind benefits and tax credits, nor does it subtract tax liabilities – as a result, it fails to reflect the full set of resources available to a family for consumption. In addition, the equivalence scale implicit in the OPM thresholds exhibits a number of anomalies, in some cases suggesting that children are more costly than adults and failing to exhibit monotonically decreasing marginal increments for additional individuals (Ruggles 1990). More generally, the OPM thresholds have been criticized for reflecting only economies of scale in food (and not other goods) and for being anchored to 1963 values that have since been updated only for price changes (which do not fully reflect standard-of-living changes in the U.S.). Finally, the OPM does not classify a cohabiting unmarried couple as a single sharing unit, even though such a couple may pool resources in similar ways as a married couple that the OPM does classify as a single sharing unit. For a fuller discussion of these and other issues, see Citro and Michael (1995) and Blank (2008).

Constructing the Supplemental Poverty Measure (SPM)

Resource Measure

Unlike the Official Poverty Measure (OPM), which uses pre-tax money income as its definition of income or resources, the SPM uses as its resource measure pre-tax money income plus non-cash transfers net of certain expenses and taxes. Specifically, the SPM resource measure adds to pre-tax money income the estimated value of benefits received through SNAP, housing assistance, school meals, the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC), and the Low Income Home Energy Assistance Program (LIHEAP). It then subtracts federal and state income tax liabilities net of credits and payroll taxes. Finally, it subtracts expenses (or their estimates) for work, childcare, child support, and health care.

In the CPS, we utilize SPM incomes calculated by the Census Bureau and available in the SPM Research File. Note that the SPM incomes in this file rely on the public-use CPS with top-coded variables, while the estimates in the official SPM report rely on the internal version of the CPS with no top-coding (Fox 2019).[107] Nevertheless, we continue to use the variables on the

[107] Therefore, an individual's poverty status when estimated using the public-use file may be different than his/her poverty status estimated using the internal file. Using the public SPM research file for 2010, there were an estimated 48,934,000 people in SPM poverty. However, on the official SPM report there were an estimated 49,094,000 people in SPM poverty.

public-use Census Bureau file because the vast majority of individuals have no top-coded incomes, and imputing variables would lead to additional non-comparabilities between our estimates and internal Census ones. Most of the SPM income components are based on amounts explicitly asked about in the CPS, while some have amounts imputed either from survey reports of receipt (for school lunch, housing assistance, and WIC) or from related variables (for work expenses and Medicare Part B premiums). Lastly, tax liabilities and credits in the CPS are imputed using the CPS tax calculator. For a detailed summary of how the components of SPM income are constructed in the CPS, see Fox (2019).

In the SIPP, we directly calculate the components of SPM income other than pre-tax money income, with our work being heavily influenced by Short (2014). For in-kind benefits, we use reported survey amounts for SNAP and LIHEAP (at the family and household level, respectively) and impute amounts for housing assistance, school meals, and WIC for those reporting program receipt.[108] We calculate the value of housing assistance by subtracting monthly rent paid (reported in the SIPP) from imputed fair market rent.[109] We impute fair market rent by matching to HUD fair market rent data based on county, year, and the number of bedrooms.[110] To calculate the annual value of free/reduced price school lunch, we multiply the number of children receiving free lunch (which is reported in the SIPP) by the daily federal reimbursement rate for free lunch ($2.68) and

---

108 For LIHEAP, note that the SIPP's estimates do not include subsidies paid directly to utility companies while the CPS asks about those subsidies.

109 Another difference between the SIPP and CPS is that amounts for rent paid are reported in the SIPP, while they must be imputed in the CPS.

110 The number of bedrooms is not reported in the SIPP, so we must impute it. We construct the number of bedrooms in the survey via the following algorithm: (1) for childless households, give 1 bedroom to households with 1 or 2 adults and add 1 bedroom for each additional adult, (2) for households with children, give 2 bedrooms to households with 1 or 2 same-gender children and add 1 bedroom for each additional child. This attempts to follow HUD's minimal requirements that a dwelling must have at least one bedroom for every 2 people and that people of opposite gender (except for couples and very young children) should not be required to occupy the same bedroom (Currie and Yelowitz 2000).

the average number of school days per year (179).[111] We then perform parallel calculations for reduced-price lunch, free breakfast, and reduced-price breakfast.[112] For WIC, we value the subsidy for every person reporting receipt using USDA's estimates of average WIC food costs per person.[113]

Next, we calculate work-related expenses in the SIPP by multiplying 85 percent of the median weekly work expenses for adult workers by the number of weeks worked per month.[114] Childcare expenses, which are asked about in the topical modules for Waves 4 and 7, are calculated as the sum over all children of the weekly amount paid on childcare multiplied by the number of weeks worked by the parent. Child support paid and medical out-of-pocket expenditures (MOOP) are reported at the person level in the topical modules for Waves 4 and 7. We also subtract Medicare Part B premiums, which we calculate from survey-reported marital status and survey-reported total income.[115] To complete the SPM resource measure, we calculate tax liabilities and credits. While the SIPP contains a topical module on taxes paid, it contains sparse and incomplete data so we calculate tax liabilities and credits on our own. We start by generating tax units based on reported family structure in the SIPP, and then – using reported incomes and demographic characteristics – we simulate taxes using the TAXSIM calculator from the National Bureau of Economic Research (NBER).[116] Using this method, we are able to calculate federal and state

---

[111] Reimbursement rates are published in the Federal Register: https://www.govinfo.gov/content/pkg/FR-2009-07-15/pdf/E9-16745.pdf.
[112] The daily reimbursement rate is $2.28 for reduced-price lunch, $1.46 for free breakfast, and $1.16 for reduced-price breakfast. To calculate the annual value of these programs, we multiply the daily reimbursement rate by 179 school days per year. Another difference between the CPS and SIPP is that the SIPP asks about receipt of both free/reduced lunch and breakfast, while the CPS asks only about receipt of free/reduced lunch. As a result, a value of zero is always assigned for school breakfast in the CPS.
[113] In 2010, the USDA estimated an average monthly food cost of $41.43 per person. See https://fns-prod.azureedge.net/sites/default/files/resource-files/wisummary-7.pdf.
[114] This follows the methodology used in the official SPM report (Fox 2019).
[115] The table of premium amounts can be found at https://www.cms.gov/newsroom/fact-sheets/cms-announces-medicare-premiums-deductibles-2010.
[116] We assume all married couples file jointly. Dependents are children under the age of 19, children under the age of 24 and enrolled in school, individuals that are permanently and totally disabled, or other relatives with income below

income and payroll tax liabilities, as well as tax credits including the Earned Income Tax Credit (EITC), Child Tax Credit (CTC), Child and Dependent Care Credit, and Making Work Pay Credit. For a more detailed discussion of how these taxes are calculated, see Meyer et al. (2020).

Resource Unit

While the OPM uses a family (defined as all people living together related by birth, marriage, or adoption) as its resource unit, the SPM resource unit additionally includes cohabiting partners, unrelated children under the age of 15, and foster children between the ages of 15 and 22. In the CPS, we use resource unit identifiers present in the SPM Research File. In the SIPP, we rely on the topical module for Wave 2 – which contains detailed household relationship information specifically identifying cohabiting partners and unrelated and foster children – to construct the SPM resource unit. Because this household relationship information is only asked about during Wave 2, we cannot use it to identify relationships for individuals entering the survey after Wave 2. For the 12% of individuals who are in our reference group but do not appear in Wave 2, we check whether they enter an existing family. Approximately half of these individuals enter into existing families, and we assign them to their respective families' SPM resource units. For the other half of these individuals who do not enter into existing families, we assign them their own SPM resource unit.

Poverty Thresholds and the Equivalence Scale

The SPM uses different poverty thresholds than the OPM does and adjusts the thresholds in a different way. The SPM thresholds are based on out-of-pocket spending on food, shelter,

$3,650. We calculate taxes at the family level, and then aggregate family-level amounts to the sharing unit level if there are multiple families per sharing unit.

clothing, and utilities. The threshold is calculated as 1.2 times the average spending of those in the 30^{th}-36^{th} percentiles of spending on these expenses, computed using five years of Consumer Expenditure Survey (CE) data. However, the SPM differentiates between three housing tenure groups (homeowners with mortgages, homeowners without mortgages, and renters) and calculates spending separately for these groups (implicitly assuming they are otherwise the same).

The SPM then applies two different adjustments to this base threshold. The first is an equivalence scale that adjusts the threshold based on the size of the reference unit. In particular, we multiply the base threshold by a three-parameter SPM equivalence scale, calculated using the demographic characteristics of the resource unit and divided by the scale for a two-adult, two-child family. The second adjustment is a geographic adjustment, which uses data from the 5-year ACS files to adjust base thresholds for differences in median gross rents (for two-bedroom apartments) across geographic areas. These medians are divided by the national median for the same unit type to calculate the MRI, and are computed for each of the 264 metropolitan statistical areas (MSAs) that can be publicly identified in the CPS ASEC. For areas that do not fall into any of the publicly identified MSAs, medians are estimated for all non-metro areas in a state as well as for a combination of all other metro areas within a state. This gives 358 adjustment factors for the year 2010: 264 from MSAs, 48 from states' non-metro areas, and 46 from states' all other metro areas.[117]

Putting these all together, the final SPM threshold is found by multiplying the base threshold together with the equivalence scale factor for family structure and the geographic

[117] We extract these 358 adjustment factors from the Census Bureau's SPM file for the CPS and merge them onto the SIPP, assigning the correct adjustment factor based on MSA, state, and metro status. Not all states have non-metro areas or other metro areas, leading to less than 50 such adjustment factors for non-metro areas and for other metro areas.

adjustment factor (which scales the MRI by the share of expenditures taken up by housing costs). More formally, the threshold for an SPM unit is given by:

$$SPM\ Threshold_{t,ac,sm} = (Base\ Threshold)_t \times \frac{(Equivalence\ Scale\ Factor)_{ac}}{E} \times [(Housing\ Share_t \times MRI_{sm}) + (1 - Housing\ Share_t)], \qquad (2.7)$$

where t is the unit's housing tenure, a and c represent the number of adults and children in the unit (respectively), and s and m denote the unit's state and MSA (respectively). The constant E is the equivalence scale for a two-adult, two-child family. The components of Equation (2.7) that vary by housing tenure and resource unit composition can be further broken down as follows:

$$Base\ Threshold_t = \begin{cases} \$25{,}018 & \text{for } t = \text{Owner with Mortgage} \\ \$20{,}590 & \text{for } t = \text{Owner without Mortgage} \\ \$24{,}391 & \text{for } t = \text{Renter} \end{cases}$$

$$Equivalence\ Scale\ Factor_{ac} = \begin{cases} (adults)^{0.5} & \text{for } a \geq 1, c = 0 \\ (adults + 0.8 \times first\ child + 0.5 \times other\ children)^{0.7} & \text{for } a = 1, c \geq 1 \\ (adults + 0.5 \times children)^{0.7} & \text{otherwise} \end{cases}$$

$$Housing\ Share_t = \begin{cases} 0.510 & \text{for } t = \text{Owner with Mortgage} \\ 0.404 & \text{for } t = \text{Owner without Mortgage} \\ 0.497 & \text{for } t = \text{Renter} \end{cases}$$

The values for the base threshold by housing tenure, all for a unit with two children and two adults, correspond to 2010.[118]

Constructing the Comprehensive Income Poverty Measure (CIPM)

<u>Resource Measure</u>

For the second poverty measure that we analyze (the CIPM), we use the CID to construct an alternative resource measure that differs from the SPM resource measure conceptually and brings in administrative data to measure incomes more accurately. We begin by discussing the conceptual differences. First, in line with the OPM and in contrast with the SPM, the CIPM resource measure does not subtract expenses for work, childcare, child support, and health care. While subtracting these expenses may theoretically yield a resource measure that better approximates the resources available for consumption, prior research has also shown that subtracting certain expenses (e.g., medical costs) identifies a poor population that appears less materially deprived (Meyer and Sullivan 2012a). Second, the CIPM resource measure estimates a flow value of services from home and car ownership as well as an annuity value of other net assets. Specifically, we estimate the rental equivalent of home ownership by multiplying home values by 8% and the service flow from vehicle ownership by multiplying vehicle values by 13%.[119] For other assets (namely business, liquid, financial, and pension assets) net of housing, vehicle, and other debt, we follow Wolff et al. (2012) in calculating an annuity value using separate rates of returns for each type of asset grouping. Because the CPS does not ask about assets in detail, we

---

[118] See https://www.census.gov/prod/2011pubs/p60-241.pdf for the base thresholds by housing tenure. See https://www.bls.gov/pir/spm/spm_threshold_200518.xlsx for the housing shares by housing tenure.

[119] The 8% of home value is close to the mean ratio of the rental equivalent to home value from the Consumer Expenditure (CE) Interview Survey. The 13% of vehicle equity is the sum of the annual depreciation rate (which is calculated from the CE Interview Survey using the sale prices of new and used vehicles over time and estimated to be 13%) and the interest rate on certificates of deposit (estimated to be close to zero in 2010 for 1- and 2-year interest rates).

are only able to estimate these asset flows in the SIPP (using information from the topical modules for Waves 4 and 7).

Next, we discuss the measurement differences. First, we bring in multiple administrative sources to construct improved measures of earnings. Specifically, we take the higher of wages in the DER, W-2s, and 1040s added to self-employment earnings from the DER to obtain a combined measure of total administrative earnings. However, we continue to use survey earnings that exceed combined administrative earnings if they reflect earnings that are plausibly missed in the tax records.[120] Next, we replace survey reports of asset income (namely interest and dividends), retirement income, Social Security, SSI, and veterans' benefits with their counterparts from the administrative data. We then use TAXSIM to calculate tax liabilities and credits, relying on tax unit structure and incomes from the IRS tax records as inputs. We also calculate taxes for units who do not link to a 1040 (in the event that they are late filers or had taxes withheld), relying on their survey-reported incomes and family structure. Finally, we replace survey values of housing assistance with administrative values, although we treat survey respondents reporting receipt that do not appear in the administrative data as true recipients (since the administrative data cover only HUD-administered programs).

[120] We incorporate survey earnings only when they are not imputed, when other employment characteristics (hours worked, weeks worked, industry of job, job occupation, and size of employer) are not imputed, and when one of the following conditions holds: 1) earnings are missing across all administrative sources, 2) the number of survey-reported employers exceeds the number of employers in the administrative data, 3) the respondent reports being self-employed in the survey, or 4) the respondent reports working for a small employer in the survey. In the CPS, we continue to use survey earnings for only 27% of all individuals whose survey earnings exceed their administrative earnings.

Resource Unit

The CIPM uses the same resource unit as the SPM – namely a family and additional cohabiting partners, unrelated children under the age of 15, and foster children between the ages of 15 and 22.

Poverty Threshold and Equivalence Scale

The CIPM also uses nearly the same poverty thresholds as the SPM, with the key exception being that the base threshold and housing share (in the geographic adjustment factor) in the CIPM no longer vary by housing tenure. Because the CIPM resource measure explicitly accounts for the flow value of home ownership, there is no longer a reason to set distinct thresholds (implicitly accounting for differences in available resources) for distinct housing status groups. The three-parameter equivalence scale used in the CIPM thresholds remains the same as that used in the SPM thresholds. More formally, the threshold for a CIPM unit is given by:

$$\begin{aligned} \textit{CIPM Threshold}_{ac,sm} &= \textit{Base Threshold} \times \frac{(\textit{Equivalence Scale Factor})_{ac}}{E} \\ &\times [(\textit{Hous. Share} \times \textit{MRI}_{sm}) + (1 - \textit{Hous. Share})], \end{aligned} \tag{2.8}$$

where a and c represent the number of adults and children in the unit (respectively) and s and m denote the unit's state and MSA (respectively). The constant E continues to designate the equivalence scale for a two-adult, two-child family. We set the base threshold in Equation (2.8) equal to $22,113 (the official poverty threshold for a two-adult, two-child family) and the housing share equal to 0.382 (which is the share of overall consumption dedicated to housing in 2010).

CHAPTER 3:
The Use and Misuse of Income Data and Extreme Poverty in the United States[121]

3.1 Introduction

There are reasons to be simultaneously concerned and skeptical about recent reports of high and rising rates of extreme poverty in the United States. Several distinguished scholars have argued that millions of Americans – many of them children – live on less than a few dollars per day. Other researchers have reported high rates of "disconnected" people, defined as those with neither earnings nor government benefits. Relying predominantly on survey reports of income, both groups claim that these problems have been rising sharply over time. On the other hand, researchers have long contended that survey reports in the tails of the income distribution have a disproportionate share of errors. Some of these scholars have pointed to evidence of increased underreporting of income in household surveys or conflicting evidence from consumption data. This paper addresses these questions by bringing to bear a combination of previously under-utilized survey data and newly linked administrative data. These data allow us to re-examine rates of extreme poverty and shed light on other questions, including the targeting of in-kind transfers, the effects of welfare reform, and the measurement of poverty.

Focusing on 2011 data from the Survey of Income and Program Participation (SIPP), we show that more than 90% of the 3.6 million households with survey-reported cash incomes below $2/person/day are misclassified. Our preferred methodology first implements a series of

---

[121] This chapter was previously accepted for publication in the *Journal of Labor Economics* on July 21, 2020. Copyright holder is the University of Chicago (0734-306X/2021/39S1-0010), and the Article DOI is https://doi.org/10.1086/711227.

adjustments using the public survey data. We begin by reclassifying households as not in extreme poverty if they received sufficiently high amounts of in-kind transfers including SNAP (Supplemental Nutrition Assistance Program), the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC), and housing assistance. We then use reported hours worked for pay to correct for underreported earnings (identifying SIPP editing problems in a large share of cases), and we also account for those who possess substantial assets. To further examine households not captured by the survey-only adjustments, we replace survey reports of earnings, asset income, retirement distributions, Old-Age, Survivors, and Disability Insurance (OASDI), Supplemental Security Income (SSI), SNAP, and housing assistance with values from linked administrative tax and program data and also account for the Earned Income Tax Credit (EITC).

In the end, our best estimate is that 0.24% of households containing 0.11% of individuals lived in extreme poverty in 2011. The difference between these percentages is explained by the finding that 90% of extreme poor households consist of a single individual. Replicating the analysis with the 2012 Current Population Survey Annual Social and Economic Supplement (CPS ASEC), we estimate that only 0.18% of households and 0.13% of individuals are in extreme poverty for the 2011 calendar year. These estimates from the two surveys are remarkably similar to rates that researchers have calculated using consumption data,[122] suggesting that improved measures of income can reconcile past inconsistencies between income and consumption measures of poverty. Furthermore, these results demonstrate that taking survey incomes in the far left tail at face value would be a misuse of the data. Yet, we suspect the true extreme poverty rate is even lower than what we estimate, given the evidence of survey underreporting for many income sources – like unemployment insurance, Temporary Assistance for Needy Families (TANF),

[122] See Chandy and Smith (2014) and Hall and Rector (2018).

workers' compensation, veterans' benefits, and informal earnings – for which we have not been able to incorporate administrative data.

The survey-only adjustments account on their own for 78% of the total decrease in extreme poverty. Using an alternative ordering that brings in the administrative data first (as some readers might prefer), we find that the administrative data adjustments account on their own for 90% of the change in extreme poverty. In addition, according to the administrative data alone, nearly 80% of the misclassified households overall are initially categorized as extreme poor due to errors or omissions in cash reports of earnings, asset income, retirement income, OASDI, SSI, or the EITC – meaning in-kind transfers play a secondary role. There is no unique way of decomposing the contribution of definitional changes and survey error, as whichever set of adjustments we incorporate first will remove the vast majority of those initially classified as in extreme poverty. The survey-only adjustments, in particular, encompass both definitional changes (incorporating in-kind transfers and assets) and corrections for underreported earnings.

Our preferred methodology implements the survey-only adjustments first, since this anchors our results to a literature that has relied only on survey data and allows many of our results to be replicated using publicly available survey data. The survey-only adjustments are also remarkably robust to a number of modifications. For example, applying either the minimum wage or half the minimum wage to hours worked removes virtually the same number of households from extreme poverty, and excluding imputed hours from the corrections for underreported earnings also yields trivial impacts. In addition, subsidy amounts for housing assistance are always above the extreme poverty threshold for the households in our analysis – but even excluding housing assistance altogether yields final estimates of extreme poverty that are almost exactly the same as those including housing assistance. We also set high enough thresholds for the survey asset adjustment, which we apply last, to suggest that the households it removes from extreme poverty have enough resources to spend at levels above $2/person/day. Yet, when it comes to identifying how far off the survey data are alone, we prefer the administrative data because they provide specific income amounts that can be used to correct misreported survey values. At the same time, the income amounts from the administrative data must be a lower bound because we are missing administrative data for a number of key income sources.

One of this paper's key methodological advances is the use of multiple sources of administrative and survey data to validate the survey-only adjustments. For the groups reclassified due to underreported earnings and substantial assets, we find that 72-93% of these households have incomes from the administrative data above the extreme poverty threshold and 47-65% have incomes above the poverty line, depending on the subgroup. Using detailed information from SIPP topical modules, we find that these groups have material well-being levels (based on measures of material hardship, appliance ownership, and housing quality) that are similar to the U.S. average. They are also comparable to the average household on a host of survey demographics, such as years of education, health insurance coverage (especially private coverage), and occupation.

Accordingly, the preponderance of evidence suggests that the households reclassified by underreported earnings and substantial assets have survey incomes that are likely to be gross errors. These results potentially explain the lack of a strong correlation found by several studies between income poverty and material hardship.[123] In contrast, the households reclassified due to receipt of in-kind transfers appear to be significantly worse off than the official poor on multiple dimensions of well-being, implying that these benefits are well targeted to the needy. These results are consistent with past findings that individuals excluded from the poverty rolls under the Census

[123] See, for example, Mayer and Jencks (1989), Meyer and Sullivan (2003, 2011, 2012a), and Short (2005).

Bureau's Supplemental Poverty Measure (which incorporates in-kind transfers into income, raising some recipients above the poverty line) appear worse off, on average, than the official income poor.[124]

It is important to keep in mind that our best estimate of the extreme poverty rate is not necessarily a final estimate for the entire population. The SIPP excludes homeless individuals and institutionalized populations (such as those living in nursing care facilities and prisons) from its survey frame, meaning our estimates of extreme poverty are understated if substantial numbers of the homeless and other institutionalized populations are in extreme poverty. We should emphasize, however, that the literature reporting high rates of extreme poverty has relied on survey data that exclude the homeless and institutionalized. If anything, these caveats further highlight the imperfect ability of most survey data, when analyzed at face value, to identify the extreme poor.

While this paper demonstrates that the rate of extreme poverty in the United States is substantially lower than what has been reported, we do not contend that there is little deprivation in the United States. Rather, we argue that focusing on such low-income thresholds as $2/person/day in the United States is likely to yield a group filled with more gross errors than households that are truly impoverished. For instance, nearly 50% of the households classified as extreme poor based on survey-reported cash have incomes above the poverty line in our administrative data sources (which are incomplete). Moreover, the households receiving means-tested in-kind transfers – who appear to be among the most materially deprived Americans – are almost all not in extreme poverty by virtue of the extreme poverty income thresholds being lower than benefit amounts. Among the households that appear to be truly extreme poor, and therefore disconnected from work or the safety net, the vast majority consist of a single childless individual.

[124] See Meyer and Sullivan (2012a) and Fox and Warren (2018).

Contrasting sharply with the focus in the literature on extreme poverty among households with children, this finding is consequential from a policy perspective as eligibility for programs is often dependent on household composition.

While our main approach begins with households that are below $2/person/day in the survey reports, as a robustness check we start from the full sample, combining the survey and administrative data. In this alternative analysis, we rely on some of the survey reports in the case of earnings and housing assistance. We do this because our source of administrative earnings data turns out to be incomplete, missing categories of earnings and individuals without work authorization. The housing data are incomplete as well, missing several million subsidized housing units that are not part of the main programs administered by the Department of Housing and Urban Development (HUD). For all other income sources, we simply replace the survey reports with the administrative records. The results using this approach do not differ appreciably from our main results. We also confirm that our main results hold at a cutoff of $4/person/day, for shorter and longer time intervals, and when imputed values of hours or income components are set to zero.

More generally, this paper is one of the first from an unprecedented new project that assembles and links survey and administrative data on income, program receipt, and closely related information (Medalia et al. 2019). The project's goals include 1) improving household surveys and statistically based tax administration and 2) better understanding poverty, inequality, and the effects of government transfers. We initially focus on extreme poverty in this paper because the results are so stark and they demonstrate the capacity of the linked data to change our understanding of poverty. There is great value in linking survey and administrative data, even relative to methods that attempt to formally adjust for misreporting within the survey. Two studies have found that sophisticated adjustments like the Urban Institute's Transfer Income Model

(TRIM) allocate SNAP and TANF benefits to those with very different survey incomes than true recipients, likely biasing any poverty estimates.[125]

The remainder of the paper is structured as follows. Section 3.2 reviews the literature on extreme poverty and discusses why the rates using survey-reported cash income are so high. Section 3.3 describes the survey and administrative data and the process used to link them. Section 3.4 discusses the methodology used to correct for errors in income reports. Section 3.5 describes the main results from the SIPP, and Section 3.6 describes results validating the survey-only corrections and adjustments. Section 3.7 replicates the analysis for the CPS and compares it to the SIPP, and Section 3.8 presents the results of robustness checks and additional caveats. Section 3.9 concludes.

3.2 Literature

3.2.1 Past Claims of Extreme Poverty and Conflicting Evidence

In a series of papers and a best-selling book, Edin and Shaefer document the prevalence of extreme poverty, which they define as having cash income less than $2/person/day. Using Wave 9 of the 2008 SIPP Panel, Shaefer and Edin (2013) find that 4.3% of all non-elderly households with children (constituting 1.65 million households and 3.55 million children) lived in extreme poverty in a given month in mid-2011.[126] Using the 2012 CPS ASEC adjusted using the TRIM, Shaefer and Edin (2017) contend that 1.3 million children (1.8% of all children) lived under $2/day based on annual cash income during the 2011 calendar year.[127] Combining quantitative analyses

[125] See Shantz and Fox (2018) and Mittag (2019).
[126] We begin our empirical work by replicating these numbers.
[127] Even though Shaefer and Edin examine reference year 2011 in both of their papers, the counts of children in extreme poverty differ rather dramatically. We think this difference is due to a few reasons. First, the higher number using the SIPP is based on the fourth reference month of a wave, rather than the monthly average in a wave. Second,

with ethnographic evidence on the day-to-day lives of the extreme poor, Edin and Shaefer (2015) further shed light on the deprivation faced by such households. Concomitantly, Deaton (2018) uses survey data from the CPS to assert that 5.3 million individuals in the United States lived under $4/day in income after taxes and in-kind transfers during the 2015 calendar year. These striking numbers have received a great deal of attention in the policymaking process and the press,[128] and they were featured in a prominent United Nations report on the state of poverty in the United States (United Nations 2018).

A related literature has arisen around the plight of "disconnected" individuals and families, who are defined as having little to no earnings and little or no government benefits (usually cash welfare). Most of these studies focus on single mothers. Turner et al. (2006) use survey data from the Women's Employment Study and find that 9% of single mothers who received cash welfare in February 1997 became disconnected for at least a quarter of the following 79 months (following welfare reform in 1996). Using data from the SIPP and CPS, Blank and Kovak (2009) find that more than 20% of single mothers who live below twice the official poverty line in the mid-2000s have no annual earnings or welfare receipt. These high rates of disconnected single mothers are echoed in Loprest (2011) and Loprest and Nichols (2011), who also utilize the SIPP.

Importantly, a number of these studies argue that rates of extreme poverty and disconnectedness have risen greatly over time in response to welfare reform. Shaefer and Edin (2013) calculate that the number of households with children in extreme poverty grew by 159% between 1996 and 2011. This rate of increase snowballs to 748% between 1995 and 2012 after

as we discuss in Section 3.7, the SIPP appears to have a non-trivial number of households with zero earnings but positive reports of hours worked for pay – an inconsistency that does not appear in the CPS. Finally, the lower CPS number relies on the Urban Institute's TRIM micro-simulation model to adjust for underreporting of cash transfers in the survey.

128 For example, see https://www.washingtonpost.com/news/wonk/wp/2018/06/25/trump-team-rebukes-u-n-saying-it-overestimates-extreme-poverty-in-america-by-18-million-people/?utm_term=.1f7ba77d349a.

using TRIM to adjust for underreporting in the CPS, with Shaefer and Edin (2017) attributing the growth entirely to cuts in cash welfare. Blank and Kovak (2009) also find that the rate of disconnected single mothers nearly doubled between 1995 and 2005 using the CPS, and Loprest and Nichols (2011) calculate a similar increase in the share of disconnected single mothers between 1996 and the 2004-2008 period.

At the same time, another literature provides evidence at odds with the results in Shaefer and Edin (2013, 2017) and related papers on disconnectedness. Some studies improve the measurement of income by including in-kind transfers and attempting to adjust for survey underreporting. Winship (2016) re-examines rates of extreme poverty by applying a number of adjustments to reported cash income in the CPS, which include incorporating in-kind transfers (non-medical and medical benefits are separated), taxes and tax credits, and a less biased price index (the Personal Consumption Expenditures (PCE) deflator) than the Consumer Price Index for All Urban Consumers (CPI-U).[129] Winship also uses TRIM3 to correct for underreporting of various transfers and divides household income by an equivalence scale to better account for resource sharing. Winship finds that the adjusted rates of extreme poverty have fallen since welfare reform to approximately 0.1% among all children and closer to 0.01% among children of single mothers in 2012. Using also the CPS, Brady and Parolin (2020) calculate that 0.40% of individuals lived in households with incomes less than $2/person/day in 2015, after accounting for taxes and transfers (including SNAP), correcting for underreporting of TANF and SSI with TRIM, and accounting for household size. Parolin and Brady (2019) employ a similar methodology (but

[129] Both the Federal Reserve's Federal Open Market Committee and the Congressional Budget Office use the PCE Deflator rather than the CPI-U to calculate inflation, because the former suffers less from a series of biases that plague the latter (Congressional Budget Office 2012, Bullard 2013).

additionally adjust for SNAP underreporting using TRIM) to find that 0.08% of children lived in households with incomes less than $2/person/day in 2015.

Rather than relying on survey-reported cash income to measure extreme and deep poverty, other studies focus on measures of consumption or hardship. In an early paper, Mayer and Jencks (1989) find that 43% of a sample of Chicagoans surveyed in the mid-1980s with incomes below the official poverty line reported expenditures on food, housing, and medical care that exceeded their incomes. For disadvantaged single mothers at the 10th percentile in the 1990s, Meyer and Sullivan (2003) also find that expenditures exceeded income by 47% and 24% in the Consumer Expenditure (CE) and Panel Study of Income Dynamics (PSID) surveys, respectively. In subsequent papers, Meyer and Sullivan (2004, 2008, 2012a) find that low percentiles of consumption rose in the period following welfare reform and that deep consumption poverty has fallen sharply over time.

Additional papers in recent years have used the CE Survey to calculate decidedly low rates of consumption-based extreme poverty (spending less than $2 or $4/person/day) or deep poverty (spending less than half the official poverty line). Chandy and Smith (2014) find that only 0.07% of the U.S. population spent less than $2/person/day in the fourth quarter of 2011. Hall and Rector (2018) examine all households interviewed in the CE survey since 1980 and similarly find that 0.08% of the U.S. population spent less than $4/person/day. They also calculate an expenditure-based deep poverty rate of 0.5% in 2017, considerably lower than the official income-based deep poverty rate of more than 6% in 2017. Much like the results in Meyer and Sullivan (2012a), Hall and Rector find that deep consumption poverty fell sharply from a rate of roughly 2% in the mid-1980s, with this fall being especially precipitous for single parents after welfare reform.

3.2.2 Why are Extreme Poverty Rates from Survey-Reported Cash Income So High?

There are several major reasons why the literature has found such high rates of extreme poverty when relying on survey reports of pre-tax cash income. First, these calculations ignore in-kind transfers and tax credits. The majority of means-tested transfer dollars are in-kind, and a broad range of authors has argued that non-medical in-kind benefits should be counted as income (see, for example, Ellwood and Summers 1985, Citro and Michael 1995, Blank 2008). In particular, SNAP benefits can be plausibly treated as cash payments, since benefit amounts usually fall below the pre-receipt food expenditures of recipient families (Hoynes and Schanzenbach, 2009, Ben-Shalom et al. 2012). The gross rents that are used to calculate housing assistance amounts have also been found to be close to market rents and thus similar to the valuation that private renters put on the units (see Olsen 2003, 2019). Several studies have even argued that the per-dollar value of benefits from transfer programs may exceed cash earnings, as transfers play an important role in insuring earnings shocks (Blundell et al. 2008, Blundell 2014, Deshpande 2016).

Given that the nature of the safety net in the U.S. has changed dramatically, it is important to account for in-kind transfers and tax credits when comparing outcomes over time. While cash welfare (Aid to Families with Dependent Children or AFDC, which became TANF) payments fell by two-thirds between 1996 and 2011, SNAP payments more than doubled and EITC benefits increased by approximately half during the same time period, both transferring more new dollars than were cut from TANF (Meyer et al. 2015). Other in-kind transfers like public and subsidized housing followed a similar upward spending trajectory over time.[130] Consequently, focusing on changes in poverty rates based solely on pre-tax cash income would be anachronistic. These concerns, in large part, motivated the reports that led the U.S. Census Bureau to start calculating

[130] See https://www.cbpp.org/research/housing/national-and-state-housing-fact-sheets-data.

the Supplemental Poverty Measure (SPM) in 2011, which takes into account many of the non-cash programs and tax credits not included in the official poverty measure. To their credit, Shaefer and Edin (2013) find that SNAP, tax credits, and housing subsidies together cut the pre-tax cash extreme poverty rate for households with children by 63% in 2011. But researchers and policymakers continue to highlight estimates that exclude these – and other – important government programs.

Another reason for high extreme poverty rates in the literature is that studies almost universally rely on survey income with substantial errors even after editing, despite many studies demonstrating significant holes in the income data that arise from survey underreporting. For example, Meyer and Mittag (2019) find that 63% of Public Assistance recipients in the CPS and 44% in the SIPP in New York do not report receipt, while 43% of SNAP recipients in the CPS and 19% in the SIPP do not report receipt. Bee and Mitchell (2017) find that 46% of pension income recipients do not report receipt in the CPS.

While the CPS has often been found to suffer from more pronounced underreporting than the SIPP, the latter is not immune to errors. Meyer and Wu (2018) find that, among single parent families, the poverty reduction effects of SSI, OASDI, and Public Assistance from the SIPP are each less than 44% of what the administrative data indicate.[131] For all families, the SIPP yields effects on near poverty of SNAP and Public Assistance that are two-thirds and one-half, respectively, what the administrative data generate (Meyer and Wu 2018). These holes in the SIPP income data have also grown over time. Since 2000, there has been a 7 percentage point increase in the share of dollars missed by the SIPP for TANF, unemployment insurance, and workers' compensation (Meyer et al. 2015). The share of SIPP dollars that are imputed has also doubled

[131] Meyer and Wu (2018) take the administrative data to be truth, though the administrative data are likely incomplete. For example, administrative tax data may miss individuals who do not file tax returns or whose employers fail to file.

since 1990, and errors in reporting amounts for SSI and OASDI rose sharply between the 1996 and 2008 SIPP panels (Gathright and Crabb 2014).

These errors in survey-reported income are likely most pronounced at the very bottom of the reported income distribution. Many studies have suspected or found errors in income reports at the tails of the distribution (Lillard et al. 1986, Blank and Schoeni 2003, Bollinger et al. 2019). Especially in the left tail, research has shown that reported expenditures are often a multiple of reported incomes. This pattern has been found not only in U.S. survey data (see Meyer and Sullivan 2004, 2008, Hall and Rector 2018), but in Canadian and British survey data as well. Brzozowski and Crossley (2011) use data from the Canadian Survey of Household Spending and the Family Expenditure Survey to show that total expenditures exceed disposable income by approximately a multiple of five in the bottom decile of the income distribution. Brewer et al. (2017) use data from the United Kingdom Living Costs and Food Survey and find that households in the bottom 1% of the income distribution (who live on less than £75/week) report a median expenditure level of £400/week, equivalent to the population median! The authors find that median expenditures are actually decreasing in income for households living on less than £110/week. They are best able to explain this puzzle by underreporting of income rather than over-reporting of expenditures or consumption smoothing over time.

3.3 Data

This section describes the detailed sources of survey and administrative data we use in this paper. The section also explains how we link these data and the advantages of using the combined data over survey or administrative sources alone.

3.3.1 Survey Data

Our survey data primarily come from the 2008 Survey of Income and Program Participation (SIPP). In Section 3.8, we also describe results using the Current Population Survey Annual Social and Economic Supplement (CPS). Each panel of the SIPP lasts several years, and individuals and households are followed longitudinally within each panel. Specifically, each respondent is interviewed every four months as part of an interview "wave". In each wave, the SIPP collects information about the income and government transfers received during the four months since the last interview wave. Nearly all of these income sources are reported at the month level. Accompanying these income data is detailed information on demographics, assets and liabilities, material well-being, and health status (among other items). Many of these characteristics are available in the SIPP topical modules. These sets of questions on a specific subject differ across interview waves and are asked on top of the core questions.

To begin from a known starting point in the literature, we focus on Wave 9 of the 2008 SIPP Panel, whose reference months include January 2011 to July 2011.[132] This sample includes the observations used by Shaefer and Edin (2013). We also link topical modules from Wave 9 and

[132] The interview wave spans 7 calendar months because of the staggered nature of the interviews. Respondents are divided into four groups, each of which has a different starting month in the wave. For example, one set of respondents in Wave 9 has reference months spanning January-April 2011. The other three sets of respondents each have reference months spanning February-May 2011, March-June 2011, and April-July 2011.

other waves that include questions on material hardships and housing quality (Wave 9), assets and liabilities (Waves 7 and 10), and disability status (Wave 6). The proximity of Wave 9 to Waves 6, 7, and 10 presents another advantage of focusing on Wave 9 in our analysis. This proximity aids the comparability of the time periods and reduces sample attrition from wave to wave. An additional benefit of examining Wave 9 is that its reference months are within a single calendar year, unlike many other waves. This choice makes linkage to tax records, which are at the annual level, more convenient.

The SIPP sample is intended to be representative of the civilian non-institutional resident population of the United States, excluding individuals living in institutions and military barracks. We use households – rather than families, as official poverty estimates do – as our units of analysis for two main reasons. First, many of the questions in the SIPP (such as equity values for specific assets and material hardship) are asked at the household level. Second, individuals who are particularly destitute may rely on the additional resources of those outside their immediate families. If so, the household may be the more natural unit for analyzing the circumstances of the extreme poor. In practice, the distinction is not especially important as 92% of all households and 94% of reported extreme poor households have one family.

3.3.2 Administrative Data

Our administrative records are derived from a number of sources, which we broadly classify into two categories: tax records from the Internal Revenue Service (IRS) and Social Security Administration (SSA), and program receipt records from various state and federal agencies. Table 3.1 describes for each income component the source of the administrative data, the income unit, the disbursement frequency, and the number of states covered.

Table 3.1. Administrative Data Sources

Income Source	Administrative Source	Income Unit	Income Frequency	States Covered
Earnings	DER (SSA)	Individual	Annual	All
Asset Income	Form 1040 (IRS)	Tax Unit	Annual	All
Retirement Distributions	Form 1099-R (IRS)	Individual	Annual	All
OASDI	PHUS (SSA)	Individual	Monthly	All
SSI	SSR (SSA)	Individual	Monthly	All
EITC	Form 1040 (IRS)	Tax Unit	Annual	All
SNAP	State Agencies	Household	Monthly	11 States
Housing Assistance	PIC & TRACS (HUD)	Household	Monthly	All

Note: This table shows – for each income source in the administrative data – the source of the data, the unit at which the dollar amounts are reported, the frequency at which the dollars are reported, and the states/years covered. Note that all of the administrative data, with the exception of SNAP, cover the universe of recipients in the United States.

Tax Records

Earnings data covering wage and salary jobs and self-employment are available from the Detailed Earnings Record (DER) database of the SSA. The DER itself is derived from IRS W-2 Forms (for wage and salary jobs) and Schedule SE of IRS 1040 Forms (for self-employment). The DER includes wage and salary earnings that are below the 1040 filings requirement, though it misses other sources of earnings that we note later. We also have data on various forms of asset income from IRS 1040 Forms, including taxable dividends and taxable and tax-exempt interest. Data on retirement distributions come from IRS 1099-R Forms, which cover gross distributions from employer-sponsored plans (defined benefit and defined contribution plans) and IRA withdrawals. Finally, we calculate eligible EITC amounts based on filing status, earned income, and qualifying dependents in the prior year's IRS 1040 Forms.

As Table 3.1 indicates, the tax data contain universe records spanning the entire United States. These data are at the level of the tax unit, which consists of a single individual or married couple filing together with any eligible dependents. Note that the tax unit is conceptually distinct from a household, even if the two units are equivalent for most people. Furthermore, we convert annual data from the administrative tax records to monthly amounts by dividing the total amounts by twelve and distributing them evenly across all months in the calendar year.

Program Participation Records

Administrative records for Social Security (OASDI) come from the SSA's Payment History Update System (PHUS) file, with our preferred measure of total benefits including any amounts that are deducted for medical insurance premiums. Data on Supplemental Security Income (SSI) come from the SSA's Supplemental Security Record (SSR) file and include all federally-administered payments that are initially split into federal payments and federally-administered state payments. OASDI and SSI benefits are paid to individuals on a monthly basis. For housing assistance, our administrative data come from the Public and Indian Housing Information Center (PIC) and Tenant Rental Assistance Certification System (TRACS) files. These data cover almost all public and subsidized housing assistance programs under the jurisdiction of HUD. We calculate the benefit amount for a household as the difference between the gross rent and actual tenant payment.[133] However, the records miss large housing programs such as those under the Department of Agriculture, which serve over a quarter million households (Scally and Lipsetz 2017), and those requiring landlords to charge below-market rents, which cover more than two million units (Scally et al. 2018). Supplemental Nutrition Assistance Program (SNAP) records come directly from various state agencies, and we have records for eleven states

[133] Because the administrative data do not include gross rent amounts for public housing units (which constitute less than a quarter of all households in the administrative data), we impute the market rent for these units based on the average rent by five-digit zip code, household size, and year. If rent is still missing, we impute by three-digit zip code, household size, and year – and subsequently by five-digit zip code and year and by three-digit zip code and year if needed. We consider a household to be receiving subsidized housing for the twelve months since the most recent certification date as long as the period is prior to any termination date.

in 2011. Housing assistance and SNAP benefits are recorded and disbursed to households on a monthly basis.

3.3.3 Linking Survey and Administrative Data

We link the administrative data to the SIPP using anonymized Protected Identification Keys (PIKs) created by the U.S. Census Bureau's Person Identification Validation System (PVS) (Wagner and Layne 2014). The PVS is based on a reference file containing Social Security Numbers, names, addresses, and dates of birth. Over 99% of most administrative records are associated with a PIK, and nearly 97% of households in Wave 9 of the SIPP contain at least one member associated with a PIK. To account for the likely small bias arising from non-random missing PIKs, we divide the household survey weights by the predicted probability that at least one member of the household has a PIK, conditional on observed characteristics in the survey (see Wooldridge 2007). This approach keeps the sample as comprehensive as possible at the expense of understating the income from administrative sources for household members who cannot be linked. The appendix provides a full discussion of the inverse probability weighting adjustment. We have also conducted bounding exercises using the survey-based extreme poverty rates of unlinked households, which are sufficiently close to the rates for linked households that the loss of 3% of the sample has a trivial impact. We link all benefit dollars from an administrative SNAP or housing case to a survey household as long as there is a common individual between each unit. For the EITC and asset income, we link only to individuals in survey households who are primary and secondary filers in the tax data.[134]

[134] If an administrative case links to multiple survey households, we distribute benefit dollars from the administrative case proportionally to the number of individuals linked to each household.

3.4 Methods

To begin, we define our baseline measure of extreme poverty (based on survey reports of cash income) and explain the decisions involved in constructing this measure. We then describe how we can improve on this reported measure using only the survey data. These adjustments involve incorporating non-medical in-kind transfers, undertaking conservative corrections for errors in reported earnings, and accounting for substantial assets. Next, we illustrate how bringing in the administrative data further improves the measurement of extreme poverty beyond what is possible in the survey. Lastly, we validate each of the survey-only adjustments by examining the administrative incomes and survey-reported material hardships, housing circumstances, and demographics of the groups removed from extreme poverty by the adjustments. Through confirming that those characterized as not extreme poor are well-off according to other indicators, the resulting measure reflects multiple dimensions of material well-being. As a check on our results, we examine the prevalence of those who are not in extreme poverty according to the survey but are reclassified as extreme poor after substituting in the administrative data. This group turns out to be miniscule.

3.4.1 Defining Extreme Poverty and Sample Construction

A number of different definitions of extreme poverty or “disconnectedness” have been used in the literature. As discussed in Section 3.2, one of the most well-known standards considers a household to be in extreme poverty if the household’s income is less than or equal to $2 per-person, per-day. Various papers use slightly different cutoffs or differ in what is included in income and the period over which income is measured, though most report results for multiple definitions.[135] We start from pre-tax money income, which includes earnings, asset and retirement income, cash transfers, and other money income that a household may receive.[136] This definition – which the Census Bureau's official poverty measure uses to calculate income – ignores in-kind transfers such as SNAP and tax credits like the EITC, though the SPM includes them. These sources of income have grown in importance over the last two decades. Subsequently, we show the degree to which a measure of extreme poverty based on this cash income definition (hereafter referred to as "reported extreme poor") holds up after various corrections and adjustments.

In this paper, we consider a household to be in extreme poverty if its average monthly income over the reference months in a wave is less than $2/person/day.[137] Empirically, we observe considerable overlap between observations with reported cash incomes below $2/person/day and who are "disconnected" based on various definitions from Blank and Kovak (2009), suggesting that our results likely generalize to analyses of disconnectedness.[138] Furthermore, we define extreme poverty at the *wave* level, though the results are very similar at the month level. There are several reasons to analyze extreme poverty at the wave level. First, a wave provides a more comparable time period than a month to use when linking with annual tax records. Second, to keep our results comparable to measures in other surveys such as the CPS, we want to use a single

[135] Deaton (2018) uses a cutoff of $4 in income after taxes and in-kind transfers per-person, per-day. Blank and Kovak (2009), in their study use three alternative definitions of disconnected single mother families: those with 1) no earnings or welfare receipt over an entire year, 2) less than $2000 in earnings and $1000 in cash welfare, or 3) the income in (2) plus annual SSI income less than $1000. United Nations (2018) relies on Deaton.

[136] "Other money income" can include sources like child support, income assistance from charitable groups, and money from friends or relatives.

[137] In Section 3.8, we also present results using a threshold of $4/person/day.

[138] Suppose we define being "disconnected" as having less than $166.67 in monthly earnings, $83.33 in cash welfare, and $83.33 in SSI income, which corresponds to one of the definitions in Blank and Kovak (2009) when the annual thresholds are converted to monthly values. Then, among single mother-headed households (the focus of the literature on disconnectedness), 86% with reported cash incomes below $2/person/day are disconnected while 42% that are disconnected have reported cash incomes below $2/person/day.

retrospective interview rather than the multiple interviews needed to construct calendar quarters or years.[139]

While previous analyses have focused on households with children (see Shaefer and Edin 2013, 2017), we study all households and investigate how the prevalence of extreme poverty differs across five mutually exclusive and exhaustive household types: households headed by someone aged 65 or older (elderly) and four non-elderly household types (single parent, multiple parents, single childless adults, and multiple childless adults).[140] This disaggregation is informative given that eligibility for transfer programs is often dependent on household type (e.g., being elderly, having children).

3.4.2 Corrections and Adjustments Using Survey and Administrative Data

We now describe the corrections and adjustments made using survey and administrative data to improve upon the reported extreme poverty rate. Our preferred specification implements the survey-only adjustments before bringing in the administrative data. We also present results using an alternative order that brings in the administrative data first.

Survey Data

Here, we describe the corrections and adjustments made using only publicly available survey data to improve upon the reported extreme poverty rate. We first incorporate the following in-kind transfers: SNAP, WIC, and housing assistance. Following the methodology in Shaefer and

---

[139] Starting with the 2014 Panel, the SIPP underwent a redesign where – among other changes – interviews are now conducted annually rather than every four months. However, the accuracy of the redesigned SIPP remains in question (National Academies of Sciences 2018).

[140] As we note in the appendix, there are some very rare cases where we classify individuals under 18 as adults – e.g., a 17-year-old single mother living on her own with her children.

Edin (2013) to account for in-kind transfers, we reclassify a household as not extreme poor if 1) its total cash income plus survey-reported values of SNAP and WIC benefits exceeds \$2/person/day or 2) it receives any form of housing assistance.[141]

Among those still in extreme poverty after incorporating in-kind transfers, we calculate lower-bound earnings based on survey reports of hours worked for pay, under the assumption that workers earn at least the federal minimum wage (\$7.25/hour). We then remove households from extreme poverty if the earnings resulting from this correction for missing dollars exceed \$2/person/day.[142] We first identify the households removed by lower-bound earnings using only reported wage and salary hours. Subsequently, we identify households removed by lower-bound earnings for reported hours worked in self-employment jobs as well. One might worry that this algorithm applies less well to off-the-books and/or self-employment jobs where the federal minimum wage does not apply. As a robustness check, we apply half the federal minimum wage to hours worked for pay, which leaves our final result unchanged. The vast majority of individuals removed by these corrections report a full set of employment characteristics but zero earnings, and they work in occupations typically paid above the minimum wage. This finding suggests that the zero earnings, rather than the positive hours worked (most of which are not imputed), are

[141] The assumption behind including housing assistance in this way is that the monetary value of public or subsidized housing is worth at least \$2/day per person. For this not to be true, the assistance amount for a two-person household would have to be less than \$120/month, which seems implausible. As a robustness check, we impute average housing assistance amounts from the administrative data based on county, household size, and year (and county and year if still missing) and designate a household as being lifted out of extreme poverty if its total cash income + survey-reported SNAP and WIC benefits + imputed housing assistance amount exceeds \$2/person/day. The results are identical.

[142] Alternatively, we could add all other (non-earnings) survey-reported income to these lower-bound earnings and compare the resulting amount to \$2/person/day. We choose to use our more conservative correction because the extreme poverty threshold is so low and, as is, very few hours are required to remove a household from extreme poverty. For example, a single person needs to work only 9 hours in a month to earn above \$2/day. In practice, whether or not we choose to add other survey-reported income to minimum wage earnings changes the results by only a few hundredths of a percentage point.

anomalous and thus recorded incorrectly.[143] Conversations with Census Bureau employees indicate that editing problems in the SIPP are responsible for at least a substantial share of these cases. As we later show, these issues appear to be unique to the SIPP and are not present in the CPS. The appendix provides a thorough discussion of how we calculate these lower-bound earnings and the occupations associated with households removed by these corrections.

Our last survey-only adjustment accounts for households holding substantial assets. Among those still left in extreme poverty after incorporating in-kind transfers and lower-bound earnings, we consider a household to not be in extreme poverty if its reported real estate equity exceeds $25,000, liquid assets exceed $5,000, or total net worth exceeds $50,000.[144] We include the restriction that households must have positive total assets to be reclassified by this adjustment. We obtain asset amounts from the topical modules to Waves 7 and 10 of the SIPP. While we acknowledge that assets are not part of cash or in-kind income, it seems inappropriate to consider households with sizeable assets that could be drawn upon to be in extreme poverty. We later show that the preponderance of these households would be removed by the administrative data corrections, but we feel it is important to show what can be done with the survey data alone.[145] It should be noted that the SPM accounts for assets in its thresholds, and the Haig-Simons definition

[143] Among all individuals aged 15 or above with zero earnings and positive wage/salary hours worked in Wave 9, 69.7% report an hourly wage, 95.4% of all hourly wage reports are above the federal minimum wage, and 99.6% of all hourly wage reports are above half the federal minimum wage (per the public use data). This implies that our minimum wage assumption is sensible and – if anything – an underestimate. We further verify that the vast majority of households removed by these corrections have incomes above $2/person/day from the administrative records, and they resemble the average household in the U.S. on various measures of reported well-being.

[144] Real estate equity includes home equity and equity in any other real estate, including mobile homes. Liquid assets include checking accounts, savings accounts, money market accounts, bonds, securities, mutual funds, debt or margin accounts, certificates of deposits, and stocks. Total net worth equals total assets (liquid assets, retirement accounts, real estate equity, vehicle equity, business equity, and the value of other financial investments - i.e., “equity in other assets”) minus secured and unsecured debt.

[145] In the end, the adjustment for substantial assets increases our final extreme poverty rate by only 0.13 percentage points. This is because 0.46% of households are removed from extreme poverty by the assets adjustment (Table 3.2a), and 28.2% of these households have incomes below $2/person/day in the administrative data (Figure 3.3). Multiplying these numbers results in 0.13% of all households that are removed from extreme poverty by the assets adjustment (our final survey-only adjustment before bringing in the administrative data) but not the administrative data.

of income – along with authoritative sources on poverty measurement like Ruggles (1990) and Citro and Michael (1995) – explicitly recognizes that not accounting for assets is problematic. In its "Guide on Poverty Measurement", the United Nations Economic Commission for Europe (2017) states that "Owning your own house or apartment in effect provides you with housing services, which should be considered as part of both income and consumption" (p. 50). Asset tests are also part of the eligibility requirements for most means-tested programs, and Ruggles (1990) states that a primary reason for not accounting for assets in poverty measurement is simply that most surveys do not ask about assets (p. 149).

Administrative Data

Given the underreporting of many types of income such as government transfers and private pensions, we bring in administrative data to further refine the extreme poverty rate. Among those still in extreme poverty after all survey-only adjustments, we consider households to not be extreme poor if their incomes exceed $2/person/day after replacing survey reports with administrative measures for various income sources. The administrative data can help account for false negatives among recipients of transfer programs and gross errors in reported amounts, among other survey errors. We first replace survey reports of earnings, interest and dividends, and retirement distributions with values from administrative tax records.[146] We then add EITC amounts

[146] Since survey reports may cover earnings from off-the-books and non-standard jobs that are not reported to the IRS (see Abraham et al. 2013, 2020), it may instead be justified to take the maximum of survey and administrative reports of earnings. By simply replacing survey reports of earnings with administrative tax earnings, we run the risk of trivially overstating the extreme poverty rate. In practice, it does not matter for this analysis whether or not we take the maximum of survey and administrative values. Since the survey values must lie below $2/person/day at the stage when we bring in the administrative data, there is effectively no difference between taking just 1) the administrative earnings values and 2) the maximum of the administrative earnings values and minuscule survey values (the vast majority of which are zero).

calculated from tax records.[147] We hereafter refer to these income sources collectively as "tax data income".

Next, we replace survey reports of OASDI, SSI, housing assistance, and SNAP with values from administrative program records.[148] We hereafter refer to these income sources collectively as "transfer income". We are able to directly incorporate OASDI, SSI, and housing assistance for all states, while we use administrative SNAP records for 11 states (covering 29% of the population) to estimate the effect of the administrative SNAP data for all states. Specifically, we calculate our final estimate of the extreme poverty rate by multiplying the rate after accounting for the survey-only adjustments, the administrative tax data, and the administrative non-SNAP transfer data (calculated over all 50 states) by the fraction of such households in the 11 states remaining in extreme poverty after bringing in the administrative SNAP data. By taking this approach, we need only assume that these 11 states are representative of the entire country in the marginal impact of the administrative SNAP data, which is weaker than assuming they are representative in the level of extreme poverty. However, these 11 states are indeed similar to the rest of the country on a number of demographic and economic characteristics.[149] As we later show, whether or not we include the administrative SNAP data at all makes very little difference for our final results.

[147] We use eligible EITC amounts calculated based on the administrative tax data rather than actual amounts. While this will overstate the true amount of the EITC disbursed, the upward bias associated with eligible EITC amounts is likely substantially smaller than the downward bias in survey-based imputations of EITC amounts. We find that the actual EITC dollars disbursed (from publicly available IRS totals) are 90% of the total eligible dollars that we calculate. In contrast, aggregate EITC dollars imputed in the CPS understate disbursements by approximately 30% (Meyer 2017).

[148] For SSI, we only have administrative data on federally-administered benefits, even though states can separately administer benefits themselves. Thus, our preferred measure of total SSI benefit amounts sums administrative values for federally-administered SSI and survey values for state-administered SSI.

[149] Households in the SNAP states have extreme, deep, and official poverty rates (measured based on cash income) that are insignificantly different from the rates in the full sample. Fewer households in the SNAP states receive OASDI and SSI, and more receive SNAP, Public Assistance, and housing assistance than those in the full sample, though only the differences for SSI, SNAP, and housing assistance are statistically significant at the 5% level.

We also employ an alternative ordering of our main results that incorporates the administrative data first. Among those with survey-reported incomes below $2/person/day, we start by replacing survey reports of earnings with administrative earnings, before bringing in administrative data on other sources of cash income, including asset income, retirement distributions, OASDI, SSI, and the EITC. We then bring in administrative program values for housing assistance and SNAP (in that order), using once again a proportional adjustment for SNAP calculated using the 11 states for which we have administrative SNAP data.

3.4.3 Validating the Survey-Only Adjustments

We recognize that our adjustments using only the publicly available survey data are imperfect. For example, some earnings, such as those from self-employment or off-the-books, are not subject to minimum wage legislation. Moreover, the survey reports of hours worked and assets may themselves be misreported. Consequently, we thoroughly validate the appropriateness of each survey-only adjustment using information from the administrative data and detailed measures of well-being from the SIPP topical modules. Poverty definitions should be examined to see how they accord with other indicators of disadvantage, but – in practice – measures are typically chosen for other reasons without such validation.[150] First, for each subgroup of the reported extreme poor removed by a survey-only adjustment, we directly calculate the share of households with incomes above $2/person/day after replacing survey values with the available administrative data values for various income sources. To investigate the extent of gross errors, we also calculate the share

[150] Exceptions include Mayer and Jencks (1989) and Meyer and Sullivan (2003, 2011, 2012a).

of households in each subgroup with incomes above half the poverty line (the deep poverty cutoff), the poverty line, and twice the poverty line based on the administrative data.[151]

As a second check on the validity of the survey-only adjustments, we compare the groups removed from extreme poverty to the official poor and all households based on survey-reported measures of hardship and housing quality.[152] Conveniently, these measures of material well-being are collected from the same interview wave (Wave 9) as the income measures. For material hardships, we examine survey answers (yes/no) to nine separate questions on a range of hardships, including not being able to pay all essential expenses, rent, mortgage, or an energy bill, having energy or telephone service disconnected, being evicted, inability to see a doctor or dentist, and a lack of food. We also examine ownership of the following eight appliances: microwaves, dishwashers, air conditioners, color televisions, computers, in-unit washers, in-unit dryers, and cell phones.[153] We further investigate whether a household faces any of seven housing quality issues, including problems with pests, a leaking roof, broken windows, exposed wires, plumbing problems, and cracks or holes in the walls, ceiling, or floors. An advantage of examining material hardships and housing problems is that these measures may be more indicative of deprivation, while an advantage of examining appliances is that they can be easily and objectively measured. The appendix provides a more detailed description of the specific hardship and material well-being variables used. Finally, we assess additional demographic and economic characteristics reported in the SIPP – such as student status, educational attainment, health insurance coverage, and asset ownership – to obtain an even better picture of each group removed from extreme poverty.

---

[151] For a single non-elderly individual, the average monthly poverty line in Wave 9 corresponds to $32.15/person/day.

[152] For the analyses on survey-reported well-being and demographics, we use the full survey sample and original survey weights (as opposed to the PIKed sample and adjusted survey weights).

[153] We exclude certain appliances (refrigerators, freezers, stoves, and regular telephones) that we think households are likely to own regardless of their material circumstances, precisely in an effort to capture those appliances that are most strongly indicative of well-being.

3.5 Main Results

3.5.1 Extreme Poverty After Adjustments

Table 3.2a displays the share of households that are left in extreme poverty after successively incorporating each adjustment.[154] The first column starts with survey-reported cash income and finds that 2.97% of all households report having less than $2/person/day of cash income.[155] However, nearly a third of these households are reclassified as not extreme poor by survey-reported in-kind transfers, with the extreme poverty rate for households falling to 2.04%. Nearly 95% of the impact of survey-reported in-kind transfers is attributable to survey-reported SNAP. Correcting for errors in reported earnings based only on reported wage and salary hours worked for pay decreases the extreme poverty rate for households to 1.83%. Further accounting for reported self-employment hours worked decreases the extreme poverty rate for households to 1.30%. All told, correcting for errors in reported earnings removes an additional 36% of households from extreme poverty. Accounting for substantial assets again reduces the extreme poverty rate by over a third, leaving us with 0.84% of households remaining in extreme poverty.

While adjustments using only the survey data eliminate most of extreme poverty, the administrative tax and program data provide additional information. Applying the administrative earnings data alone removes an additional 50% of those remaining in extreme poverty and cuts the extreme poverty rate to 0.42%. Incorporating the administrative data on asset and retirement income decreases the extreme poverty rate to 0.35%, and adding the EITC further reduces the rate

[154] Standard errors are calculated using replicate weights corresponding to Wave 9 of the 2008 SIPP.

[155] The reported extreme poverty rate of 2.97% calculated using the PIKed subsample and adjusted survey weights is nearly identical to the reported extreme poverty rate of 3% calculated using the entire sample and original survey weights.

Table 3.2a. Percentage of Households in Extreme Poverty

	All Households (1)	Elderly (2)	Single Parents (3)	Multiple Parents (4)	Single Childless Adults (5)	Multiple Childless Adults (6)
Survey-Reported Cash	2.97	0.46	8.99	2.04	6.85	1.90
	(0.13)	(0.08)	(0.91)	(0.17)	(0.44)	(0.15)
Survey-Only Adjustments						
Add In-Kind Transfers	2.04	0.42	2.80	1.16	5.58	1.58
	(0.11)	(0.08)	(0.41)	(0.13)	(0.41)	(0.14)
Correct Wage/Salary Earnings	1.83	0.37	2.66	0.94	5.12	1.39
	(0.10)	(0.08)	(0.40)	(0.12)	(0.39)	(0.13)
Correct Self-Emp. Earnings	1.30	0.35	1.97	0.53	4.04	0.75
	(0.08)	(0.08)	(0.38)	(0.08)	(0.32)	(0.09)
Account for Assets*	0.84	0.13	1.35	0.27	2.86	0.44
	(0.07)	(0.05)	(0.32)	(0.06)	(0.31)	(0.07)
Administrative Data Adjustments						
Correct Earnings	0.42	0.11	0.54	0.11	1.63	0.08
	(0.05)	(0.04)	(0.20)	(0.04)	(0.22)	(0.03)
Correct Assets/Retirement Inc.	0.35	0.06	0.46	0.11	1.35	0.08
	(0.04)	(0.03)	(0.19)	(0.04)	(0.19)	(0.03)
Add EITC	0.31	0.06	0.10	0.08	1.29	0.08
	(0.04)	(0.03)	(0.07)	(0.03)	(0.19)	(0.03)
Correct OASDI/SSI	0.27	0.01	0.10	0.06	1.19	0.07
	(0.04)	(0.01)	(0.07)	(0.03)	(0.18)	(0.03)
Correct Housing Assistance	0.27	0.01	0.10	0.06	1.18	0.07
	(0.04)	(0.01)	(0.07)	(0.03)	(0.18)	(0.03)
Correct SNAP	0.24	0.00	0.00	0.00	1.12	0.07
	(0.04)	(0.00)	(0.00)	(0.00)	(0.18)	(0.03)
Population Estimates (millions)						
U.S.	118.6	26.1	6.9	31.7	22.5	31.4
SNAP States	34.4	7.4	2.0	9.3	6.6	9.0
Sample Sizes						
U.S.	31,500	8,000	1,500	8,300	5,200	8,500
SNAP States	10,000	2,500	500	2,700	1,600	2,700

*Real estate equity > $25,000, liquid assets > $5,000, or total net worth > $50,000.
Sources: Wave 9 of the 2008 SIPP Panel, spanning January-July 2011. Administrative data sources described in text.
Note: Standard errors calculated using replicate weights in parentheses. Households in "extreme poverty" are those with average income across the four months of the wave less than or equal to $2/person/day. Sample consists of households with at least one member with a PIK and present in reference month 4. Reference month 4 survey weights are adjusted for missing PIKs. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization numbers: CBDRB-FY18-324 and CBDRB-FY19-173.

Table 3.2b. Percentage of Individuals in Extreme Poverty

Specification	All Households (1)	Elderly (2)	Single Parents (3)	Multiple Parents (4)	Single Childless Adults (5)	Multiple Childless Adults (6)
Survey-Reported Cash	2.60	0.47	9.56	2.11	6.85	1.83
	(0.14)	(0.21)	(0.98)	(0.17)	(0.44)	(0.15)
Survey-Only Adjustments						
Add In-Kind Transfers	1.57	0.43	2.65	1.20	5.58	1.50
	(0.09)	(0.21)	(0.45)	(0.14)	(0.41)	(0.14)
Correct Wage/Salary Earnings	1.37	0.36	2.53	0.95	5.12	1.33
	(0.09)	(0.18)	(0.44)	(0.12)	(0.39)	(0.14)
Correct Self-Emp. Earnings	0.90	0.35	1.93	0.52	4.04	0.75
	(0.06)	(0.18)	(0.45)	(0.09)	(0.32)	(0.11)
Account for Assets*	0.57	0.11	1.49	0.28	2.86	0.46
	(0.05)	(0.04)	(0.45)	(0.06)	(0.31)	(0.10)
Administrative Data Adjustments						
Correct Earnings	0.24	0.09	0.64	0.10	1.63	0.09
	(0.03)	(0.04)	(0.28)	(0.04)	(0.22)	(0.04)
Correct Assets/Retirement Inc.	0.21	0.06	0.56	0.10	1.35	0.09
	(0.03)	(0.03)	(0.25)	(0.04)	(0.19)	(0.04)
Add EITC	0.17	0.06	0.12	0.07	1.29	0.09
	(0.03)	(0.03)	(0.10)	(0.03)	(0.19)	(0.04)
Correct OASDI/SSI	0.14	0.01	0.12	0.05	1.19	0.07
	(0.02)	(0.01)	(0.10)	(0.02)	(0.18)	(0.04)
Correct Housing Assistance	0.14	0.01	0.12	0.05	1.18	0.07
	(0.02)	(0.01)	(0.10)	(0.02)	(0.18)	(0.04)
Correct SNAP	0.11	0.00	0.00	0.00	1.12	0.07
	(0.02)	(0.00)	(0.00)	(0.00)	(0.18)	(0.04)
Population Estimates (millions)						
U.S.	305.6	45.5	18.8	138.3	22.5	80.5
SNAP States	88.6	13.0	5.4	40.0	6.6	23.6
Sample Sizes						
U.S.	82,500	14,500	4,500	37,000	5,500	21,000
SNAP States	26,500	4,500	1,400	12,000	1,700	6,700

*Real estate equity > $25,000, liquid assets > $5,000, or total net worth > $50,000.

Sources: Wave 9 of the 2008 SIPP Panel, spanning January-July 2011. Administrative data sources described in text.

Note: Standard errors calculated using replicate weights in parentheses. Individuals in "extreme poverty" are in households with average income across the four months of the wave less than or equal to $2/person/day. Sample consists of individuals in households with at least one member with a PIK and in reference month 4. Reference month 4 survey weights are adjusted for missing PIKs. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization numbers: CBDRB-FY18-324 and CBDRB-FY19-173.

to 0.31%. Bringing in administrative data on OASDI and SSI lowers the extreme poverty rate to 0.27%, which decreases insignificantly to 0.24% after bringing in the administrative housing assistance and SNAP data.[156] Of the additional households removed from extreme poverty by the administrative data, 67% have incomes above half the poverty line and 55% have incomes above the poverty line.[157] This finding suggests that there are still non-trivial gross errors in the extreme poverty rate after the survey-only adjustments. Together, the adjustments reduce extreme poverty by 92% from a reported rate of 2.97%, with more than three-quarters of the total reduction due to corrections and adjustments using the survey data alone.

We observe a similar pattern for individuals (Table 3.2b), with in-kind transfers cutting the extreme poverty rate the most and each of the other adjustments also removing a sizable portion of individuals out of extreme poverty. When looking at only reported cash income, we find that 2.60% of individuals live on less than $2/day. After accounting in the survey for in-kind transfers, reported hours worked, and substantial assets, the extreme poverty rate falls by more than three-quarters to 0.57%. Bringing in the administrative tax and transfer data further reduces the extreme poverty rate to 0.11%. The extreme poverty rates for individuals are lower than those for households because extreme poor households tend to have fewer members.

3.5.2 Corrections and Adjustments Bringing in the Administrative Data First

Figure 3.1 illustrates the results of an alternative order to the adjustments that incorporates the administrative data first. We can see that 55.1% of households initially classified as having

---

[156] Standard errors at the final step (bringing in administrative SNAP values) are calculated using the delta method.

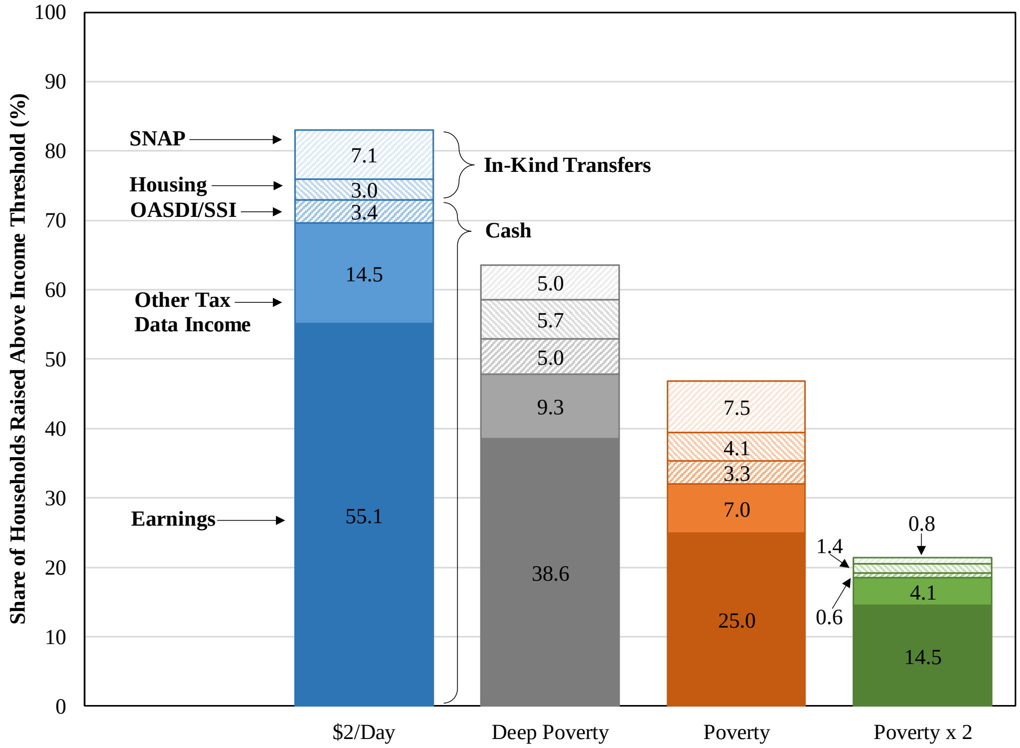
[157] These numbers are calculated from Figure 3.3. Among the remaining extreme poor (after survey-only adjustments), 71% and 47.9% have incomes above the extreme and deep poverty lines, respectively. Because those with incomes above the deep poverty line are a subset of those with incomes above the extreme poverty line, we can calculate that 67% (49.1% divided by 72.3%) of those above the extreme poverty line are also above the deep poverty line. A similar logic follows for the poverty line.

incomes below $2/person/day based on survey cash reports have incomes above $2/person/day from the administrative earnings records alone. In other words, correcting earnings using the administrative data singlehandedly decreases the extreme poverty rate from a survey base of 2.97% to 1.33%. Bringing in the other administrative sources of money income shows that a full 73% of reported extreme poor households are misclassified due simply to errors in cash reports of earnings, asset income, retirement distributions, OASDI, and SSI and the omission of the EITC. This change further brings down the extreme poverty rate to 0.80%. When we add in administrative sources of in-kind transfers (housing assistance and SNAP), 83.1% of those originally classified as in extreme poverty have incomes above the extreme poverty threshold. In sum, this implies that the administrative data alone can decrease the extreme poverty rate from a reported cash base of 2.97% to 0.50%. The adjustments using only the survey data then take the extreme poverty rate down by an additional 0.26 percentage points to 0.24%. Thus, when incorporated first, the administrative data can account for 90% of the change in extreme poverty due to all adjustments.

Furthermore, nearly half of all households initially classified as in extreme poverty based on survey cash reports have incomes above the poverty line and over a fifth have incomes above twice the poverty line. This finding makes it clear that there is a vast amount of error associated with classifying households as extreme poor based solely on their survey-reported cash income. We can also show from Figure 3.1 that 79% of all misclassified households are initially categorized as extreme poor due to errors in reports of cash income including earnings, retirement income, OASDI, and SSI and the omission of the EITC.[158] This share must be a lower bound for those misclassified due to all errors in cash income, because the Detailed Earnings Record or DER (our

[158] To calculate this share, note that 92% of all reported cash extreme poor households are misclassified (since the adjustments decrease the rate from a base of 2.97% to 0.24%). We also know from Figure 3.3 that 73% of reported

Figure 3.1. Share of Reported Cash Extreme Poor Households Raised Above Income Thresholds by Administrative Data



Sources: Wave 9 of the 2008 SIPP Panel, spanning January-July 2011. Administrative data sources described in text.
Note: Sample consists of households initially classified as having incomes below $2/person/day based on survey-reported cash income, with at least one member with a PIK and present in reference month 4. Reference month 4 survey weights are adjusted for missing PIKs. Other administrative tax data income includes asset income (taxable dividends, taxable and tax-exempt interest), retirement distributions (gross distributions from employer-sponsored plans and IRA withdrawals), and the EITC. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization numbers: CBDRB-FY18-324 and CBDRB-FY19-173.

source of administrative earnings data) misses the income of low-paid household workers and undocumented immigrants, as well as other income such as tips not reported to an employer. The DER also records only Medicare-taxable self-employment earnings, which is defined as 92.35% of total net self-employment income minus health insurance costs, and it additionally misses off-the-books income. Consequently, in-kind transfers play a secondary role relative to errors in cash reports in explaining the high reported extreme poverty rate.

---

extreme poor households are misclassified due to errors in cash reports of tax data income, OASDI, and SSI and the omission of the EITC. Because these households are a subset of all misclassified households, we divide 73% by 92% to obtain the share of all misclassified households initially among the reported cash extreme poor due to errors or omissions in cash reports.

3.5.3 Extreme Poverty by Household Type

We now analyze how extreme poverty differs by household type. Shaefer and Edin focus on two of the five household types (those with children), while the "disconnected families" literature focuses on single parent households. We first consider elderly households, who tend to have a significantly lower extreme poverty rate than other household types.[159] The elderly begin with a reported extreme poverty rate of 0.47%, less than one-sixth the rate for all households (Table 3.2a, Column 2). After incorporating each of the survey-only adjustments, 0.11% of elderly households remain in extreme poverty. Bringing in the administrative tax and non-SNAP transfer data removes nearly 90% of the remaining households from extreme poverty. Perhaps not surprisingly for these elderly households, the role of the administrative data is driven entirely by improved measures for three income sources: retirement distributions, OASDI, and SSI (Table 3.3, Column 2). The final estimate of the elderly extreme poverty rate is 0.01% prior to bringing in administrative SNAP records, and it becomes zero after incorporating SNAP.

We next consider single parent households (Table 3.2a, Column 3), whose reported extreme poverty rate of 8.99% is more than three times and statistically significantly above the rate for all households. However, about two-thirds of single parent households are reclassified as not extreme poor by survey-reported in-kind transfers. The extreme poverty rate for single parent households then declines to 1.97% and 1.35% after correcting for underreported earnings and accounting for substantial assets in the survey, respectively. After bringing in the administrative tax data, the extreme poverty rate for single parents falls to 0.1%. Most of this reduction is due to administrative earnings and the EITC (calculated from prior year earnings). After including the administrative SNAP data, no single parent households remain in extreme poverty. Among single

[159] The only exception to this rule is that the elderly do not have a significantly lower extreme poverty rate than multiple parent households after the self-employment correction.

parent households in the remaining extreme poor after the survey-only corrections, 75% have positive earnings from the tax records and 69% receive at least one transfer – usually SNAP or the EITC – per the administrative data (Table 3.3, Column 3).

Unlike single parent households, multiple parent households start with a reported extreme poverty rate of 2.04% significantly below that of all households (Table 3.2a, Column 4). In-kind transfers noticeably decrease their extreme poverty rate by 43%, and the subsequent adjustments for reported hours worked and assets bring down their extreme poverty rate to 0.27%. Like single parents, multiple parent households have an estimated extreme poverty rate of zero after incorporating the administrative data. This impact of the administrative data is driven again by the role of earnings and transfers, with more than 70% of the remaining extreme poor after survey-only adjustments having positive earnings and 79% receiving a transfer (Table 3.3, Column 4).

Households containing a single non-elderly individual have a reported extreme poverty rate of 6.85%, which – while lower than that of single parent households – is still 2.3 times and statistically significantly higher than that of all households (Table 3.2a, Column 5). Single individuals are not nearly as impacted by the adjustments and therefore have the highest extreme poverty rate of any household type after every adjustment. We are left with 2.86% of single individual households in extreme poverty after the survey-only adjustments, which is almost as high as the overall reported extreme poverty rate. Bringing in the administrative data also has a smaller effect on single individuals, removing just 61% of single individuals remaining in extreme poverty after the survey-only adjustments. This relatively smaller reduction is due to several factors. First, single individuals appear to be particularly disconnected from the safety net, with

Table 3.3. Income Receipt Rates (%) for Remaining Extreme Poor Households After Survey-Only Adjustments

Specification	All Households (1)	Elderly (2)	Single Parents (3)	Multiple Parents (4)	Single Childless Adults (5)	Multiple Childless Adults (6)
Earnings	55.93	15.41	74.76	70.15	47.92	81.24
	(3.34)	(14.97)	(10.01)	(11.57)	(4.29)	(7.05)
Asset Income	9.77	15.41	0.00	6.67	9.17	19.61
	(2.30)	(14.97)	(0.00)	(3.73)	(2.89)	(7.93)
Retirement Distributions	11.75	43.08	6.59	3.90	12.75	7.88
	(2.25)	(18.77)	(7.07)	(4.04)	(3.40)	(4.06)
OASDI	6.14	58.50	0.00	7.18	5.15	1.58
	(1.69)	(19.83)	(0.00)	(5.04)	(2.06)	(1.71)
SSI	2.83	35.47	0.00	14.46	0.00	2.69
	(1.18)	(19.21)	(0.00)	(9.00)	(0.00)	(2.20)
Housing Assistance	3.55	20.60	1.11	11.00	2.79	0.00
	(1.47)	(15.54)	(1.08)	(9.69)	(1.69)	(0.00)
EITC	30.19	0.00	67.95	73.34	19.19	36.18
	(3.40)	(0.00)	(11.41)	(9.18)	(3.74)	(11.38)
SNAP	19.06		61.40	73.40	3.80	21.79
	(4.26)		(14.34)	(19.23)	(2.49)	(11.19)
Any Transfer	42.68	89.51	69.26	78.73	30.97	51.40
	(6.81)	(11.12)	(20.81)	(20.08)	(8.14)	(18.07)

Sources: Wave 9 of the 2008 SIPP Panel, spanning January-July 2011. Administrative data sources described in text.
Note: Standard errors calculated using replicate weights in parentheses. These shares reflect the percent of households among the remaining extreme poor after the survey-only adjustments that receive each source of income in the administrative data. Sample consists of households with at least one member with a PIK and present in reference month 4. Reference month 4 survey weights are adjusted for missing PIKs. All income sources are calculated over all 50 states except for SNAP and "Any Transfer", which are calculated over the 11 states for which we have administrative SNAP data. For the elderly, we omit SNAP from the "Any Transfer" category. Approved for release by the Census Bureau's Disclosure Review Board, authorization numbers CBDRB-FY18-324 and CBDRB-FY19-173.

only 31% of the remaining extreme poor after the survey-only corrections receiving at least one transfer.[160] Second, the majority of these remaining extreme poor single individuals do not have earnings (Table 3.3, Column 5). We are therefore left with a final extreme poverty rate of 1.12% for non-aged single individuals.

Multiple childless adult households have a reported extreme poverty rate (1.90%) not far from that of multiple parent households (Table 3.2a, Column 6). The survey adjustments for in-

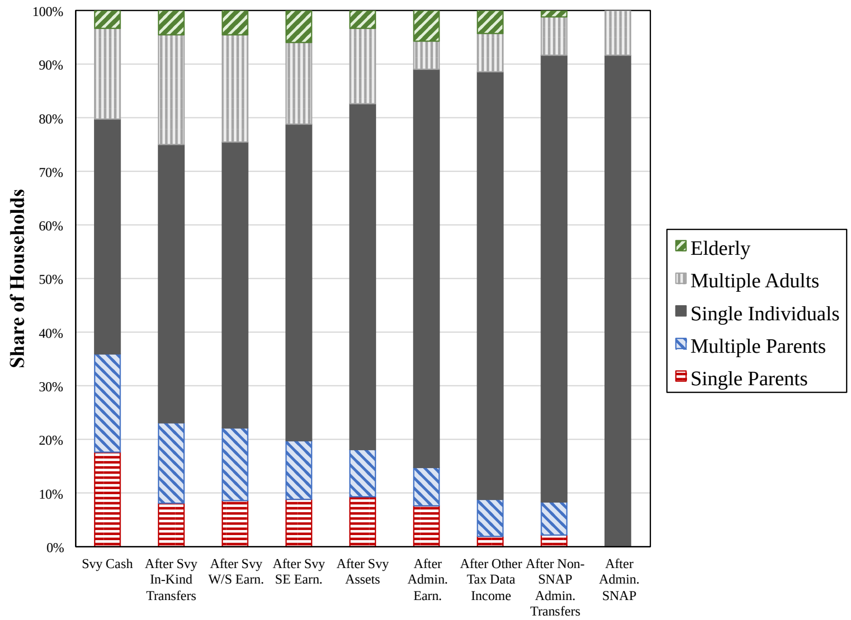
[160] Tests of differences across household types are often indecisive when restricted to the smaller set of states with SNAP data. Focusing on transfers besides SNAP and those remaining in extreme poverty after the survey corrections, single individuals have a significantly lower receipt rate than all other household types except multiple childless adults.

kind transfers, reported hours worked, and substantial assets together decrease their extreme poverty rate by more than three-quarters to 0.44%. After adding in administrative tax and transfer data, the extreme poverty rate for multiple childless adults becomes 0.07%. Among the remaining extreme poor after survey-only adjustments, multiple childless adults have far more earnings than single individuals (Table 3.3, Column 6) and have a high in-kind transfer receipt rate. Consequently, 85% of multiple childless adult households among the remaining extreme poor after survey-only adjustments are removed by the administrative data (compared to only 61% of single individuals). In summary, the combined survey and administrative data indicate that extreme poverty is extremely rare for the elderly, families with children, and multiple childless adults.

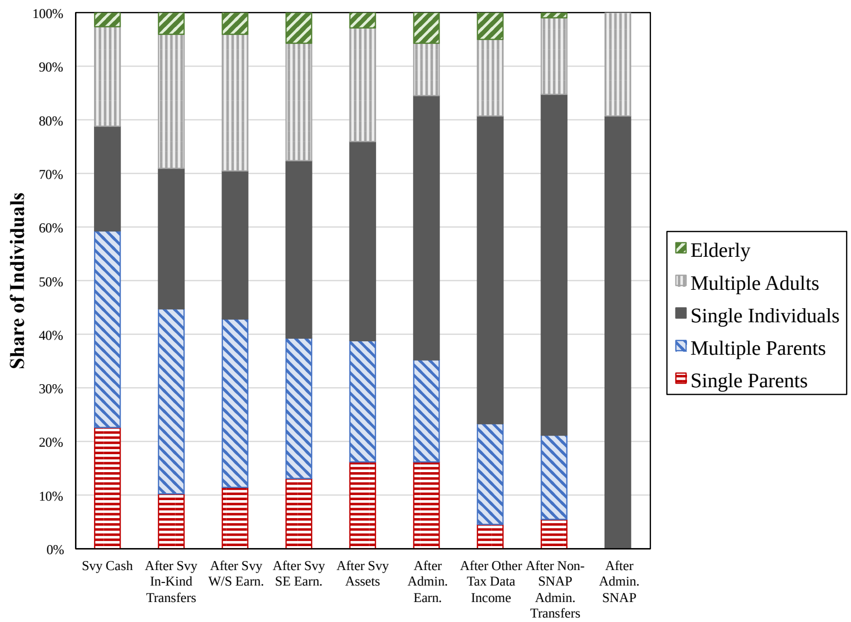
3.5.4 Distribution of Household Types

Not only do the errors in the income data exaggerate the level of extreme poverty, but they also lead to a distorted image of the type of households likely to be at the very bottom. Among the reported extreme poor, single individuals make up the largest share at nearly 44% (Figure 3.2a). Households with children form the next largest shares, with single and multiple parent households together making up about 36% of the reported cash income extreme poor (about 18% each). Multiple childless adult households also make up about 17% of the reported extreme poor, while elderly households contribute only a little over 3%. However, as we add each adjustment, single individual households constitute an increasingly larger share. After the survey-only adjustments, they constitute 65% of the remaining extreme poor households. Once we incorporate the administrative tax and non-SNAP transfer data, single individuals make up more than 83% of all

Figure 3.2. Household Type Distribution of Extreme Poor Subgroups After Adjustments



(a) Share of Households



(b) Share of Individuals

Sources: Wave 9 of the 2008 SIPP Panel, spanning January-July 2011. Administrative data sources described in text.
Note: Sample consists of households with at least one member with a PIK and present in reference month 4. Reference month 4 survey weights are adjusted for missing PIKs. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization numbers: CBDRB-FY18-324 and CBDRB-FY19-173.

extreme poor households, with this share rising to nearly 92% after bringing in the administrative SNAP data.

While single individuals constitute a disproportionate share of the households in extreme poverty, we may also want to consider the composition of extreme poverty in terms of the share of individuals living in extreme poor households (Figure 3.2b). When analyzing extreme poverty at the individual level, we find that only about 19% of the reported extreme poor are single individuals, while 59% are members of households with children (about 23% single parent and 36% multiple parent). Multiple childless adults make up another 19%, and the elderly contribute 2.7%. Nonetheless, we see with each adjustment the same, albeit less dramatic, pattern that we saw with households, as single individuals make up an ever-larger share of extreme poor individuals. Specifically, single individuals make up 37% of the remaining extreme poor individuals after the survey-only adjustments, and they make up almost 81% of extreme poor individuals after all adjustments.

3.6 Validation of Survey-Only Adjustments

We now describe the results obtained from validating each of the survey-only adjustments through comparisons to administrative income data and survey reports of material well-being and selected demographics. Throughout this subsection, the "remaining extreme poor" subgroup refers to households that are left in extreme poverty after the survey-only adjustments.

3.6.1 Administrative Income Data

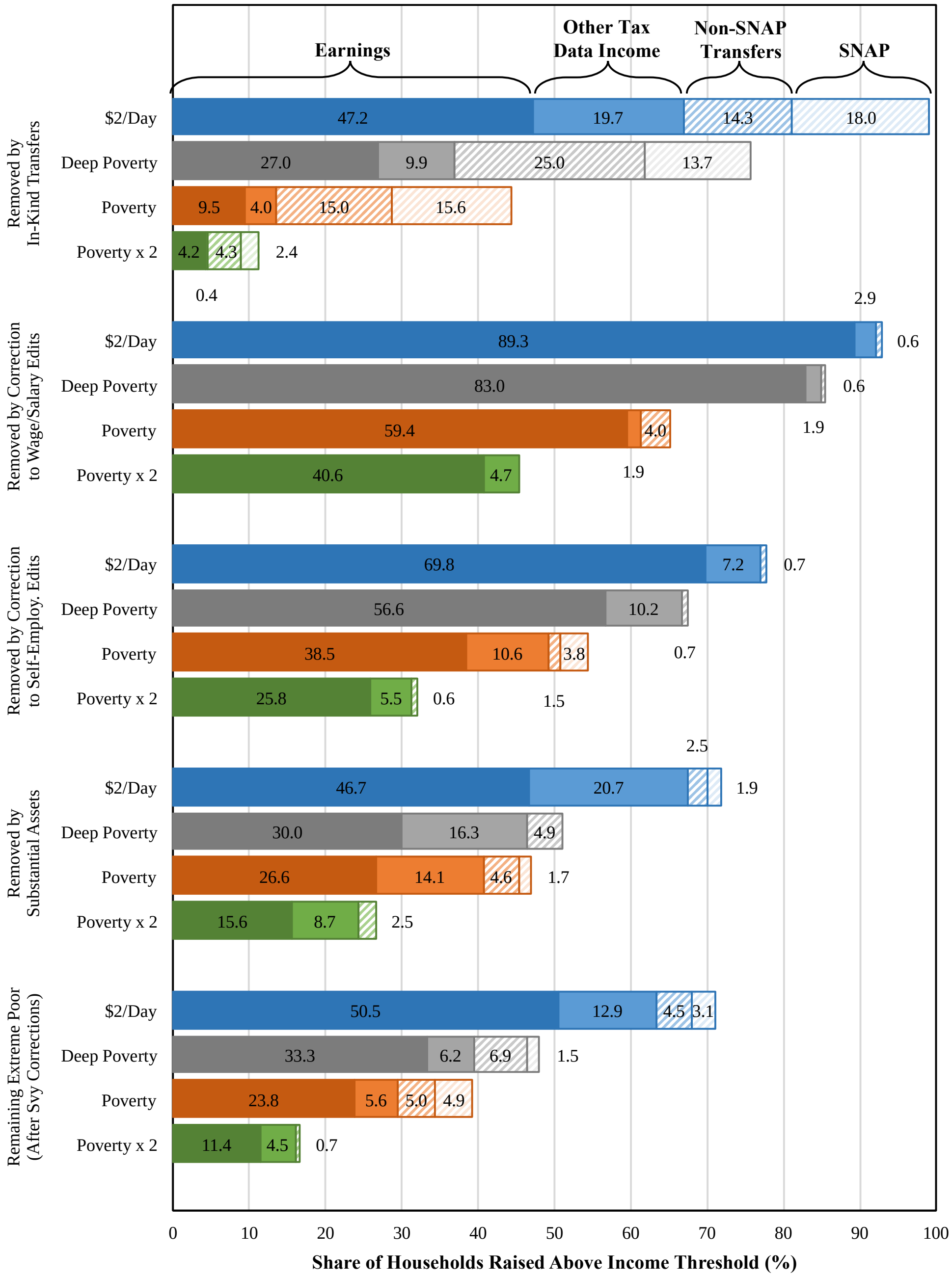
We first examine the share of households in each extreme poor subgroup with incomes above $2/person/day (and other thresholds) according to the administrative data (see Figure 3.3).

While the administrative data are our most accurate source for many income components, they still have important gaps. Yet, we are able to confirm the vast majority of our corrections and adjustments with these incomplete administrative data.

First, for the households reclassified as not extreme poor by survey-reported in-kind transfers, over 99% have incomes above \$2/person/day according to the administrative tax and transfer data. As expected, the administrative transfer data play a relatively large role for this subgroup, with nearly a third of its households removed from extreme poverty by administrative transfers after accounting for income from the administrative tax records. The tiny percentage of households not raised above \$2/person/day by the administrative data may be due to incomplete administrative data, incomplete linkage, or survey false positives. The correction for reported wage and salary hours is similarly robust, with 93% of the households removed by this correction having incomes from the administrative data confirmed above the extreme poverty threshold. Convincingly, 89% of these households are removed from extreme poverty by the administrative earnings alone. This subgroup also appears to have substantial gross errors, with 65% of households having incomes above the poverty line and 45% having incomes above twice the poverty line.

There is slightly less confirmation of the corrections for reported self-employment hours and substantial assets. Among the households removed due to reported self-employment hours, 70% have incomes above \$2/person/day based on administrative earnings alone and 78% have incomes above the threshold using all administrative tax and transfer data. While these shares are still large, they are smaller than those for the groups reclassified by in-kind transfers and wage and salary hours. This discrepancy could be due in part to the underreporting of self-employment

Figure 3.3. Share of Households in Extreme Poor Subgroups Raised Above Income Thresholds by Administrative Data



Sources: Wave 9 of the 2008 SIPP Panel, spanning January-July 2011. Administrative data sources described in text.
Note: Sample consists of households with at least one member with a PIK and present in reference month 4. Reference month 4 survey weights are adjusted for missing PIKs. Other administrative tax data income includes asset income (taxable dividends, taxable and tax-exempt interest), retirement distributions, and EITC. Non-SNAP transfers include OASDI, SSI, and housing assistance. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization numbers: CBDRB-FY18-324 and CBDRB-FY19-173.

earnings on tax returns (Internal Revenue Service 2016). The error is unlikely to result from minimum wage earnings being too high of a lower bound for self-employment earnings, given the relatively high-earning self-employment occupations and industries reported for this group and the similarity of the results using half the minimum wage (Table 3.5). Among the households removed from extreme poverty due to substantial assets, 67% are not extreme poor based on the administrative tax data and 72% are not extreme poor based on the administrative tax and transfer records. However, note once again the high fraction of gross errors in this subgroup, with 47% of households having incomes above the poverty line and 27% above twice the poverty line.

We therefore find strong evidence that these survey-only adjustments are by and large confirmed by the administrative data. The adjustments for in-kind transfers and reported wage and salary hours are particularly robust, with nearly all reclassified households having incomes above $2/person/day in the administrative data. The adjustments for reported self-employment hours and substantial assets are less strongly validated by the administrative data, but it is important to remember that our administrative data do not completely cover all income sources and do not cover assets at all.

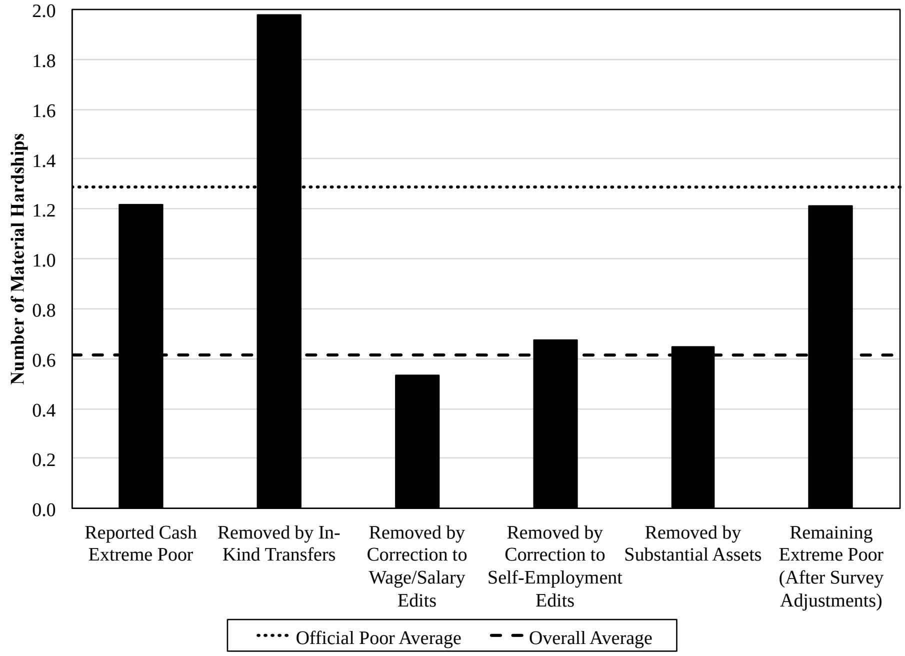
3.6.2 Survey-Reported Material Well-Being

Next, we assess the material well-being of the potentially misclassified reported extreme poor households to perform yet another test of the validity of our survey-only adjustments. Figure 3.4 displays the mean number of material hardships experienced by households among the reported extreme poor and the groups removed by each adjustment. The dotted line shows the mean for all official poor households, while the dashed line shows the mean for all households.[161] Figure 3.5

[161] Official poor households are defined as having incomes below official poverty thresholds, which vary by household size and composition.

shows the share of households with at least one hardship, while Figure 3.6 shows housing problems and Figure 3.7 appliance ownership.

Figure 3.4. Mean Number of Material Hardships for Extreme Poor Subgroups



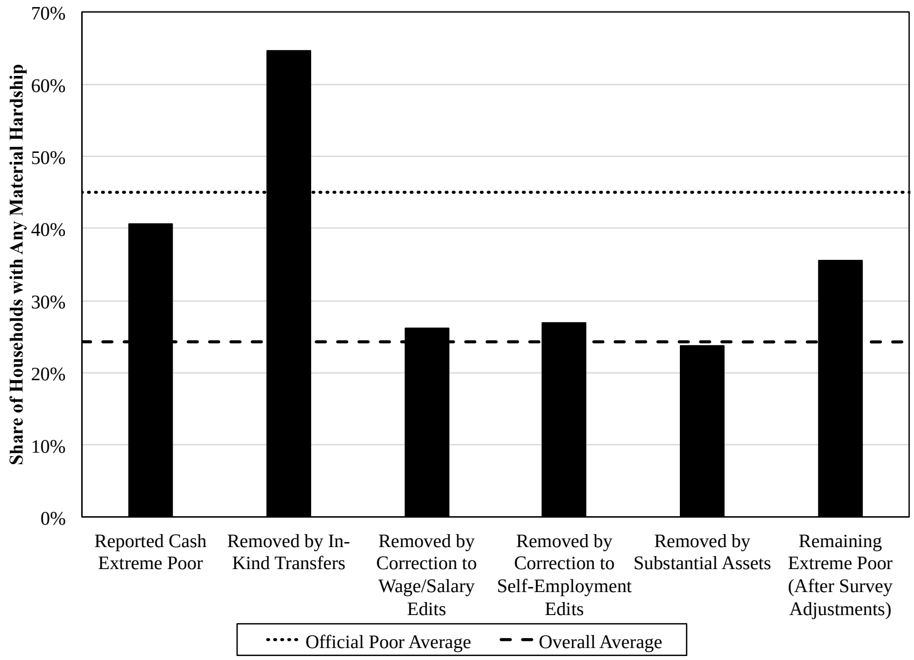
Source: Wave 9 of the public use 2008 SIPP Panel, spanning January-July 2011.

Looking first at the number of hardships, which ranges from 0 to 9, a clear pattern appears. The reported extreme poor experience 1.22 hardships on average. This count is slightly below but insignificantly different from the number of material hardships, 1.29, experienced by official poor households. Assuming that the truly extreme poor should experience more hardships than the official poor, this finding suggests that there could be substantial classification error in the reported extreme poor. Indeed, we see that sharp differences between subgroups of the reported extreme poor add up to the overall result. Most of the subgroups do not experience hardships at a level commensurate with extreme poverty, although one of the subgroups does.

Focusing first on the subgroup that appears especially disadvantaged, households that are removed from extreme poverty by in-kind transfers experience on average of 1.98 hardships, 53% more hardships than official poor households (with this gap being statistically significant). Recipients of in-kind transfers are clearly among the worst-off non-institutionalized Americans, suggesting that these transfer programs are well-targeted. On the other hand, the groups removed from extreme poverty by wage and salary hours, self-employment hours, and substantial assets experience about the same number of hardships as a typical household in the U.S. Specifically, those removed by wage and salary hours have 0.53 hardships, 13% fewer than the average of 0.61 over all households, while those removed by self-employment hours and substantial assets have 0.67 and 0.65 hardships, only 10% and 7% more, respectively, than the average household. For none of these subgroups is the difference in mean hardships from all households statistically significant. Thus, rather than being extreme poor or even poor, these households appear to be close to average or better.

The remaining extreme poor after the survey-only corrections average 1.21 hardships, which is insignificantly different from the average number of hardships experienced by the official poor. This last result suggests that some substantial errors still remain, which is not surprising since the remaining group includes those households that are not in extreme poverty when incorporating the administrative data. Similar patterns emerge when we analyze the share of households reporting least one hardship (Figure 3.5).

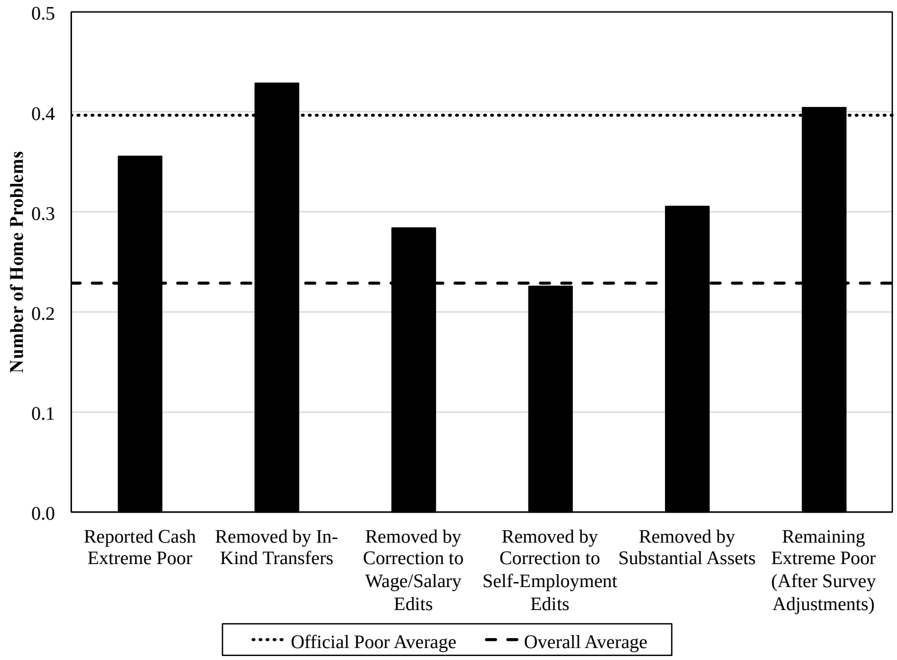
Figure 3.5. Share of Households with Any Material Hardship for Extreme Poor Subgroups



Source: Wave 9 of the public use 2008 SIPP Panel, spanning January-July 2011.

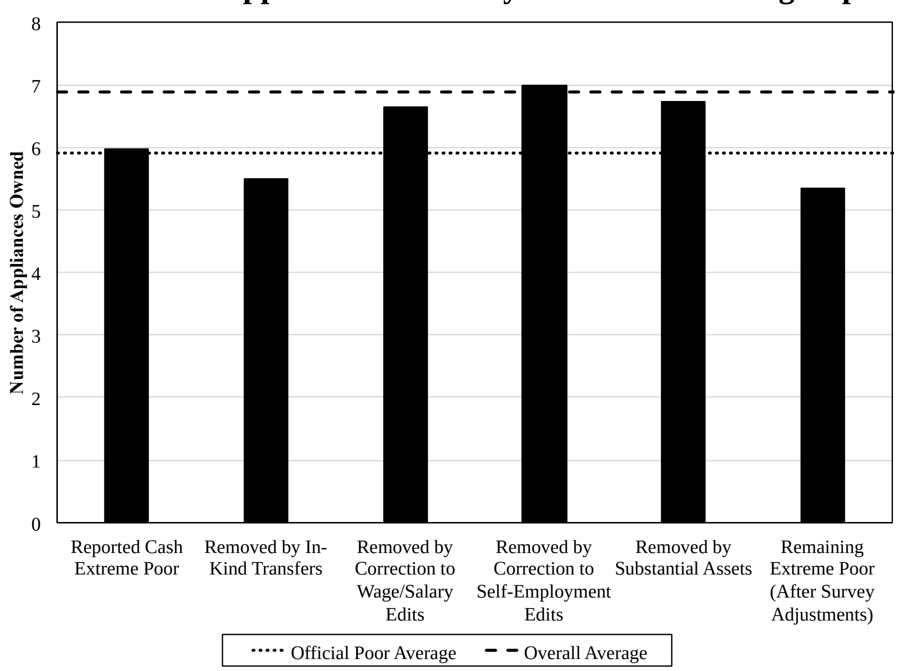
Examining housing quality issues and appliance ownership reveals patterns that are similar but often less dramatic than those for material hardships. The reported extreme poor have on average 0.36 housing quality problems and 5.98 appliances, which is insignificantly different from the 0.40 housing problems and 5.91 appliances of the official poor. These two comparisons again suggest that there are problems with the designation of extreme poverty in the raw reported data. Those households reclassified as not extreme poor due to in-kind transfers have a very high rate of housing problems, 0.43 on average (though this is insignificantly different from the official poor), and they own 5.49 appliances, significantly fewer than the official poor. Again, these gaps demonstrate the good targeting of in-kind transfers.

Figure 3.6. Mean Number of Home Problems for Extreme Poor Subgroups



Source: Wave 9 of the public use 2008 SIPP Panel, spanning January-July 2011.

Figure 3.7. Mean Number of Appliances Owned by Extreme Poor Subgroups



Source: Wave 9 of the public use 2008 SIPP Panel, spanning January-July 2011.

On the other hand, the households removed from extreme poverty due to reported hours worked or substantial assets have a mean number of housing problems closer to the overall average. Households removed by wage and salary hours worked and by substantial assets have on average 0.28 and 0.30 problems respectively, which is in between (and insignificantly different from) the average level for all households and the official poor. Those removed by self-employment hours have on average 0.23 problems, significantly lower than those of the official poor and insignificantly different from those of all households. The households removed by self-employment hours also own 6.90 appliances on average, a high level compared to the other subgroups,[162] and the households removed by wage and salary hours and substantial assets have 6.64 and 6.73 appliances, respectively, or 4% and 2% fewer than all households. For none of these subgroups is the difference from all households in mean appliances statistically significant. Finally, the remaining extreme poor again are insignificantly different from the official poor on the mean number of housing problems, but they own statistically significantly fewer appliances (almost 10% less) than the official poor.

We also test whether the differences in material well-being between the groups removed from extreme poverty by the survey-only adjustments remain after controlling for demographic covariates. To do so, we regress an indicator of well-being (mean number of hardships, appliances, or housing problems) on a dummy for whether a household is poor based on pre-tax cash income, separate dummies for whether a household is removed from extreme poverty by a given adjustment, and covariates for the age of the household head and the number of children and adults in the household. Even after the inclusion of covariates, we find that the households removed from

[162] This number is statistically significantly higher at the 1% level than that of households removed by in-kind transfers and the remaining extreme poor after survey corrections, but it is not significantly different than that of households removed by wage/salary hours or assets.

extreme poverty by in-kind transfers continue to be significantly worse off than poor households while those removed by the earnings and asset adjustments have hardships insignificantly different from the average non-poor household.

In sum, while some of the indicators of well-being may be imperfect on their own, we observe the same pattern across every measure – that the households removed from extreme poverty by in-kind transfers are materially worse off than the official poor, while the households removed from extreme poverty by the earnings and assets adjustments have a level of material well-being similar to the average over U.S. households.

3.6.3 Survey-Reported Demographics

Finally, we briefly discuss the demographic characteristics of households among the reported extreme poor. First, the patterns in educational attainment across each of the extreme poor subgroups reflect the patterns observed for material well-being. The heads of households that are considered to not be extreme poor due to in-kind transfers have the fewest years of education of any group, while the households removed by hours worked and substantial assets have education levels at least between those of the official poor and all households. Likewise, the subgroup reclassified as not extreme poor by in-kind transfers has the highest rate of reported Medicaid coverage (1.5 times the rate of official poor households), while the households removed by reported hours or substantial assets have relatively high rates of *private insurance* coverage (with near or above 50% of these households covered).[163] The households reclassified as not extreme poor by in-kind transfers also have very low asset ownership rates, while those removed by wage and salary earnings and self-employment earnings have similar asset ownership rates, respectively,

[163] Note that our time period is before the Affordable Care Act broadened Medicaid eligibility.

to the official poor and all households. Among the households considered to not be extreme poor after accounting for substantial assets, more than 63% have total net worth exceeding $100,000 and 31% have total net worth over $250,000.[164] Finally, we find that full-time students head 18.1% of households among the remaining extreme poor. Student status could proxy for access to other sources of financial support for which we do not account, such as financial aid (cash or in-kind), unreported assistance from parents, and student loans. Indeed, more than half of student-headed households among the remaining extreme poor report receiving educational assistance not included in cash income.[165]

3.7 Comparison to Current Population Survey Results

While we focus on the SIPP in this paper, we are also interested in examining whether our results generalize to the CPS ASEC (hereafter referred to as the CPS). In addition to serving as the official source of poverty and income statistics in the United States, the CPS is one of the most widely used surveys. Because the CPS collects a sparser set of information on income and well-being than the SIPP, we can only incompletely replicate our analysis.

3.7.1 Data and Methods

We use the 2012 CPS, which interviewed 74,383 households in March 2012 about their annual incomes in the previous calendar year. Thus, the reference period for the CPS includes the seven months that comprise Wave 9 of the 2008 SIPP. Our sample consists of households that

[164] There are even a handful of households in this subgroup with net worth in the millions (i.e., extremely gross errors).
[165] Our survey-reported measure of cash income includes GI bill education benefits, but none of our survey or administrative income sources includes other measures of educational assistance. See p. 3-7 of the SIPP Users' Guide: https://www2.census.gov/programs-surveys/sipp/guidance/SIPP_2008_USERS_Guide_Chapter3.pdf

have at least one member with a PIK and no members that are whole imputes, with survey weights adjusted to account for missing PIKs and the presence of whole imputes (see the appendix for more information). We employ a set of adjustments similar to those used for the SIPP but proceed in a slightly different order, allowing us to better compare the estimates of extreme poverty in the two surveys. We start with households living on less than $2/person/day over the course of 2011 according to their survey-reported cash income. We first correct for underreported earnings based on reported hours worked. We multiply a household's annual hours worked (as reported in the survey) by the federal minimum wage and remove households from extreme poverty if these lower-bound earnings are above the extreme poverty threshold. This is done separately for wage and salary hours and for total hours worked (which include self-employment hours).

Next, we incorporate in-kind transfers by reclassifying a household as not extreme poor if 1) its total cash income plus survey-reported SNAP benefits exceeds $2/person/day or 2) it receives housing subsidies.[166] We do not include WIC payments because WIC amounts are not reported in the CPS. We subsequently account for substantial assets in the CPS in a slightly different manner than in the SIPP. The CPS does not contain as detailed information as the SIPP on the specific amounts of various types of assets, but it does ask about home value and whether or not a household has a mortgage. We therefore remove a household from extreme poverty if it has *no* mortgage and a home value greater than $25,000, or if a household has a mortgage and a home value greater than $100,000. Finally, we remove a household from extreme poverty if its annual income from cash and in-kind transfers from the administrative data exceeds $2/person/day. The reference period of the CPS actually aligns better than the SIPP with the administrative tax records

[166] We assume that if a household reports receiving housing assistance, then it receives housing assistance for all 12 months of 2011. This follows the assumption that the U.S. Census Bureau makes when calculating the Supplemental Poverty Measure (Johnson et al. 2011).

because both are for the calendar year. We follow these alternative methods in this Section for both the SIPP and the CPS to allow a close comparison.

3.7.2 Results

Table 3.4 reports the extreme poverty rate after each adjustment in the CPS and compares it to the rate after the same aligned adjustment in the SIPP. The reported extreme poverty rate in the CPS of 2.08% (Column 1) is statistically significantly less than the corresponding rate of 2.97% in the SIPP (Column 2). Correcting for underreported wage and salary earnings reduces the gap in the estimates between surveys, with the extreme poverty rate declining by 0.29 percentage points to 2.68% in the SIPP and falling slightly by 0.03 percentage points to 2.05% in the CPS. Correcting for underreported self-employment earnings remarkably closes this gap, with the extreme poverty rate dropping to 2.07% in the SIPP and remaining almost unchanged at 2.03% in the CPS. We are unable to reject the null hypothesis that these rates are equal to each other (see Column 3).

Including SNAP and housing assistance cuts the extreme poverty rate by approximately a third in both surveys, yielding insignificantly different extreme poverty rates of 1.35% and 1.30% in the CPS and SIPP, respectively. Note that SNAP and housing assistance are by far the most important in-kind transfers, since including WIC (as our original SIPP adjustment for in-kind transfers does) has virtually no impact on the extreme poverty rate (Column 4). Accounting for substantial assets further reduces the extreme poverty rate to 0.96% in the SIPP and 0.80% in the CPS.[167] Bringing in administrative earnings cuts the extreme poverty rate by about half in both surveys, and bringing in the additional tax data on asset income, retirement income, and the EITC

[167] Due to the more limited asset information available in the CPS, the adjustment for substantial assets that we utilize here is narrower than what we utilize for the main SIPP results. Specifically, the extreme poverty rate in the SIPP after accounting for assets is 0.96% using this more limited adjustment, compared to 0.84% using the original adjustment (which also accounts for liquid and total assets).

Table 3.4. Comparison of CPS and SIPP Extreme Poverty Estimates for All Households

Specification	CPS	SIPP – Aligned Adjustments		SIPP – Original Adjustments	
	Rate (%)	Rate (%)	Diff. in Rates, CPS – SIPP (pp)	Rate (%)	Diff. in Rates, CPS – SIPP (pp)
	(1)	(2)	(3)	(4)	(5)
Survey-Reported Cash	2.08	2.97	-0.89***	2.97	-0.89***
	(0.08)	(0.13)		(0.13)	
Survey-Only Adjustments					
Correct Wage/Salary Earn	2.05	2.68	-0.63***	2.68	-0.63***
	(0.06)	(0.13)		(0.13)	
Correct Self-Emp. Earnings	2.03	2.07	-0.04	2.07	-0.04
	(0.06)	(0.11)		(0.11)	
Add In-Kind Transfers[1]	1.35	1.30	0.05	1.30	0.05
	(0.06)	(0.08)		(0.08)	
Account for Assets[2]	0.80	0.96	-0.16*	0.84	-0.04
	(0.05)	(0.07)		(0.07)	
Admin. Data Adjustments					
Correct Earnings	0.41	0.48	-0.07	0.42	-0.01
	(0.04)	(0.05)		(0.05)	
Correct Asset/Retire. Inc.	0.34	0.40	-0.05	0.35	-0.01
	(0.03)	(0.04)		(0.04)	
Add EITC	0.34	0.35	-0.01	0.31	0.03
	(0.03)	(0.04)		(0.04)	
Correct OASDI/SSI	0.23	0.31	-0.08*	0.27	-0.04
	(0.03)	(0.04)		(0.04)	
Correct Housing Assistance	0.21	0.31	-0.10**	0.27	-0.06
	(0.03)	(0.04)		(0.04)	
Correct SNAP	0.18	0.29	-0.11**	0.24	-0.06
	(0.03)	(0.04)		(0.04)	

*** p<0.01, ** p<0.05, * p<0.1

1. SNAP, WIC, and housing assistance in SIPP (original adjustments). SNAP and housing in CPS, SIPP (aligned adjustments).

2. For SIPP (original adjustments), owns real estate equity > $25,000, liquid assets > $5,000, or total net worth > $50,000. For CPS and SIPP (aligned adjustments), household has home value > $25,000 and has no mortgage, or has home value > $100,000 and has a mortgage.

Sources: For SIPP, we use Wave 9 of the 2008 SIPP Panel, spanning January-July 2011. For CPS, we use the 2012 CPS ASEC corresponding to reference year 2011. Administrative data sources described in text.

Note: Standard errors calculated using replicate weights in parentheses. For SIPP, sample consists of households with at least one member with a PIK and present in reference month 4. Reference month 4 survey weights are adjusted for missing PIKs. For CPS, sample consists of households with at least one member with a PIK and no members that are whole imputes. Survey weights adjusted for missing PIKs and whole imputes. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization numbers: CBDRB-FY18-324 and CBDRB-FY19-173.

decreases the extreme poverty rate to 0.35% in the SIPP and 0.34% in the CPS. After incorporating the administrative transfer data, we obtain a final extreme poverty rate among households of 0.18%

in the CPS and 0.29% in the SIPP, which are statistically significantly different at the 5% significance level. Much of the final gap between the two surveys is due to the administrative data for OASDI and SSI playing a larger role in reducing the extreme poverty rate in the CPS. However, the difference in the final extreme poverty rate between the CPS and the original adjustments in the SIPP is statistically insignificant (Column 5). We also calculate an extreme poverty rate of 0.13% for individuals in the CPS, compared to the rate of 0.11% rate we reported for the SIPP.

Consequently, the results for the two surveys are far more alike than they initially seem. The sizeable errors that we find in the left tail of the SIPP income distribution appear with almost the same frequency in the CPS. The primary difference is that households almost never report positive hours worked and extremely low earnings in the CPS, while such a pattern is relatively common among the reported cash extreme poor in the SIPP.[168] Our final estimates of the extreme poverty rate in Column 7 are also consistent with the idea that poverty over the course of an entire year (CPS) should be less frequent than poverty over the course of four months (SIPP). Finally, the larger impact of the administrative transfer data on the CPS estimate is in line with work showing greater underreporting of transfer programs in the CPS relative to the SIPP (Meyer et al. 2015).

[168] In the CPS, all households that report 0 earnings also report 0 hours worked across all members. In the SIPP, 7.94% of all households that report 0 earnings report positive average monthly hours worked for pay. Also in the SIPP, 72% of households lifted out of extreme poverty by wage and salary hours reported 0 earnings, as did 88% of households lifted out of extreme poverty by self-employment hours.

3.8 Robustness Checks and Caveats

3.8.1 Robustness Checks

We conduct a series of robustness checks to examine the sensitivity of our results to a wide set of alternative specifications. Table 3.5 presents the results of several key checks. First, we apply half the federal minimum wage (rather than the full minimum wage) to reported wage/salary and self-employment hours worked. This modification increases the extreme poverty rate by a mere 0.01 percentage points after the survey corrections for underreported earnings, and leaves the final extreme poverty rate after all adjustments unchanged (Column 1). We also calculate estimates excluding survey-reported and administrative values of housing assistance, given concerns about the lack of fungibility of housing assistance. The final extreme poverty rate after all adjustments, but prior to bringing in the administrative SNAP records, is once again 0.01 percentage points above the comparable rate accounting for housing assistance (Column 2). As expected, SNAP is responsible for most of the impact of in-kind transfers.

We also examine whether our extreme poverty results extend to a higher income cutoff – specifically \$4/person/day (see Allen 2017, Deaton 2018). Column 3 of Table 3.5 shows that they do. As expected, the extreme poverty rates are higher when measured using a higher income threshold. The reported rate of 3.68% using \$4/person/day is 24% higher than the rate using \$2/person/day, and the final rate of 0.34% after incorporating all adjustments using \$4/person/day is 42% higher than the rate using \$2/person/day (but still very low). The similar patterns using \$4/person/day are consistent with a relatively large number of household reports of zero income, many of which are likely to be gross errors. The survey-only adjustments also cut the reported extreme poverty rate by 73% when using the \$4/person/day threshold, which is comparable to the 72% cut by the survey-only adjustments when using the \$2/person/day threshold.

Table 3.5. Robustness Checks (Percentage of Households in Extreme Poverty)

Specification	**$2/Day: Half Minimum Wage (1)**	**$2/Day: No Housing Assistance (2)**	**$4/Day (3)**
Survey-Reported Cash	2.97	2.97	3.68
	(0.13)	(0.13)	(0.13)
<u>Survey-Only Adjustments</u>			
Add In-Kind Transfers	2.04	2.08	2.48
	(0.11)	(0.10)	(0.12)
Correct Wage/Salary Earn.	1.83	1.86	2.17
	(0.10)	(0.10)	(0.11)
Correct Self-Emp. Earnings	1.31	1.32	1.54
	(0.08)	(0.07)	(0.08)
Account for Assets	0.84	0.86	1.00
	(0.07)	(0.07)	(0.07)
<u>Admin. Data Adjustments</u>			
Correct Earnings	0.42	0.43	0.50
	(0.05)	(0.05)	(0.05)
Correct Asset/Retire. Income	0.35	0.36	0.42
	(0.04)	(0.04)	(0.05)
Add EITC	0.31	0.32	0.39
	(0.04)	(0.04)	(0.04)
Correct OASDI/SSI	0.27	0.28	0.35
	(0.04)	(0.04)	(0.04)
Correct Housing Assistance	0.27		0.34
	(0.04)		(0.04)
Correct SNAP	0.24		0.34
	(0.04)		(0.04)

Sources: Wave 9 of the 2008 SIPP Panel, spanning January-July 2011. Administrative data sources described in text.
Note: Standard errors calculated using replicate weights in parentheses. Sample consists of households with at least one member with a PIK and present in reference month 4. Reference month 4 survey weights are adjusted for missing PIKs. In the "Half Minimum Wage" column, half of the federal minimum wage ($3.625/hour) is used when correcting wage/salary earnings and self-employment earnings based on reported hours. In the "No Housing Assistance" column, housing assistance is not included in the adjustment for in-kind transfers. In the "$4/Day" column, households in extreme poverty are those with average incomes across the four months of the wave less than or equal to $4/person/day. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization numbers: CBDRB-FY18-324 and CBDRB-FY19-173.

Next, we examine whether our methodology misses households that should be extreme poor but have survey-reported incomes greater than \$2/person/day due to imputation or over-reporting.[169] We find that a minuscule number of households have incomes greater than \$2/person/day in the survey but fall under \$2/person/day after setting imputed earnings equal to zero and applying the administrative data – specifically, taking the maximum of survey and administrative values for earnings and housing assistance (since the administrative values for these sources are incomplete as noted earlier), replacing survey with administrative values for interest and dividends, retirement income, OASDI, SSI, and SNAP, and adding EITC amounts calculated from the administrative tax records. We also analyze how our results change after basing our survey correction for underreported earnings only on non-imputed hours worked and ignoring retirement accounts in the survey adjustment for assets. Once again, the results barely budge following each modification.

We also calculate estimates of extreme poverty at the level of the fourth reference month, which is regarded as having the most accurate survey reports (Moore 2008) and follows the reference period in Shaefer and Edin (2013). Estimates using the fourth reference month are only slightly higher than our wave-level estimates. For example, using the fourth reference month, the reported extreme poverty rate of 3.82% and the extreme poverty rate of 1.09% after accounting for substantial assets are each 25-30% higher than the comparable wave-level rates. The similarity of the estimates across the month and wave reference periods reflects the tendency of survey responses to be strongly correlated when taken within the same interview (Moore 2008).

[169] We consider an amount to be imputed if it is statistically imputed (i.e., hot or cold deck). We do not consider logically imputed amounts to be imputed, because they are based on previous wave information that is likely to be of good quality.

It is not clear what is the appropriate time interval for measuring income poverty. Most of the literature on income and well-being has argued for looking over a full year, given transitory fluctuations in income that may not be reflected in consumption or other outcomes. There is also a long literature that emphasizes the persistence of poverty or argues for looking at income over multiple years (see Duncan and Rogers 1991 or Solon 1992, for example). As well as looking at a shorter time interval than a four month wave, we considered examining SIPP estimates over a calendar year. While there are advantages to a longer period, we do not do so in the SIPP because of attrition across interview waves. Furthermore, responses would be taken across three or four interviews (rather than one), making it less comparable to surveys like the American Community Survey (ACS) and CPS ASEC that cover a reference year in a single interview. However, the results in the CPS strongly suggest that the patterns and levels of annual estimates mirror those of wave-level estimates in the SIPP.

We also examine more closely the households removed from extreme poverty due to reported hours worked. Among the top 10 occupations of workers in households removed by wage and salary and self-employment hours, most are not exempt from the federal minimum wage and, in fact, would seem to have average earnings that are generally far above minimum wage. For example, almost 9% of workers in households removed from extreme poverty by wage and salary hours are computer scientists or engineers (compared to less than 2% of all wage and salary workers). Additionally, while 1.41% of workers in this subgroup are waiters and waitresses (an occupation that could conceivably earn less than minimum wage), a higher rate (1.82%) of all wage and salary workers are waiters and waitresses. Additionally, the three most common occupations for workers in households removed from extreme poverty by self-employment hours are various kinds of managers (14% of such workers, as compared to 11.55% of all self-employed workers). To get a sense of the extent to which these subgroups are likely to have error-ridden earnings reports, we find that 72% of households removed by wage and salary hours report zero earnings and that 91% report zero, single-digit, or double-digit earnings. Of the households removed by self-employment hours, 88% report zero earnings and 94% report zero, single-digit, or double-digit earnings.

Finally, we check to see that imputed values in the survey only have a minor effect on our estimates. First, almost no households were initially classified as extreme poor because their survey incomes were imputed to be zero. Furthermore, 59% of households reclassified by survey-reported in-kind transfers have SNAP amounts imputed, but only 3% have receipt imputed. Encouragingly, 99% of the households reclassified by survey-reported in-kind transfers have incomes above $2/person/day based on the administrative data. We also find that 9% of households removed by wage and salary hours have imputed values for overall hours worked. Once again, 93% of the households removed by wage and salary hours have incomes above the extreme poverty threshold based on the administrative data.

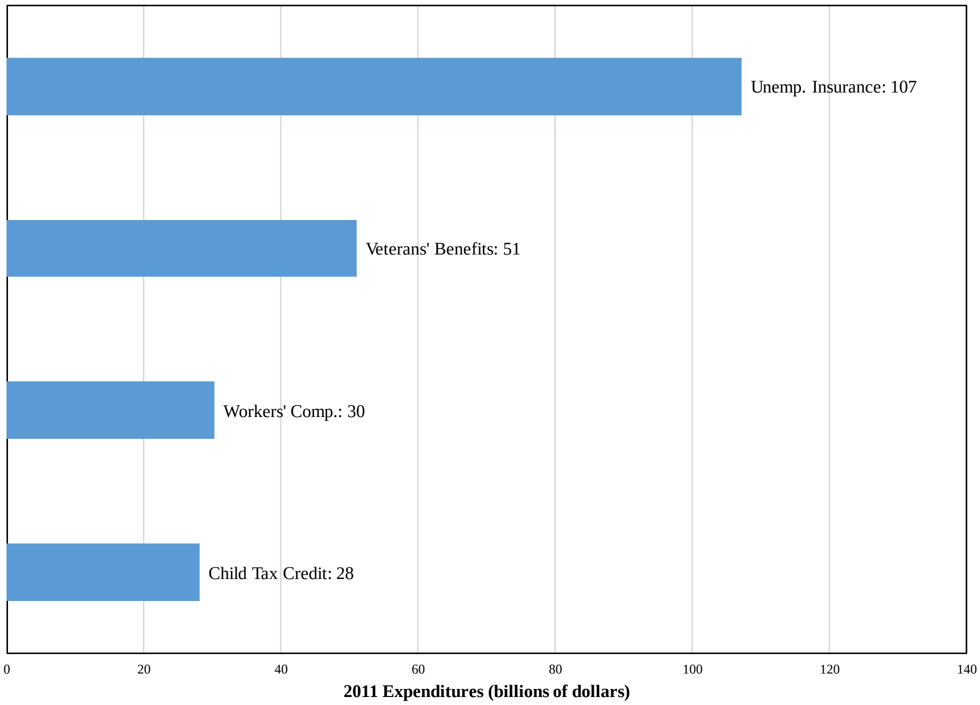
3.8.2 Caveats

In this subsection, we discuss some weaknesses in our data and their likely effect on our results. We first discuss reasons why we may have understated the extent of extreme poverty and then reasons why we may have overstated it. First, we rely on annual administrative tax data that we allocate evenly across the four months of the SIPP wave. If the months of the year with low other income are also those with low taxable income, then we will understate extreme poverty. The results from the CPS suggest that this bias is small, since the CPS yields strikingly similar results to those in the SIPP even though CPS income is annual and does not suffer from this potential misalignment. Furthermore, we may be less worried about this possibility as it could only occur when a household is only transitorily extreme poor. Given that most of those who we remove from extreme poverty via the administrative data have income over twelve months above the poverty line, to be extreme poor for four months would require these households to have income one and a half times the extreme poverty line over the remaining months of the year. It is worth emphasizing that the potential misalignment between annual and wave-level data does not apply to the administrative program data, which are all at the month level.

Second, and probably most importantly, the SIPP and CPS survey frames cover only resident households, meaning they miss homeless individuals (among other institutionalized populations). Given that there were 636,000 homeless individuals in 2011 (based on HUD estimates) and that the homeless are among the most destitute members of our communities, our final estimate of the extreme poverty rate may be an understatement for the entire population.[170] While the extreme poverty estimates in the literature discussed in Section 3.2 also rely on surveys that do not include the homeless, a broader view of extreme poverty would include them. Moreover, if homeless individuals are more likely than the non-homeless to be single and childless, then incorporating the homeless might further amplify the already large share of single childless extreme poor individuals.

[170] See https://www.hudexchange.info/resources/documents/2011AHAR_FinalReport.pdf.

Figure 3.8. Expenditures on Four Largest Transfer Programs Not in Administrative Data



Note: Does not include Medicare and Medicaid. OASDI, EITC, SSI, housing assistance, and SNAP are in the administrative data. Unemployment insurance, veterans' benefits, workers' compensation, child tax credit, Public Assistance, school food programs, WIC, and LIHEAP are *not* in the administrative data. Expenditures data from National Income and Product Account Table 3.12 and other sources.

There are also reasons why our data may overstate extreme poverty. First, we are unable to include administrative data for a number of income sources, such as Temporary Assistance for Needy Families, General Assistance, the Child Tax Credit, unemployment insurance, workers' compensation, non-DER earnings, non-HUD housing assistance, and veterans' benefits.[171] Figure 3.8 shows that we miss $216 billion from the four largest transfer programs or tax credits not in our administrative data, with $107 billion attributable to unemployment insurance alone. The total expenditures for these "excluded" programs are similar to the total expenditures for the non-

[171] While we have access to administrative TANF data and have used them in previous work (see Meyer and Wu 2018), we only have these data for 30 states and are hesitant to base our extreme poverty estimates on the seven states for which we have both administrative SNAP and TANF data.

OASDI programs for which we have administrative data.[172] We likely also miss income from sources like off-the-books employment and money from relatives (Jencks 1997).[173] Incorporating administrative sources for these other income components may lead to further reductions in the extreme poverty rate. In fact, we find that more than 20% of the remaining extreme poor households contain veterans, which may be due in part to our administrative data excluding information on veterans' benefits.

Second, we are unable to access asset information for a substantial share of our households since the asset information is from either a few months before or after our Wave 9 reference period. Because of survey attrition and slight changes in household composition across waves, not all households that appear in Wave 9 match to topical modules from the other waves (especially later waves).[174] Therefore, any households that do not link to the topical modules for Waves 7 or 10 are not removed from extreme poverty by substantial assets. In fact, one-sixth of the un-weighted households left in extreme poverty after accounting for substantial assets cannot be linked to the Wave 7 or 10 topical modules. This missing data problem leads to an understatement of the asset adjustment and therefore an overstatement of the final extreme poverty rate. We should also note that incomplete linking of individuals means we cannot bring in administrative data for all survey respondents, likely understating their income. Finally, the year we examine was near the peak of the most severe recession in 70 years, so incomes were atypically low. Overall, we expect that the incompleteness of our data – especially the linked administrative data – leads to an overstatement

[172] According to Meyer et al. (2015), the expenditures in 2011 are $48.9 billion for SSI, $72.8 billion for SNAP, and $62.9 billion for the EITC. Based on Table 8.7 in the OMB's Historical Tables, the expenditures for housing assistance in calendar year 2011 (calculated from amounts for fiscal years 2011 and 2012) are $46.8 billion.
[173] There are good reasons to believe that the non-tax earnings in the survey are themselves underreported (Hurst et al. 2014, Hokayem et al. 2015).
[174] From the publicly available SIPP data, 92.55% of the households (weighted) we use in Wave 9 link to the Wave 6 topical module. 97.88% of households we use in Wave 9 link to either Wave 10 or Wave 7: we link 90.83% to Wave 10 and 7.05% to Wave 7 (we only link households to Wave 7 if we could not link them to Wave 10).

of households in extreme poverty, but the omission of the homeless (who are not in households) is an important gap in our information.

3.9 Conclusion

Through closely examining the SIPP and augmenting the survey data with administrative tax and transfer data, we find that 92% of the households categorized as extreme poor based on survey-reported cash income are misclassified. Our methodology yields a similar finding in the CPS, where 91% of the households categorized as extreme poor based on survey-reported cash income are misclassified. Consequently, we estimate that 0.24% of households in the United States lived on $2/person/day or less over a four-month period in 2011 (SIPP) and 0.18% of households lived on $2/person/day or less over the course of the entire 2011 reference year (CPS). The corresponding share of individuals in the SIPP is 0.11%. The survey-only adjustments can explain 78% of the total decrease in extreme poverty, while the administrative data adjustments can explain 90% of the change in extreme poverty. Thus, whichever set of adjustments we incorporate first will remove the vast majority of those initially classified as extreme poor.

Many of the households included in survey-reported extreme poverty appear to be better off than the average American household based on numerous indicators of material well-being. These results indicate that survey data at the very bottom of the income distribution are especially error-ridden, and it would be a misuse of the data to take them at face value. The results also reflect the very low rates of extreme and deep consumption poverty that various studies have found. Importantly, we may yet overstate the true rate of extreme poverty, because our administrative data miss a number of important income sources for which surveys underreport or miss altogether.

This paper further demonstrates that the face of extreme poverty is quite different from what the literature has previously emphasized. Among the 285,000 households left in extreme poverty, 90% are made up of single individuals. Households with multiple childless individuals make up the other 10% of the extreme poor. Strikingly, after implementing all adjustments, no SIPP-interviewed households with children have incomes below $2/person/day. This result lies in stark contrast to the focus in academic and policy circles on the plight of extreme poor households with children. This result also indicates that extreme poverty among such households, given its low current level, cannot have risen in an economically meaningful way due to welfare reform. It is worth re-emphasizing that these dramatic results hold even in the absence of administrative data for TANF, which is targeted towards single-parent households and is heavily underreported in surveys.

Our results also indicate that means-tested transfers – especially in-kind benefits – are well-targeted to the needy, as the households reclassified as not extreme poor by in-kind transfers appear to be considerably worse off than those in official poverty. We therefore provide an explanation for the imperfect ability of the U.S. Census Bureau's Supplemental Poverty Measure to select those with low material well-being (see Meyer and Sullivan 2012a): it likely reclassifies as non-poor those receiving in-kind transfers, who are very needy, and leaves as poor those who are misclassified because of substantial assets or unreported income.

This paper leaves room for a number of extensions. First, one could examine post-tax measures of extreme poverty. While this paper does calculate the EITC from administrative tax records, it does not account for all tax credits (like the Child Tax Credit) and tax liabilities. Second, one could incorporate more complete administrative data as they become available (e.g., on veterans' benefits, unemployment insurance, and workers' compensation). Doing so would also help us understand how many of the single individuals we categorize as being in extreme poverty are misclassified because of missing administrative data. The wide variety of information in the SIPP may also allow greater understanding of the significant number of extreme poor single individuals.

More generally, we lay out a novel methodology for how income data can be better used to measure poverty. We show that our adjustments using exclusively the publicly available survey data go a long way in addressing survey errors, accounting for more than three-quarters of the change in extreme poverty due to the combination of public and administrative data adjustments. Although this paper focuses on extreme poverty, we can apply a similar methodology to address survey errors at higher income cutoffs such as the deep and official poverty thresholds. By combining the accuracy of the administrative data with the detail of the SIPP data, one may also be able to better understand the barriers to success faced by households that are truly poor.

3.A Appendix

3.A.1 Survey-Only Adjustments in the SIPP

We make use of the extensive information available in the SIPP to adjust for errors and misclassifications in reported income and capture a fuller picture of the resources respondents have access to. The methods for the four adjustments we employ are described in detail immediately below. We also make an analogous set of adjustments when analyzing extreme poverty in the CPS; the methodologies specific to the CPS are described further below.

Reported Cash Extreme Poor

From the overall sample, we exclude any households that do not have a household head or that are not present in reference month 4. We then collapse to the household-wave level and use reference month 4 weights for households and individuals. To calculate each household's mean income per-person per-day in the wave, we divide the mean total household monthly income (*thtotinc* averaged across the four months of the wave) by the mean number of days in a month and the mean number of people in the household during the wave (*ehhnumpp* averaged across the four months of the wave).

We define a household as being in extreme poverty if its mean income per-person, per-day is less than or equal to $2. We define a household as *not* being in extreme poverty if its total income is negative in *any* of the four months of the wave. Negative household income may indicate that the household suffered losses in self-employment and/or asset income. Therefore, households with negative income are not the kinds of households we would like to consider as extreme poor – or, more generally, disconnected from earnings or the safety net.

In-Kind Transfers

We add total household SNAP amounts (*thfdstp*) and WIC amounts (*t25amt*, summed across all household members) received in a given month to total household income (*thtotinc*), and then repeat the calculation described in the above subsection. From the set of reported cash extreme poor households, we remove a household from extreme poverty if 1) its income after adding WIC and SNAP is strictly greater than $2/person/day or 2) it receives any kind of housing assistance: public housing (*epubhse*) or government subsidized rent (*egvtrnt*). We are unable to add housing subsidy amounts to household income because the SIPP only reports a binary indicator indicating whether or not a household received assistance.

Hours Worked

Respondents can report hours for up to two wage and salary jobs (*ejbhrs1*, *ejbhrs2*) and two self-employment jobs (*ehrsbs1, ehrsbs2*). A respondent reporting positive hours for multiple jobs might have held the jobs at the same time or might have started one job after leaving the other.[175] The respondents report their average weekly hours worked in each job during the weeks that they worked in the wave; they are also allowed to simply report "hours vary." For each month in the wave, the respondents also report the number of weeks in the month they had a job (*rmwkwjb*) and the number of weeks they were absent without pay from their job(s) (*rmwksab*). We calculate the number of weeks the respondent actually worked in the month by subtracting *rmwksab* from *rmwkwjb*. Note that the wage and salary jobs are classified as "private for profit employee," "private not for profit employee," "local government worker," "state government

[175] For example, a respondent who works two 30-hour wage and salary jobs each week of the month (such that the respondent works 60 hours total per week) will have 30 for *ejbhrs1* and 30 for *ejbhrs2*; a respondent who works one 30-hour job for the first two weeks of the month, quits, and starts a new 30-hour job for the last two weeks of the month (such that the respondent works 30 hours total per week) will also have 30 for *ejbhrs1* and 30 for *ejbhrs2*.

worker," "federal government worker," or "family worker without pay" (see *eclwrk1*, *eclwrk2*); we ignore hours reported for jobs classified as unpaid family work.

Respondents can also report their average weekly hours worked across all jobs (*ehrsall*). The variable for overall hours worked, *ehrsall*, does not distinguish between hours worked in a wage/salary job and hours worked in a self-employment job, so we need to use the job-specific hours (*ejbhrs1*, *ejbhrs2*, *ehrsbs1*, *ehrsbs2*) to categorize the weekly hours worked as wage/salary or self-employment hours. However, because the respondent does not necessarily work all of the jobs for which they report positive hours every week of the month, we cannot simply add *ejbhrs1* and *ejbhrs2* to get weekly wage/salary hours and *ehrsbs1* and *ehrsbs2* to get weekly self-employment hours – if we do so, we risk over-counting hours. To address this, we first add *ejbhrs1* and *ejbhrs2* to get an initial sum of wage/salary hours and *ehrsbs1* and *ehrsbs2* to get an initial sum of self-employment hours. If the sum of wage/salary hours and self-employment hours is greater than *ehrsall*, we scale down each set of hours in proportion to the share of the initial sum of hours:

$$\textit{adjusted wage \& salary hours} = \textit{ehrsall} * \frac{\textit{wage \& salary hours}}{\textit{wage \& salary hours} + \textit{self-employment hours}}$$

$$\textit{adjusted self-employment hours} = \textit{ehrsall} * \frac{\textit{self employment hours}}{\textit{wage \& salary hours} + \textit{self-employment hours}}$$

As a result, the scaled-down hours add up to *ehrsall*:

$$\textit{adjusted wage \& salary hours} + \textit{adjusted self-employment hours} = \textit{ehrsall}$$

If the sum across hours reported for each individual job is less than *ehrsall*, we stick with the sum of *ejbhrs1* and *ejbhrs2* for wage/salary hours and the sum of *ehrsbs1* and *ehrsbs2* for self-employment hours (i.e., we do *not* scale up the sum of hours to match *ehrsall*). The important exception, however, is if "hours vary" was reported instead of an actual number of hours for any job. In this case, if the sum of hours across jobs for which actual hours were reported is less than *ehrsall*, we assign the difference between *ehrsall* and the sum of actual hours to the category of the job for which "hours vary" was reported (e.g. to wage/salary hours if "hours vary" was reported for *ejbhrs1* or *ejbhrs2*, to self-employment hours if "hours vary" was reported for *ehrsbs1* or *ehrsbs2*).[176] If "hours vary" was reported for multiple jobs that span both categories, we assign the full difference to self-employment hours. We do this because the assumption that self-employment jobs pay at least the federal minimum wage is more debatable, so we choose to assign any ambiguous hours to this category so as to have an upper bound on the impact of including self-employment hours. If one wanted to ignore the correction based on self-employment hours, they could look just at the wage/salary correction, which would not include any hours that could have been assigned to self-employment jobs.

In order to get the total hours worked in each type of job in a month, we multiply the hours worked in each job type by the number of weeks worked in the month (*rmwkwjb* minus *rmwksab*). We then multiply total average monthly wage/salary hours by the federal minimum wage ($7.25) to obtain our measure of minimum wage earnings. After adjusting for in-kind transfers, we remove a household from extreme poverty if its minimum wage earnings exceed $2/person/day. To be clear, we do *not* add these calculated earnings to survey reported income or transfers, but rather consider whether the calculated earnings on their own are sufficient to lift a household out of extreme poverty.[177] We next remove a household from extreme poverty if its combined average minimum wage earnings (i.e., combined average monthly self-employment hours *and* wage/salary hours multiplied by the minimum wage) exceed $2/person/day.

[176] Note that respondents can also report "hours vary" when asked about their overall hours worked across all activities. If a respondent reports "hours vary" for the number of hours worked in at least one job *and* for their overall hours worked across all activities, *ehrsall* is imputed using a hot-deck imputation process that takes into account hours worked in the prior wave, full-/part-time status in the current wave, and demographic characteristics.

[177] We do this because a rather low number of hours worked is sufficient to lift a household out of extreme poverty (e.g., 9 hours worked raises a single individual above the extreme poverty threshold).

Assets

After adjusting for in-kind transfers and under-reported earnings, we remove households from extreme poverty if they have strictly greater than $25,000 in real estate equity, $5,000 in liquid assets, or $50,000 in net worth. Real estate equity is defined as the sum of total household home equity (*thhtheq*, which includes equity in mobile homes) and total household equity in other real estate (*thhore*). Liquid assets consist of checking accounts, savings accounts, money market accounts, bonds, securities, debt or margin accounts, and certificates of deposits (i.e., total household interest earning assets held at banking institutions, *thhintbk*, total household interest earning assets at other institutions, *thhintot*, and equity in stock and mutual funds, *rhhstk*). Total household net worth (*thhtnw*) consists of real estate equity, liquid assets, retirement equity (which consists of IRA and KEOGH accounts, *thhira*, and equity in 401k and Thrift savings accounts, *thhthrif*), business equity (*thhbeq*), equity in vehicles (*thhvehcl*), and any other financial investments (*thhotast*) minus unsecured debt (*thhuscbt*).

For the most part, assets are not reported in the core file, so we pull them in from the topical modules. Assets are reported in the topical modules for Waves 4, 7, and 10; we make use of the responses pertaining to Waves 7 and 10. We first attempt to link the households in the Wave 9 core file to the Wave 10 topical module responses. Because of attrition, not all households in Wave 9 appear in Wave 10. Furthermore, a household may split, meaning the members of one household in Wave 9 may now be found among multiple households in Wave 10. Consequently, we link Wave 9 households to Wave 10 households at the person level. We match Wave 9 individuals to Wave 10 individuals, and assign the total household assets in Wave 10 to the household of the individual in Wave 9. This means that if the individuals in the Wave 9 household split into two households in Wave 10, we assign the assets from both Wave 10 households to the single Wave 9 household. In this manner, we link 90.83% of Wave 9 public use sample households to Wave 10 households.[178]

If no one from a Wave 9 household matches to any Wave 10 households (e.g., the household dropped out of the survey after Wave 9), we try to match that household to a Wave 7 household, again at the individual level.[179] We link an additional 7.05% of households to the Wave 7 topical module, such that in total we link 97.88% of all households in the public use sample to topical module assets. Note, however, that of the households remaining in extreme poverty after all survey-only adjustments, about one-sixth did not link to the Wave 7 or Wave 10 topical modules.

Variables Used in SIPP Survey-Only Adjustments

For more information on the variables used for the in-kind transfers and under-reported survey earnings corrections, see the SIPP 2008 Panel Wave 09 Core File Technical Documentation.[180] For the variables used in the assets adjustment, see the SIPP 2008 Panel Wave 10 Topical Module Technical Documentation[181] and the SIPP 2008 Panel Wave 07 Topical Module Technical Documentation.[182]

---

[178] The linkage rates in this section are calculated on the public-use SIPP data.

[179] To be clear, if *any* individual in a Wave 9 household appeared in Wave 10, we do not try to match the Wave 9 household to Wave 7, even if some Wave 9 individuals did not match to Wave 10.

[180] See https://census.gov/content/dam/Census/programs-surveys/sipp/tech-documentation/complete-documents/2008/SIPP%202008%20Panel%20Wave%2009%20-%20Core%20File.pdf.

[181] See https://www.census.gov/content/dam/Census/programs-surveys/sipp/tech-documentation/complete-documents/2008/SIPP%202008%20Panel%20Wave%2010%20-%20Topical%20Module.pdf.

[182] See https://www.census.gov/content/dam/Census/programs-surveys/sipp/tech-documentation/complete-documents/2008/SIPP%202008%20Panel%20Wave%2007%20-%20Topical%20Module.pdf.

3.A.2 Survey-Only Adjustments in the CPS

While we use similar adjustments in both our SIPP and CPS analyses, there are some differences in our CPS methods due to the different structure of the survey. Here, we describe the approaches taken in the CPS to make analogous adjustments to the ones made in the SIPP.

Reported Cash Extreme Poor

From the initial sample, we drop households in group quarters and collapse to the household-year level. We consider a household to be in extreme poverty if its total annual household income (*htotval*) divided by 365 and by the number of people in the household (*h_numper*) is less than or equal to $2. As with the SIPP, we do not consider a household to be in extreme poverty if the total annual household income is negative.

Hours Worked

In the CPS, hours worked are not reported separately for each job. For individuals working multiple jobs, we therefore cannot distinguish between hours worked in wage and salary jobs and self-employment jobs. Consequently, we assign hours based on the category of the longest held job. Respondents report their average total weekly hours worked in the weeks that they worked in the past year (*hrswk*). We obtain their total annual hours worked by multiplying *hrswk* by the number of weeks they worked in the past year (*wkswork*). We assign the total annual hours to wage/salary work if the class of the longest job (as indicated by *a_clswkr*) is "private," "federal government," "state government," or "local government." We assign the total annual hours to self-employment work if the class of the longest job is "self-employed – incorporated" or "self-employed – not incorporated." We do not use the hours reported if *a_clswkr* indicates that the job was without pay. For both wage/salary and self-employment jobs, we multiply the annual hours worked by the federal minimum wage ($7.25) to calculate minimum wage earnings. We remove a household from extreme poverty if its minimum wage earnings across all members strictly exceed $2/person/day.

In-Kind Transfers

Unlike in the SIPP, the amount of WIC received is not reported in the CPS, so in our CPS analysis we only analyze SNAP amounts received and housing assistance receipt. Of the households that remain in extreme poverty after correcting for under-reported earnings, we remove a household from extreme poverty if it has 1) strictly greater than $2/person/day after adding total household annual SNAP amounts (*hfdval*) to total household annual reported cash income or 2) reports living in a public housing project (*hpublic*) or paying lower rent because of government assistance (*hlorent*). Note that the CPS does not ask how long the household has been receiving housing assistance and asks only if the household is receiving housing assistance at the time of the interview. We assume that the household received housing assistance for the full reference year, which is the same assumption that the Census Bureau makes when calculating the Supplemental Poverty Measure.[183]

Assets

The CPS does not contain as much detail about asset ownership as the SIPP. While we cannot obtain liquid or total asset amounts in the CPS, we can attempt to identify households with real estate equity by examining property values and mortgage ownership. Accordingly, after

[183] See p. 6 of https://www.census.gov/content/dam/Census/library/working-papers/2010/demo/spm-housingassistancejuly2011.pdf.

adjusting for under-reported earnings and in-kind transfers, we remove a household from extreme poverty if (1) the household has no mortgage (according to *hpres_mort*) but has a property value (*hprop_val*) strictly greater than $25,000, *or* (2) the household has a mortgage and has a property value strictly greater than $100,000.

Variables Used in CPS Survey-Only Adjustments

For more information on the variables used in the CPS survey-only adjustments, see the CPS 2012 Annual Social and Economic (ASEC) Supplement Technical Documentation.[184]

Adjustments to SIPP Survey-Only Adjustments to Align with CPS Adjustments

In Table 3.5, we present the extreme poverty rates under the CPS adjustments, the original SIPP adjustments, and a set of SIPP adjustments that align with the CPS adjustments. For the aligned SIPP adjustments, we do not include WIC among in-kind transfers and we use the same asset thresholds as we use in the CPS (i.e. we remove a household from extreme poverty if it has no mortgage (*ehmort*) and a property value (*tpropval*) strictly greater than $25,000, *or* if it has a mortgage and a property value strictly greater than $100,000).

3.A.3 Inverse Probability Weighting Adjustment for Incomplete PIK Linkage

We link the administrative tax and transfer data to the SIPP using Protected Identification Keys (PIKs) created by the Person Identification Validation System (PVS) of the U.S. Census Bureau. For every record in the administrative and survey data, the PVS searches for a matching record by name, date of birth, sex, and address in a reference file derived from the SSA's

[184] See https://www2.census.gov/programs-surveys/cps/techdocs/cpsmar12.pdf.

Numerical Identification file (Numident). The Numident contains all transactions recorded against every Social Security Number, which is scrambled and transformed into a PIK if a matching record is found. Approximately 99% of the administrative records are associated with a PIK, and approximately 97% of households in Wave 9 of the 2008 SIPP Panel contain at least one member associated with a PIK. Because we cannot link to administrative data the survey households in which no individual has a PIK, we use inverse probability weighting (IPW) to account for this incomplete linkage. This bias should not be particularly pronounced given that only 3% of survey households do not have a PIK.

We estimate a probit model using a binary indicator for whether anyone in the survey household is PIKed as the dependent variable and a variety of observable characteristics of the household and household head as independent variables. Our model controls for household type (dummy variables for single childless adult, single parent, and multiple childless adults), the number of children in the household, the number of adults in the household, survey-reported household income (dummy variables for brackets of $20,000), survey-reported transfer receipt (dummy variables for OASDI, SSI, TANF, WIC, housing assistance, and SNAP), rural status, and presence of a household member with a disability. It also controls for the following characteristics of the household head: age (dummy variables for brackets of five years), education (dummy variables for less than high school, some high school, high school graduate, some college, and college graduate), race and ethnicity (dummy variables for Hispanic, black non-Hispanic, white non-Hispanic, and Asian non-Hispanic), citizenship status, nativity, employment status (dummy variables for unemployed and not in labor force, and works in a construction-related occupation), average weekly hours worked, health insurance receipt (dummy variables for public insurance only, private insurance only, and public and private insurance together), gender, and marital status (dummy variables for married, married and spouse absent, and never married).

This model is estimated over all SIPP households in Wave 9, weighted using their original survey weights. Using the estimated coefficients, we construct for each household the predicted probability that it contains a member associated with a PIK. We then multiply the survey household weights by the inverse of these predicted probabilities, delivering adjusted household weights correcting for the bias from incomplete linkage of SIPP records to PIKs. This adjustment yields consistent estimates so long as linkage to a PIK is uncorrelated with poverty status and incomes, conditional on observable characteristics (see Wooldridge 2007). One caveat is that an individual may have a PIK in the administrative data but not in the survey data, but this person belongs to a survey household where at least one member has a PIK. We would understate the amount of income linked from the administrative data for that household, because we cannot link information for the individual without a PIK in the survey.

For the households in the 2012 CPS ASEC, we adjust not only for missing PIKs but also for whole imputes (designated by the variable *fl665*). We carry out the adjustment process in two stages. In the first stage, we estimate a probit model using a binary indicator for whether no one in the survey household is a whole impute as the dependent variable and a variety of observable characteristics of the household and household head as independent variables. These independent variables are very similar to the set of independent variables used in the SIPP and described above. We estimate this model over all households in the CPS, weighted using their original survey weights. We then construct adjusted survey weights for the non-whole imputed households by multiplying a household's original survey weight by the inverse of the predicted probability that it contains no member that is a whole impute. In the second stage, we estimate a probit model using a binary indicator for whether anyone in the survey household is PIKed as the dependent variable and the same set of independent variables as what we use in the first stage. We estimate this model over only the households in the CPS with no whole imputes, weighted using their adjusted survey weights obtained from the first stage IPW. We then construct re-adjusted survey weights for the non-whole imputed and PIKed households by multiplying a household's adjusted survey weight after only the first stage adjustment (for whole imputes) by the inverse of the predicted probability that it contains a member associated with a PIK.

3.A.4 Household Types

When analyzing the SIPP data, we divide households into five exclusive and exhaustive types:

- Elderly households are headed by a person age 65 or older. They can be childless or have children.
- Multiple childless adult households consist of multiple people who are all age 18 or older. The head of the household is also under age 65.
- Single childless adults have only one person, and that person is age 18 or older and under age 65.
- Multiple parent households have multiple people age 18 or older, and at least one child under age 18. The head of the household is also under age 65.
- Single parent households have exactly one person age 18 or older (but under age 65), and at least one child under age 18.

We assign household types according to the composition of the household and age of household members in reference month four. In Wave 9, there is one household in which all people are under age 18; we label this household as “single parent household” because it consists of a 17-year-old mother and her children.

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association*, 105(490): 493-505.

Abraham, Katherine G., John Haltiwanger, Kristin Sandusky, and James R. Spletzer. 2013. "Exploring Differences in Employment between Household and Establishment Data." *Journal of Labor Economics*, 31(2): S129-S172.

_____. 2020. "Measuring the Gig Economy: Current Knowledge and Open Issues." In *Measuring and Accounting for Innovation in the 21st Century*, ed. C. Corrado, J. Haskel, J. Miranda, and D. Sichel. Chicago: University of Chicago Press.

Alatas, Viva, Abhijit Banerjee, Rema Hanna, Benjamin A. Olken, Ririn Purnamasari, and Matthew Wai-Poi. 2016. "Self-Targeting: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy*, 124(2): 371-427.

Allen, Robert C. 2017. "Absolute Poverty: When Necessity Displaces Desire." *American Economic Review*, 107(12): 3690-3721.

Aman, Alfred C. 2013. "Globalization and the Privatization of Welfare Administration in Indiana." *Indiana Journal of Global Legal Studies*, 20(1): 377-424.

Armour, Philip. 2018. "The Role of Information in Disability Insurance Application: An Analysis of the Social Security Statement Phase-In." *American Economic Journal: Economic Policy*, 10(3): 1-41.

Aten, Bettina H. 2005. "Report on Interarea Price Levels." BEA Working Paper No. 2005-11.

Aten, Bettina H. and Roger J. D'Souza. 2008. "Regional Price Parities: Comparing Price Level Differences across Geographic Areas." *Survey of Current Business*, 88(November): 64–74.

Baker, Michael, Janet Currie, and Hannes Schwandt. 2019. "Mortality Inequality in Canada and the United States: Divergent or Convergent Trends?" *Journal of Labor Economics*, 37(S2): S325-353.

Barr, Andrew and Sarah Turner. 2018. "A Letter and Encouragement: Does Information Increase Postsecondary Enrollment of UI Recipients?" *American Economic Journal: Economic Policy*, 10(3): 42-68.

Bee, Adam and Joshua Mitchell. 2017. "Do Older Americans Have More Income Than We Think?" SESHD Working Paper 2017-39. Washington, D.C.: U.S. Census Bureau.

Ben-Shalom, Yonatan, Robert Moffitt, and John K. Scholz. 2012. "An Assessment of the Effectiveness of Anti-Poverty Programs in the United States." In *The Oxford Handbook of the Economics of Poverty*, ed. P.N. Jefferson. Oxford: Oxford University Press.

Besley, Timothy and Stephen Coate. 1992. "Workfare versus Welfare Incentive Arguments for Work Requirements in Poverty-Alleviation Programs." *American Economic Review*, 82(1): 249-261.

Bertrand, Marianne, Sendhil Mullainathan, and Eldar Shafir. 2004. "A Behavioral-Economics View of Poverty." *American Economic Review*, 94(2): 419-423.

Bettinger, Eric P., Bridget Terry Long, Philip Oreopoulos, and Lisa Sanbonmatsu. 2012. "The Role of Application Assistance and Information in College Decisions: Results from the H&R Block FAFSA Experiment." *Quarterly Journal of Economics*, 127(3): 1205-1242.

Bhargava, Saurabh and Dayanand Manoli. 2015. "Psychological Frictions and the Incomplete Take-Up of Social Benefits: Evidence from an IRS Field Experiment." *American Economic Review*, 105(11): 3489-3529.

Black, Sandra E. 1999. "Do Better Schools Matter? Parental Valuation of Elementary Education." *Quarterly Journal of Economics*, 114(2): 577-599.

Blank, Rebecca M. 2008. "Presidential Address: How to Improve Poverty Measurement in the United States." *Journal of Policy Analysis and Management*, 27(2): 233-254.

Blank, Rebecca and Brian Kovak. 2009. "The Growing Problem of Disconnected Single Mothers." In *Making the Work-Based Safety Net Work Better*, ed. C. J. Heinrich and J. K. Scholz. New York: Russell Sage Press.

Blank, Rebecca M. and Robert F. Schoeni. 2003. "Changes in the Distribution of Children's Family Income over the 1990's." *American Economic Review*, 93(2): 304-308.

Blomquist, Glenn C., Mark C. Berger, and John P. Hoehn. 1988. "New Estimates of Quality of Life in Urban Areas." *American Economic Review*, 78(1): 89-107.

Blundell, Richard. 2014. "Income Dynamics and Life-Cycle Inequality: Mechanisms and Controversies." *The Economic Journal*, 124 (May): 289-318.

Blundell, Richard, Luigi Pistaferri, and Ian Preston. 2008. "Consumption Inequality and Partial Insurance." *American Economic Review*, 98(5): 1887-1921.

Bollinger, Christopher R., Barry T. Hirsch, Charles M. Hokayem, and James P. Ziliak. 2019. "Trouble in the Tails? What We Know about Earnings Nonresponse Thirty Years after Lillard, Smith, and Welch." *Journal of Political Economy*, 129(5): 2143-2185.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess. 2021. “Revisiting Event Study Designs: Robust and Efficient Estimation.” Working Paper.

Brady, David and Zachary Parolin. 2020. “The Levels and Trends of Deep and Extreme Poverty in the U.S., 1993-2016.” *Demography*, 57(6): 2337-2360.

Brewer, Mike, Ben Etheridge, and Cormac O’Dea. 2017. “Why are Households that Report the Lowest Incomes so Well-Off?” *The Economic Journal*, 127 (October): F24-F49.

Brzozowski, Matthew and Thomas F. Crossley. 2011. “Viewpoint: Measuring the Well-Being of the Poor with Income or Consumption: A Canadian Perspective.” *Canadian Journal of Economics*, 44(1): 88-106.

Brueckner, Jan K. 2011. *Lectures on Urban Economics.* MIT Press.

Bullard, James. 2013. “President's Message: CPI vs. PCE Inflation: Choosing a Standard Measure.” *The Regional Economist*. Federal Reserve Board of St. Louis.

Callander, Steven and B a ˚ rd Harstad. 2015. “Experimentation in Federal Systems.” *Quarterly Journal of Economics*, 130(2): 951-1002.

Callaway, Brantly and Pedro H.C. Sant'Anna. 2020. “Difference-in-Differences with Multiple Time Periods.” *Journal of Econometrics.*

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. “Bootstrap-Based Improvements for Inference with Clustered Errors.” *Review of Economics and Statistics*, 90(3): 414-427.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer. 2019. “The Effect of Minimum Wages on Low-Wage Jobs.” *Quarterly Journal of Economics*, 134(3): 1405-1454.

Chandy, Laurence and Cory Smith. 2014. “How Poor Are America’s Poorest? U.S. $2 A Day Poverty In A Global Context.” Washington, D.C.: Brookings Institute.

Chay, Kenneth Y. and Michael Greenstone. 2005. “Does Air Quality Matter? Evidence from the Housing Market.” *Journal of Political Economy*, 113(2): 376-424.

Citro, Constance F. and Robert T. Michael (eds.). 1995. *Measuring Poverty: A New Approach.* Washington, D.C.: National Academy Press.

Congressional Budget Office. 2012. “The Distribution of Household Income and Federal Taxes, 2008 and 2009.” Washington, D.C.: Congressional Budget Office.

Cunningham, Scott and Manisha Shah. 2018. "Decriminalizing Indoor Prostitution: Implications for Sexual Violence and Public Health." *Review of Economic Studies*, 85(3): 1683-1715.

Currie, Janet. 2003. "U.S. Food and Nutrition Programs." In R. Moffitt (ed.), *Means-Tested Transfer Programs in the United States*, 199-289. Chicago: University of Chicago Press.

_____. 2006. "The Take-Up of Social Benefits." In A. Auerbach, D. Card, and J. Quigley (eds.), *Public Policy and the Income Distribution*, 80-148. New York: Russell Sage Foundation.

Currie, Janet, and Jeffrey Grogger. 2001. "Explaining Recent Declines in Food Stamp Program Participation." In W. Gale and J. Rothenberg-Pack (eds.), *Brookings-Wharton Papers on Urban Affairs*. Washington, D.C.: Brookings Institution.

Daponte, Beth, Seth Sanders, and Lowell Taylor. 1999. "Why Do Low-Income Households Not Use Food Stamps? Evidence from an Experiment." *Journal of Human Resources*, 34(3): 612-628.

Davis, Lucas. 2004. "The Effect of Health Risk on Housing Values: Evidence from a Cancer Cluster." *American Economic Review*, 94(5): 1693-1704.

De Chaisemartin, Clement and Xavier D'Haultfoeuille. 2020. "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects." *American Economic Review*, 110(9): 2964-2996.

Deaton, Angus. 2018. "The U.S. Can No Longer Hide From Its Deep Poverty Problem." *New York Times*, January 24.

Deshpande, Manasi. 2016. "Does Welfare Inhibit Success? The Long-Term Effects of Removing Low-Income Youth from the Disability Rolls." *American Economic Review*, 106(11): 3300–3330.

Deshpande, Manasi and Yue Li. 2019. "Who Is Screened Out? Application Costs and the Targeting of Disability Programs." *American Economic Journal: Economic Policy*, 11(4): 213-248.

Duggan, Mark, Jonathan Gruber, and Boris Vabson. 2018. "The Consequences of Health Care Privatization: Evidence from Medicare Advantage Exits." *American Economic Journal: Economic Policy*, 10(1): 153-186.

DuMond, J. Michael, Barry T. Hirsch, and David A. Macpherson. 1999. "Wage Differentials Across Labor Markets and Workers: Does Cost of Living Matter?" *Economic Inquiry*, 37(4): 577-598.

Duncan, Greg J. and Willard Rodgers. 1991. Has Children's Poverty Become More Persistent? *American Sociological Review*, 56(4): 538-550.

Ebenstein, Avraham and Kevin Stange. 2010. "Does Inconvenience Explain Low Take-Up? Evidence from Unemployment Insurance." *Journal of Policy Analysis and Management*, 29(1): 111-136.

Edin, Kathryn J. and H. Luke Shaefer. 2015. *$2.00 A Day: Living on Almost Nothing in America.* Boston: Houghton Mifflin Harcourt.

Ellwood, David T. and Lawrence H. Summers. 1985. "Measuring Income: What Kind Should be In?" Conference on the Measurement of Noncash Benefits, U.S. Bureau of the Census.

Epple, Dennis. 2008. Tiebout Hypothesis. In S.N. Durlauf and L.E. Blume (Eds.), *The New Palgrave Dictionary of Economics*, 2^{nd} edition.

Eubanks, Virginia. 2018. "Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor." New York: St. Martin's Press.

Finkelstein, Amy and Matthew J. Notowidigdo. 2019. "Take-Up and Targeting: Experimental Evidence from SNAP." *Quarterly Journal of Economics*, 134(3): 1505-1556.

Fox, Liana. 2019. The Supplemental Poverty Measure: 2018. Current Population Report P60-268. Washington, D.C.: U.S Census Bureau.

Fox, Liana and Lewis Warren. 2018. Material Well-Being and Poverty: New Evidence Across Poverty Measures. APPAM Presentation Slides. Washington, D.C.: U.S. Census Bureau.

Gathright, Graton and Tyler A. Crabb. 2014. "Reporting of SSA Program Participation in SIPP." Working Paper. Washington, D.C.: U.S. Census Bureau.

Ganong, Peter and Jeffrey B. Liebman. 2018. "The Decline, Rebound, and Further Rise in SNAP Enrollment: Disentangling Business Cycle Fluctuations and Policy Changes." *American Economic Journal: Economic Policy*, 10(4): 153-176.

General Accounting Office. 1995. Poverty Measurement: Adjusting for Geographic Cost-of-Living Difference. Washington, D.C.: U.S. General Accounting Office, GAO/GGD-95-64.

Glaeser, Edward L. 1998. "Should Transfer Payments Be Indexed to Local Price Levels?" *Regional Science and Urban Economics*, 28(1): 1-20.

_____. 2011. "Measuring Local Poverty Rates." *Working Paper*.

Glassman, Brian. 2019. "Multidimensional Deprivation in the United States: 2017," Report No. ACS-40. Washington, D.C.: U.S. Census Bureau.

Golden, Olivia and Amelia Hawkins. 2011. "TANF Child-Only Cases." Washington, D.C.: Urban Institute.

Goodman-Bacon, Andrew. 2021. “Difference-in-Differences with Variation in Treatment Timing.” *Journal of Econometrics*.

Gordon, Tracy, Richard Auxier, and John Iselin. 2016. “Assessing Fiscal Capacities of States: A Representative Revenue System-Representative Expenditure System Approach, Fiscal Year 2012.” Research Report. Washington, D.C.: Urban Institute.

Gray, Colin. 2019. “Leaving Benefits on the Table: Evidence from SNAP.” *Journal of Public Economics*, 179(C): 1-15.

Greenstone, Michael. 2017. “The Continuing Impact of Sherwin Rosen’s ‘Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition.’” *Journal of Political Economy*, 125(6): 1891-1902.

Gyourko, Joseph and Joseph Tracy. 1991. “The Structure of Local Public Finance and the Quality of Life.” Journal of Political Economy, 99(4): 774-806.

Hall, Jamie B. and Robert Rector. 2018. “Examining Extreme and Deep Poverty in the United States.” Backgrounder No. 3285. Washington, D.C.: The Heritage Foundation.

Haurin, Donald R. 1980. “The Regional Distribution of Population, Migration, and Climate.” *Quarterly Journal of Economics*, 95(2): 293-308.

Herd, Pamela and Donald Moynihan. 2018. *Administrative Burden: Policymaking by Other Means*. New York: Russell Sage Foundation.

Hirsch, Barry. 2011. “Adjusting Poverty Thresholds When Area Prices Differ: Labor Market Evidence.” Working Paper.

Hokayem, Charles, Christopher Bollinger, and James P. Ziliak. 2015. “The Role of CPS Nonresponse in the Measurement of Poverty.” *Journal of the American Statistical Association*, 110(511): 935-945.

Homonoff, Tatiana and Jason Somerville. 2019. “Program Recertification Costs: Evidence from SNAP.” Working Paper.

Hoynes, Hilary W. and Diane Whitmore Schanzenbach. 2009. “Consumption Responses to In-Kind Transfers: Evidence from the Introduction of the Food Stamp Program.” *American Economic Journal: Applied Economics*, 1(4): 109-39.

Hurst, Erik, Geng Li, and Benjamin Pugsley. 2014. “Are Household Surveys Like Tax Forms? Evidence from Income Underreporting of the Self-Employed.” *Review of Economics and Statistics*, 96(1): 19-33.

Iceland, John, Claire Kovach, and John Creamer. 2021. "Poverty and the Incidence of Material Hardship, Revisited." *Social Science Quarterly*, 102(1): 585-617.

Internal Revenue Service. 2016. "Tax Gap Estimates for Tax Years 2008-2010." Washington, D.C.: Internal Revenue Service.

Jencks, Christopher. 1997. "Introduction to *Making Ends Meet*, by K. Edin and L. Lein." New York: Russell Sage Foundation.

Johnson, Paul D., Trudi J. Renwick, and Kathleen Short. 2011. "Estimating the Value of Federal Housing Assistance for the Supplemental Poverty Measure." SEHSD Working Paper 2010-13. Washington, D.C.: U.S. Census Bureau.

Jones, Damon and Ioana Marinescu. 2018. "The Labor Market Impacts of Universal and Permanent Cash Transfers: Evidence from the Alaska Permanent Fund." NBER Working Paper No. 24312.

Kahn-Lang, Ariella and Kevin Lang. 2019. "The Promise and Pitfalls of Differences-in-Differences: Reflections in *16 and Pregnant* and Other Applications." *Journal of Business and Economic Statistics*, 38(3): 613-620.

Kaplow, Louis. 1995. "Regional Cost-of-Living Adjustments in Tax-Transfer Schemes." *Tax Law Review*, 51(2): 175-198.

Kleven, Henrik Jacobsen and Wojciech Kopczuk. 2011. "Transfer Program Complexity and the Take-Up of Social Benefits." *American Economic Journal: Economic Policy*, 3(1): 54-90.

Kopczuk, Wojciech and Cristian Pop-Eleches. 2007. "Electronic Filing, Tax Preparers, and Participation in the Earned Income Tax Credit." *Journal of Public Economics*, 91(7-8): 1351-1367.

Lillard, Lee, James P. Smith, and Finis Welch. 1986. "What Do We Really Know about Wages? The Importance of Nonreporting and Census Imputation." *Journal of Political Economy*, 94(3): 489-506.

Loprest, Pamela J. 2011. "Disconnected Families and TANF." *Temporary Assistance for Needy Families Program—Research Synthesis Brief Series*. Washington, D.C.: The Urban Institute.

Loprest, Pamela and Austin Nichols. 2011. "Dynamics of Being Disconnected from Work and TANF." Washington, D.C.: The Urban Institute.

Madrian, Brigitte C. and Dennis F. Shea. 2001. "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior." *Quarterly Journal of Economics*, 116(4): 1149-1187.

Mani, Anandi, Sendhil Mullainathan, Eldar Shafir, and Jiaying Zhao. 2013. "Poverty Impedes Cognitive Function." *Science*, 341(6149): 976-980.

Manoli, Dayanand and Nicholas Turner. 2014. "Nudges and Learnings: Evidence from Informational Interventions for Low-Income Taxpayers." NBER Working Paper No. 20718.

Mayer, Susan E. and Christopher Jencks. 1989. "Poverty and the Distribution of Material Hardship." *Journal of Human Resources*, 24(1): 88-114.

Medalia, Carla, Bruce D. Meyer, Amy O'Hara, and Derek Wu. 2019. "Linking Survey and Administrative Data to Measure Income, Inequality, and Mobility." *International Journal of Population Data Science*, 4(1): 1-8.

Meyer, Bruce D. 2017. "The Earned Income Tax Credit." In *A Safety Net That Works: Improving Federal Programs for Low-Income Americans*, ed. R. Doar. Washington, D.C.: American Enterprise Institute.

Meyer, Bruce D. and Nikolas Mittag. 2019. "Using Linked Survey and Administrative Data to Better Measure Income: Implications for Poverty, Program Effectiveness, and Holes in the Safety Net." *American Economic Journal: Applied Economics*, 11(2): 176-204.

Meyer, Bruce D., Wallace K.C. Mok, and James X. Sullivan. 2015. "Household Surveys in Crisis." *Journal of Economic Perspectives*, 29(4):199–226.

Meyer, Bruce D. and James X. Sullivan. 2003. "Measuring the Well-Being of the Poor Using Income and Consumption." *Journal of Human Resources*, 38(S):1180-1220.

_____. 2004. The Effects of Welfare and Tax Reform: The Material Well-Being of Single Mothers in the 1980s and 1990s. *Journal of Public Economics*, 88:1387-1420.

_____. 2008. Changes in the Consumption, Income, and Well-Being of Single Mother Headed Families. *American Economic Review*, 98(5): 2221-2241.

_____. 2011. "Viewpoint: Further Results on Measuring the Well-Being of the Poor using Income and Consumption." *Canadian Journal of Economics*, 44(1): 52-87.

_____. 2012a. "Identifying the Disadvantaged: Official Poverty, Consumption Poverty, and the New Supplemental Poverty Measure." *Journal of Economic Perspectives*, 26(3): 111-136.

_____. 2012b. "Winning the War: Poverty from the Great Society to the Great Recession." *Brookings Papers on Economic Activity*, Fall 2012: 133-200.

Meyer, Bruce D. and Derek Wu. 2018. "The Poverty Reduction of Social Security and Means-Tested Transfers." *Industrial and Labor Relations Review*, 71(5): 1106-1153.

Meyer, Bruce D., Derek Wu, and Brian Curran. 2021. "Does Geographically Adjusting Poverty Thresholds Improve Poverty Measurement and Program Targeting?" Working Paper.

Meyer, Bruce, Derek Wu, Grace Finley, Patrick Langetieg, Carla Medalia, Mark Payne, and Alan Plumley. 2020. "Estimating Tax Liabilities and Credits Using Survey and Administrative Data." NBER Working Paper No. 28229.

Meyer, Bruce D., Derek Wu, Victoria Mooers, and Carla Medalia. 2021. "The Use and Misuse of Income Data and Extreme Poverty in the United States." *Journal of Labor Economics*, 39(S1): S5-S58.

Mills, Gregory, Tracy Vericker, Heather Koball, Kye Lippold, Laura Wheaton, and Sam Elkin. 2014. "Understanding the Rates, Causes, and Costs of Churning in the Supplemental Nutrition Assistance Program (SNAP)." Washington, D.C.: Urban Institute.

Mittag, Nikolas. 2019. "Correcting for Misreporting of Government Benefits." *American Economic Journal: Economic Policy*, 11(2): 142-164.

Moffitt, Robert. 1983. "An Economic Model of Welfare Stigma." *American Economic Review*, 73(5): 1023-1035.

Moore, Jeffrey C. 2008. "Seam Bias in the 2004 SIPP Panel: Much Improved, but Much Bias Still Remains." Survey Methodology 2008-3 Report. Washington, D.C.: U.S. Census Bureau.

Moretti, Enrico. 2011. "Local Labor Markets." In. O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics, Volume 4B,* 1237-1313. Amsterdam: Elsevier.

Mukherjee, Anita. 2021. "Impacts of Private Prison Contracting on Inmate Time Served and Recidivism." *American Economic Journal: Economic Policy*, 13(2): 408-438.

Mullainathan, Sendhil and Eldar Shafir. 2013. *Scarcity: Why Having Too Little Means So Much*. New York: Times Books, Henry Holt and Company.

National Academies of Sciences, Engineering, and Medicine. 2018. "The 2014 Redesign of the Survey of Income and Program Participation: An Assessment." Washington, D.C.: The National Academies Press.

Needels, Karen, Walter Corson, Tim Meier, Ira Harley, and Karen Blass. 2000. "Evaluation of the Impact of Telephon Initial Claims Filing." Princeton, NJ: Mathematica Policy Research, Inc.

Nichols, Albert L. and Richard J. Zeckhauser. 1982. "Targeting Transfers through Restrictions on Recipients." *American Economic Review,* 72(2): 372-377.

Nichols, Donald, Eugene Smolensky, and Nicolaus Tideman. 1971. "Discrimination in Waiting Time in Merit Goods." *American Economic Review*, 61(3): 312-323.

Oates, Wallace E. 1969. "The Effects of Property Taxes and Local Public Spending on Property Values: An Empirical Study of Tax Capitalization and the Tiebout Hypothesis." *Journal of Political Economy*, 77(6): 957-971.

Olsen, Edgar O. 2003. "Housing Programs for Low-Income Households." In *Means-Tested Transfer Programs in the United* States, ed. R. Moffitt. Chicago: University of Chicago Press.

_____. 2019. "Does HUD Overpay for Voucher Units, and Will SAFMRs Reduce the Overpayment?" Working Paper.

Parolin, Zachary and David Brady. 2019. "Extreme Child Poverty and the Role of Social Policy in the United States." *Journal of Poverty & Social Justice*, 27(1): 3-22.

Renwick, Trudi, Bettina Aten, Eric Figueroa, and Troy Martin. 2014. "Supplemental Poverty Measure: A Comparison of Geographic Adjustments with Regional Price Parities vs. Median Rents from the American Community Survey." SEHSD Working Paper No. 2014-22. Washington, D.C.: U.S. Census Bureau.

Renwick, Trudi, Eric Figueroa, and Bettina Aten. 2017. "Supplemental Poverty Measure: A Comparison of Geographic Adjustments with Regional Price Parities vs. Median Rents from the American Community Survey: An Update." SEHSD Working Paper No. 2017-36. Washington, D.C.: U.S. Census Bureau.

Renwick, Trudi. 2018. "Incorporating Amenities into Geographic Adjustments of the Supplemental Poverty Measure Thresholds." SEHSD Working Paper No. 2018-32. Washington, D.C.: U.S. Census Bureau.

_____. 2019. "Supplemental Poverty Measure: Alternative Geographic Adjustments." Presentation Slides.

Roback, Jennifer. 1982. "Wages, Rents, and the Quality of Life." *Journal of Political Economy*, 90(6): 1257-1278.

_____. 1988. "Wages, Rents, and Amenities: Differences Among Workers and Regions." *Economic Inquiry*, 26(1): 23-41.

Rosen, Sherwin. 1974. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy*, 82(1): 34-55.

Rossin-Slater, Maya. 2013. "WIC in Your Neighborhood: New Evidence on the Impacts of Geographic Access to Clinics." *Journal of Public Economics*, 102(C): 51-69.

Rowe, Gretchen, Carolyn O'Brien, Sam Hall, Nancy Pindus, Lauren Eyster, Robin Koralek, and Alexandra Stancyzk. 2010. "Enhancing Supplemental Nutrition Assistance Program (SNAP) Certification: SNAP Modernization Efforts." Washington, D.C.: Urban Institute.

Ruggles, Patricia. 1990. *Drawing the Line: Alternative Poverty Measures and their Implications for Public Policy*. Washington D.C.: The Urban Institute Press.

Scally, Corianne P., Amanda Gold, Carl Hedman, Matt Gerken, and Nicole Dubois. 2018. "The Low-Income Housing Tax Credit." Washington, D.C.: Urban Institute.

Scally, Corianne P. and David Lipsetz. 2017. "New Public Data Available on USDA Rural Housing Service's Single-Family and Multifamily Programs." *Cityscape*, 19(1): 295-304.

Shaefer, H. Luke and Kathryn J. Edin. 2013. "Rising Extreme Poverty in the United States and the Response of Federal Means-Tested Transfer Programs." *Social Service Review*, 87(2): 250-268.

_____. 2017. "Welfare Reform and the Families it Left Behind." *Pathways*, Fall 2017: 22-27. Stanford Center on Poverty and Inequality.

Schwabish, Jonathan A. 2012. "Downloading Benefits: The Impact of Online Food Stamp Applications on Participation." Working Paper.

Shantz, Kathryn and Liana E. Fox. 2018. "Precision in Measurement: Using State-Level Supplemental Nutrition Assistance Program and Temporary Assistance for Needy Families Administrative Records and the Transfer Income Model (TRIM3) to Evaluate Poverty Measurement." Washington, D.C.: U.S. Census Bureau.

Shepard, Mark and Myles Wagner. 2021. "Reducing Ordeals through Automatic Enrollment: Evidence from a Subsidized Health Insurance Exchange." Working Paper.

Short, Kathleen S. 2005. "Material and Financial Hardship and Income-Based Poverty Measures in the USA." *Journal of Social Policy*, 34(1): 21-38.

_____. 2014. "The Supplemental Poverty Measure in the Survey of Income and Program Participation: 2009." Report Number P60-254. Washington, D.C.: U.S. Census Bureau.

Solon, Gary. 1992. "Intergenerational Income Mobility in the United States." *American Economic Review*, 82(3): 393–408.

Sun, Liyang and Sarah Abraham. 2020. "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects." *Journal of Econometrics*.

Tiebout, Charles M. 1956. "A Pure Theory of Local Expenditures." *Journal of Political Economy*, 64(5): 416-424.

Turner, Lesley J., Sheldon Danziger, and Kristin S. Seefeldt. 2006.“Failing the Transition from Welfare to Work: Women Chronically Disconnected from Employment and Cash Welfare.” *Social Science Quarterly*, 87(2): 227–249.

United Nations. 2018. “Report of the Special Rapporteur on Extreme Poverty and Human Rights on His Mission to the United States of America.”

United Nations Economic Commission for Europe. 2017. “Guide on Poverty Measurement.”

Unrath, Matthew. 2021. “Targeting, Screening, and Retention: Evidence from California's Food Stamps Program.” Working Paper.

U.S. Department of Agriculture (USDA). 2019. “SNAP Policy Database, SNAP Policy Data Sets.” Economic Research Service. https://www.ers.usda.gov/data-products/snap-policy-data-sets/

U.S. Department of Health, Education, and Welfare. 1976. “The Measure of Poverty: A Report to Congress as Mandated by the Education Amendments of 1974.” Washington, D.C.: U.S. Government Printing Office.

Wagner, Deborah and Mary Layne. 2014. “The Person Identification Validation System (PVS): Applying the Center for Administrative Records Research & Applications' (CARRA) Record Linkage Software.” CARRA Working Paper 2014-01. Washington, D.C.: U.S. Census Bureau.

Winship, Scott. 2016. “Poverty After Welfare Reform.” Manhattan Institute.

Wooldridge, Jeffrey M. 2007. “Inverse Probability Weighted Estimation for General Missing Data Problems.” *Journal of Econometrics*, 141: 1281–1301.

Yagan, Danny. 2019. “Employment Hysteresis from the Great Recession.” *Journal of Political Economy*, 127(5): 2505-2558.

Ziliak, James P. 2016. “Temporary Assistance for Needy Families.” In R. Moffitt (ed.), *Economics of Means-Tested Transfer Programs in the United States (Volume 1)*, 303-393. Chicago: University of Chicago Press.