Innovative Methods for Using Census Data to Study Poverty, Labor Markets, and Policy

A THESIS
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

José Daniel Pacas Viscarra

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

JOSEPH RITTER

September, 2017

Acknowledgements

I have spent nearly my entire life nestled comfortably in institutes of education. And now that this phase of my life is over, I have been reflecting on the amount of privilege that I have enjoyed throughout this journey. As President Obama said: "If you were successful, somebody along the line gave you some help." A lot, and I mean a lot, of people helped me along the way.

My graduate studies started at the Humphrey School of Public Affairs. Not only did the professors deeply impact my thinking but also paved the way to my starting the Ph.D. In particular, I thank the professors who took a chance on me: Katherine Fennelly, Deborah Levison, and Samuel Myers. I did not think a Ph.D. was in the cards for me but they convinced me otherwise.

Of course, getting through the Ph.D. would not be possible without the support of my advisor, Joe Ritter. I thank Joe for the countless hours and ideas he provided me throughout the years. I especially thank him for his patience during my time in D.C. I also thank my committee members for their support. I'm forever indebted to Aaron Sojourner for helping me through an entire chapter of my dissertation. Elizabeth Davis was instrumental in pushing me towards the Ph.D. and was always a source of good advice. Lastly, I thank Steve Ruggles for not just serving on my committee but also for creating the Minnesota Population Center. Without it, I'd be unemployed.

Graduate school is expensive and I was fortunate enough to have been funded by the Interdisciplinary Center for the Study of Global Change. I thank, the Andrew W. Mellon Foundation and ICGC for providing the financial support necessary to enroll and succeed in this endeavor. I also thank the Applied Economics department for providing financial support throughout the way. In particular, I thank the Center for International Food and

Agriculture Policy for funding a summer of research.

My academic life took a turn for the better when I started working as a graduate research assistant at the MPC. Not only did the MPC provide funding for my studies, it also developed the skills and interests that have defined my academic and professional career. Above all, I thank my boss and mentor, Sarah Flood, for giving me the freedom to explore my interests and truly believing in my abilities. Her support has been influential in my career. I also thank Katie Genadek for always being a willing listener and providing guidance in both my academic and professional career.

Along the way, I received support from a lot of different people and organizations. I thank Bob Warren for taking me under his wing during my time with the Center for Migration Studies of New York. I thank Carolyn Liebler for always keeping me in mind for collaborations. I thank the Center on Labor, Human Services, and Population at the Urban Institute, particularly Margaret Simms, Liz Peters, and Breno Braga. I am most indebted to the Social, Economic, and Housing Statistics Division of the Census Bureau. They provided a supportive environment for my research while also allowing me to start my professional career. I thank my colleagues in the Poverty Statistics Branch and the Longitudinal Research, Evaluation, and Outreach Branch. There were many a long metro ride that Michael Gideon and Marta Murray-Close listened to my dissertation rants. I thank Misty Heggeness for providing great mentorship and support. The final push on my dissertation would not have been possible without her. I thank Liana Fox for always checking in on me, reading my work, and offering words of encouragement on an almost daily basis. Finally, I thank Trudi Renwick for seeing potential in me and gambling that I could make positive contributions at the Census Bureau. Her belief in me enabled this dissertation and, if I ever amount to anything professionally, I'll have her to thank for giving me my professional start.

Many friends helped along the way by keeping me sane and providing perspective on the whole thing. I thank all my Williams friends, the WRFC, and Bruce Stephenson. To name a few, I thank Michael Hagerty, Pierre Bordeaux, James DiCosmo, Nick Greer and Tom Derbish. I also thank all my friends from the Humphrey School and the Applied Economics department. In particular, I thank Kari and Max Heerman for all the warm meals and welcome distractions, Marc Dettmann for all the jokes, and Greg Beaver for sharing the pain of the process.

Throughout my journey, my family has grown considerably. It seems fitting to quote Hubert Humphrey who once said that "behind every successful man is a proud wife and a surprised mother-in-law." When Martha and I got married, I promised to take care of her (though the truth is she mostly takes care of me) and then I promptly went back to school for six years. I thank my in-laws, Jenny, Fred, Matt and Maggie, for always offering words of support and not losing patience with me. Uncle Glenn and Dominique were also encouraging supporters of my finishing this dissertation.

I am most proud to be able to thank the Pacas family. I thank Andrea and Laura for always encouraging me to follow my heart. Since we were very young, I have always felt their belief in me. They made me feel like I can do anything even though I only do a few things well. I also thank them for marrying great people. I thank Mike and Ardit for being nice to my sisters. It is a lot easier to live away from my sisters knowing that they are loved.

To my mother and my father, I give thanks beyond words. Everything I have been able to do in my life has been a result of their sacrifice, their hard work, and their never-ending love and commitment to their children. I remember trying to figure out what college to go to my junior year of high school. What did we know about this whole process? But they pushed me to reach as high as I could and gave me as much support, emotional and financial, as was possible. I realize now the immense effort it took to get us there. As a team, they made all of this possible. Dad, you pushed all of us to do our best in school, at work, in all aspects of our life. You taught us to never to be complacent. You did your best and I want you to know that your best was certainly good enough. Mom, you taught us how to be compassionate in our daily lives and taught us modesty and humility. But you also taught us to be resolute in our passions and to always listen to our hearts. Your voices live in me forever.

And, of course, I thank Martha.

Dedication

To my patient, supportive, and loving partner in life, Martha. From the day I met you, I've strived to be a better version of myself. To that end, I wrote and rewrote this dissertation. Here we are, eight years later, a little older, slightly more tired, but stronger than ever. Let's pick something shorter next time.

Abstract

This dissertation studies poverty, labor markets, and policy. Integral to this effort are innovative methods for using Census data to study these topics. The dissertation consists of three chapters. The first chapter studies year-to-year poverty transitions in the United States. The second chapter measures the extent to which individuals' union membership status affects the levels of taxes they pay and the cost of public benefits they receive. The third chapter analyzes how the electronic employment verification system, known as E-Verify, affects the labor market outcomes of unauthorized immigrants.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

This dissertation is focused on developing innovative methods for using commonly used data sources to analyze policy regarding labor markets. In particular, this dissertation substantively studies the factors associated with poverty transitions, the role of labor unions on taxes paid and benefits received, and the effect of immigration policy on labor market outcomes. To do so, this dissertation develops new methods for using Census Bureau data, namely the Current Population Survey (CPS) and the American Community Survey (ACS). This introduction gives an overview of the three chapters that constitute this dissertation.

The first chapter studies year-to-year poverty transitions in the United States. This chapter provides the foundation for studying poverty transitions using linked Annual Social and Economic Supplement of the Current Population Survey (CPS-ASEC) data. By doing so, the chapter shows that poverty transitions are sensitive to the definition of poverty. For example, the Supplemental Poverty Measure (SPM) demonstrates a higher level of people entering and exiting poverty than the Official Poverty Measure (OPM). Focusing on the SPM, in order to measure the relative importance of particular household resource components, a simple framework is developed to test the degree to which poverty transitions can be explained by one resource. Income from wages/salaries are the most influential resource component for poverty transitions but it is shown that other resources are important in explaining poverty transitions. Finally, this framework is used in four real-life contexts with the following main results. First, consider families who have a householder aged 69 or older and who exit poverty. More than 50 percent of these families exit poverty because of an increase in the Social Security benefits they receive. Second, large increases in medical expenditures ($20,000 or more) are not a common occurrence in the data but are important predictors of families entering poverty. Third, consider families that had a householder working full-time year-round in one year but became unemployed in the next year. One-third of these families will enter poverty and the majority of these families fall into poverty solely because of the loss in wages/salaries. Fourth, when a male adult leaves the household, about 20 percent of those families will enter poverty while only about 10 percent of families enter poverty when a female leaves the family.

The second chapter, co-authored with Aaron Sojourner (Associate Professor, Department of Work and Organizations), measures the extent to which individuals' union membership status affects the levels of taxes they pay and the cost of public benefits they receive. A positive effect of unions on individual wages and employer-provided fringe benefit levels has been well-established, especially at the low-end of the wage distribution. If hours do not fall much, this should raise labor income. This positive effect might have a positive impact on individual net fiscal impact (NFI), i.e. taxes paid less the cost of public benefits received. On the other hand, union membership may reduce net fiscal impact by raising receipt of earned income tax credits, unemployment insurance, and workers compensation. This chapter uses CPS data between 1994 and 2015 to study the effect of union membership on net fiscal impact overall and give evidence on the importance of various channels. Using both pooled cross-sections and individual first-differences, the chapter documents that union members pay more in taxes and receive less in public benefits, implying a a positive net fiscal impact through the worker-level channels studied here. Other channels by which union membership affects NFI, such as by decreasing firm profits or affecting policy, are described in the context of available, relevant evidence.

The third chapter analyzes how the electronic employment verification system, known as E-Verify, affects the labor market outcomes of unauthorized immigrants. E-Verify, is widely considered an important component of immigration reform. Lacking federal action, various states have passed laws requiring the use of E-Verify to certain employers. As the measurement of unauthorized immigrants is a critical component of assessing these effects, this chapter uses "logical edits" micro-data which is widely considered the most reliable data on unauthorized immigrants. Exploiting the variation in policy implementation across states, this paper uses difference-in-difference estimation with CPS-ASEC data to estimate the effect of universal E-Verify on the employment levels of unauthorized immigrants, naturalized Hispanics, and US-born non-Hispanics. Results suggest that universal E-Verify lowers the likelihood of employment for all groups though the effect is largest for unauthorized immigrants and naturalized Hispanics. Furthermore, using the ACS and a similar difference-in-difference estimation, it is shown that universal E-Verify also reduces the population size of unauthorized immigrants in E-Verify states.

Chapter 2

More than just a job?: Characterizing poverty transitions in the U.S.

2.1 Introduction

Poverty is a human condition that, when boiled down to a single number, we fail to describe in human terms. For example, a common description of poverty in policy circles may be: "In 2015, the official poverty rate had one of its highest one-year drops in decades; the 13.5 percent poverty rate represented a drop of nearly 3.5 million people from the 14.8 percent of 2014" (Proctor, Semega and Kollar, 2016). This reductiveness is inevitable in statistics but need not be in the construction of the statistic itself. The official poverty measure (OPM) highlights this point. In the simplest terms, the measure adds the total cash a family brings in throughout one year and compares that to a poverty line. But how does the family who suffers a medical emergency and must pay thousands of dollars in medical bills fit into this equation? How do we account for the single mother who loses her job and relies on food stamps for her family's livelihood? There have been extensive efforts to expand the definition of poverty for decades, most notably the National Academy of Sciences efforts in the 1990s that culminated in what is known today as the Supplemental Poverty Measure (SPM). This measure provides a solid framework for analyzing poverty but, to date, it has mainly been studied in a cross-sectional framework. Events such as death, divorce, or even large increases in medical expenditures, are identifiable in a longitudinal setting that are not in a cross-sectional setting. This paper's aim is to leverage the data available to describe poverty in relation to life events and to move poverty statistics beyond such black and white terms.

For over fifty years, the Annual Social and Economic Supplement of the Current Population Survey (CPS-ASEC) has been the source of the official poverty statistics for the U.S. Although sampling methodology of the CPS is a rotating panel covering 16 months, the CPS has rarely been used to study household-level poverty transitions. Indeed, the CPS is typically disregarded in the poverty dynamic literature because of its 16-month panel duration, although no documentation exists to prove such a dismissal is warranted. By leveraging the longitudinal component of the survey, this paper provides the foundation and justification for using the CPS-ASEC to study poverty transitions. Using the CPS-ASEC in this setting is desirable for three reasons. First, it is important to characterize these transitions using the CPS-ASEC because the dataset is the official source of income and poverty data in the US. Second, the CPS-ASEC is the only dataset that includes the SPM and thus must be used to study the SPM. And third, because the CPS-ASEC dates back to the 1960s, establishing a framework for studying poverty transitions can ultimately be extended back in time and be useful for the historical analysis of poverty transitions. Therefore, this paper demonstrates how to leverage different definitions of poverty, such as the SPM, to more fully characterize poverty transitions using the CPS-ASEC.

In doing so, it makes six contributions to the literature. First, by linking the CPS-ASEC across two year periods, poverty rates are decomposed into families entering, exiting or remaining in poverty. (i.e. the dynamics of poverty). Quantifying the churn gives new insight into the most vulnerable populations. Second, this paper analyzes how changes in family composition affect poverty rates and, by doing so, provides a more general framework for analyzing family composition changes in the CPS. Third, since the OPM and SPM are both measured using the CPS-ASEC, this paper looks at how poverty transitions and economic vulnerability are different across the two measures. Fourth, focusing primarily on the SPM, the paper characterizes how changes in resource components influence poverty transitions. By leveraging the panel of the CPS, the paper is the first to document the distribution of resource component changes for poverty transitions. Fifth, this paper develops a simple methodology for measuring the relative importance of resource changes to poverty transitions. And finally, having established the foundation for using the CPS-ASEC to study poverty transitions, the paper looks at how life-events, like divorce or increases in medical expenditures, are associated with entering poverty or, in the case of government subsidies, keeping families out of poverty.

The remainder of the paper is structured as follows. Section 2.2 reviews the relevant literature on poverty dynamics and highlights the lack of work using the CPS-ASEC. The section also gives an overview of how the OPM and SPM are constructed with particular emphasis on the key differences between the two definitions. Section 2.3 provides a general overview of the CPS-ASEC and how the panel nature of the CPS-ASEC allows for year-to-year linking for family-level analysis. Section 2.4 demonstrates the reliability of using linked CPS-ASECs to analyze poverty transitions by looking at the patterns of poverty transitions across both OPM and SPM. Section 2.5 then looks at how family composition changes can be studied in this framework. Section 2.6 decomposes the poverty transitions into its resource component changes and provides a method for analyzing the relative influence of resource component changes on poverty transitions. Section 2.7 demonstrate how this methodology can be leveraged to quantify the role real-life events (such as job loss, death, divorce, retirement, receipt of government subsidies) have in pushing people in and out of poverty. Section 2.8 considers the role of imputations and attrition on the findings of this paper. Section 2.9 gives an overall discussion of the findings, areas for future research and concludes.

2.2 Review of literature and definitions

2.2.1 Literature review

The literature on poverty dynamics highlights three important points that are relevant for the present study. First, the literature focuses on how certain events affect the probabilities of entering or exiting poverty as defined by the OPM. Since the OPM is based on cash income, most of the literature finds that employment status and earnings are strongly associated with transitions into and out of poverty. Second, research on poverty dynamics has been conducted primarily with more traditional longitudinal datasets such as the Panel Study of Income Dynamics (PSID) or the Survey of Income and Program Participation (SIPP). Little research has been conducted exploring the longitudinal nature of the CPS-ASEC as it relates to poverty transitions. Third, the literature emphasizes the limitations of the OPM, further showing the need to conduct longitudinal research using the SPM.

A first-order issue in the poverty dynamics literature is estimating the overall probability of entering and exiting poverty in a given year. This literature is most in line with what is done in this study. Mainly researched with data from the 1990s, the literature tends to find a likelihood for the U.S. population of about 4 percent. That is, the unconditional likelihood of any family entering poverty is about 4 percent (Eller, 1996; Naifeh, 1998). Specifically, two studies using SIPP data from the 1980s find annual entry rates of about 3 percent while McKernan and Ratcliffe (2005) find that this rate increased to about 4 percent using PSID from the 1990s. From a poverty exit perspective, Eller (1996) and Naifeh (1998) find that about 23 percent of people in poverty exit poverty on an annual basis in the 1980s, while McKernan and Ratcliffe (2005) find this to be closer to 35 percent during the 1990s. Importantly, the studies use longitudinal surveys with longer timeframes and therefore provide a good benchmark for the results using the CPS-ASEC.

Leveraging the full time frame of these longitudinal survey such as SIPP and PSID, a large part of the literature also focuses on the probability of experiencing poverty across the lifecycle. In particular, the literature shows that, although the unconditional probability of entering poverty is about 4 percent, the likelihood of experiencing poverty at some point in one's lifetime is much higher. Rank and Hirschl (2001) show that the likelihood of experiencing poverty increases significantly with age. Specifically, they find that nearly 30 percent of adults experience poverty by age 30, about 40 percent by age 50, and nearly half of adults by age 65. Moreover, studies find that the likelihood of entering poverty is highest for those under the age of 25 (McKernan and Ratcliffe, 2005; Ribar and Hamrick, 2003; Rank and Hirschl, 2001). Because the CPS captures only a 16-month window, these sorts of transitions are not studied in this study.

Focusing in on the events that are associated with poverty entries and exits, various studies have confirmed that labor supply and earnings changes are the events most commonly associated with poverty entries and exits. McKernan and Ratcliffe (2005) point out that the most important factor leading to poverty entries is loss of employment by household heads as corroborated by various studies (Bane and Ellwood, 1983; Blank and Holzer, 1997). While changes in household composition can also increase or decrease the likelihood of entering poverty, the overwhelming evidence shows that employment is the leading event leading to poverty. Cellini, McKernan and Ratcliffe (2008) provide a thorough review of the poverty dynamics literature and corroborate the finding that employment changes are the most important factor explaining poverty entry and exits. They write, "descriptive analyses using both the SIPP and PSID find that changes in labor supply and earnings are more commonly associated with poverty exits than changes in household structure and composition." The findings of this paper corroborate these findings but extend the literature by quantifying the number of families that fall into poverty because of these household employment and composition change. That is, rather than focusing on changes in likelihood, this paper focuses on describing the actual proportion of families that enter or exit poverty.

Cellini, McKernan and Ratcliffe (2008) make it evident that most work on poverty transitions has been done using the SIPP and PSID. Of the 21 papers they reviewed, only four use datasets outside of the SIPP and PSID, two of them using the CPS-ASEC and two of them using the National Longitudinal Survey of Youth (1979). Taking advantage of the panel nature of the CPS-ASEC is typically dismissed; as Cellini, McKernan and Ratcliffe (2008) note: "with a maximum of two years of annual poverty status on individuals, the CPS-ASEC is generally not well-suited for these types of studies." This paper argues that the CPS-ASEC is in fact useful for studying year-to-year transitions. Since the CPS-ASEC is the source of the official poverty rate for the U.S., it is particularly important to understand how the CPS-ASEC can be used to study year-to-year transitions in poverty status. Moreover, without a thorough analysis of poverty transitions in the CPS-ASEC, it is unclear exactly what can be learned from using the CPS-ASEC.

Using the longitudinal aspect of the CPS-ASEC is common, particularly in studies of labor status transitions, but these are more commonly conducted using the monthly-to-monthly transitions with Basic Monthly CPS samples. Using the CPS-ASEC to analyze transitions is less common though a few studies are relevant in this study. Feng (2013) links CPS-ASECs and speaks to the identification and statistical inference issues regarding a linked CPS with an application to poverty dynamics. While the focus is not on the factors influencing poverty transitions, Feng shows that CPS-ASEC can be used for poverty transition analyses. Hokayem and Heggeness (2014) analyze how people move into and out of near poverty which they define as individuals between 100 and 125 percent of the official poverty threshold. By creating two-year panels of the CPS-ASEC, they find that education levels, labor force status, homeownership and marriage are important factors explaining entrance and exits into near poverty. In sum, though not typically used for the study of poverty transitions, there is some precedence for using the CPS-ASEC to study poverty transitions.

Lastly, the literature continually points out the need to analyze transitions using alternative measures such as the SPM. For example, McKernan and Ratcliffe (2005) argue that "future research needs to leverage alternative definitions of poverty as they can address limitations of the current official poverty measure by accounting for in-kind transfers, taxes, medical expenditures, owner-occupied housing, geographic variation in the cost of living, and cohabitation." Hokayem and Heggeness (2014) also point out that future work should look at the SPM directly.

2.2.2 Defining OPM and SPM

Central to the analysis in this paper are the definitional differences between the OPM and SPM. This section covers the main differences between the two poverty measures (for a detailed review refer to Fox et al. (2015) and Renwick and Fox (2016)).

At its core, the SPM is a series of extensions to the OPM. As explained by the U.S. Census Bureau:

> Concerns about the adequacy of the official measure culminated in a congressional appropriation in 1990 for an independent scientific study of the concepts, measurement methods, and information needed for a poverty measure. In response, the National Academy of Sciences (NAS) established the Panel on Poverty and Family Assistance, which released its report, "Measuring Poverty: A New Approach," in the spring of 1995. In March of 2010, an Interagency Technical Working Group on Developing a Supplemental Poverty Measure (ITWG) listed suggestions for a new measure that would supplement the current official measure of poverty (Renwick and Fox, 2016).

Three main concepts define both poverty measures. First, there are a series of thresholds that define the minimum level of income needed for a family and these depend on family size and family composition. Second, families are defined to capture units of people who share resources. Third, resources for a family are then compared to the poverty threshold to define the poverty status of a given family.

Thresholds

Developed in the early 1960s by Mollie Orshansky, the OPM is an absolute poverty measure that compares the resources of a family to a poverty threshold. When they were developed, the OPM thresholds looked to monetize the minimum yet adequate level of family consumption. The key insight in the process was a finding from the 1955 Household Food Consumption Survey which estimated that the average expenditure on food for families, regardless of income level, accounted for one third of their total family income after taxes. With this, Orshansky then derived the cost of a minimum yet adequate diet for a family of four (2 parents and 2 children) from food plans developed by the Department of Agriculture. Multiplying this cost by three then gave the poverty threshold for a family of four. Adjustments for family characteristics were then made, varying on the family composition (number of adults versus number of children) and family size. Since 1963, for all intents and purposes, these thresholds have only been adjusted annually to account for inflation which today is done using the Consumer Price Index for All Urban Consumers (CPI-U).[1]

The poverty thresholds for 2016 are presented in Table 2.1. The poverty thresholds are a function of the number of people in the family crossed with the number of related children under 18 which is, in essence, the number of dependents in the family. As the family size increases, the poverty threshold increases. And, within a given family size, as the number of children increases, the poverty threshold decreases. The logic here is that younger children tend to need less food than older children or adults. Finally, for families of one or two people, the number of people over 65 decreases the threshold. The assumption here is that these people tend to be retired and thus need less energy for their daily activities. The underlying logic of the table demonstrates the outdated assumptions about workforce; that

[1] For a complete history, see The Development of the Orshansky Poverty Thresholds and Their Subsequent History as the Official U.S. Poverty Measure by Gordon M. Fisher.

is, today there tends to be less physical labor on average than fifty years ago.

On the other hand, the SPM estimates a threshold based on the 33rd percentile of expenditures on food, clothing, shelter and utilities calculated from the Consumer Expenditure Survey (CE) for the U.S.. By doing so, the SPM thresholds more accurately capture the costs of living of the typical U.S. household. As with the OPM, the thresholds are adjusted for family size and composition. Most importantly, these SPM thresholds are adjusted geographically while the OPM thresholds are the same throughout the US. The SPM thresholds are adjusted using 5-year estimates of median gross rents for different metropolitan statistical areas and non-metro areas for a total of 358 adjustment factors.[2]

Family units

The two poverty definitions differ in their treatment of families, or more specifically, family units. It is necessary here to distinguish a few technical definitions. In the CPS-ASEC, households consist of the people within a housing unit or dwelling. A household can consist of one or more families. Family units are individuals who self-report to living within the same household and are related by birth, marriage, or adoption (Renwick and Fox, 2016). All other individuals within a household who are not members of a family are considered unrelated individuals.

The goal of defining a family unit is to capture a group of people that share resources. The OPM considers two distinct resource-sharing units: families and unrelated individuals. Treating families as a resource-sharing unit is uncontroversial. However, one of the criticisms of the OPM is that excluding unrelated individuals from families excludes people who are arguably sharing resources with the rest of the family. The most noticeable omission of the OPM family definition are unmarried partners (or cohabiting couples). The SPM adapts the OPM definition to account for different possible resource arrangements. Specifically, SPM families include unmarried partners and their relatives, co-residing unrelated children, and foster children. In essence, the SPM takes people considered to be unrelated individuals under the OPM and classifies them as a resource-sharing unit when appropriate. All other unrelated individuals in a household are considered as their own resource-sharing unit.

[2] These thresholds are available at https://www.census.gov/hhes/povmeas/methodology/supplemental/overview.html.

For example, a renter living within someone's household would be considered an unrelated individual under both the OPM and SPM.[3]

In this paper, the unit of observation will be the family as defined by the OPM and the SPM. Effectively, the sample will consist of families of two or more people and individuals (families of one). Moreover, to be complete with the terminology, households will refer to the group of people residing together while the term dwelling will refer to the physical unit in which people live (i.e. households are people while dwellings are the homes in which those people live).

Resources

The definition of resources of a family also vary across OPM and SPM, though the OPM-defined resources are completely encompassed in the SPM. Back to the history of the OPM, when Orshansky was looking to measure the total resources of a family, the CPS was the only reliable source of nationally representative data on income. At the time, and to this day, the CPS asks individuals about their before-tax cash income from various sources in the previous calendar year. In large part, it was the availability of data that led to the OPM adopting a definition of total resources that used before-tax income. While the major component of most families' total resources are wages/salaries, the CPS-ASEC also asks about 19 other components. In this paper, wages/salaries income will be analyzed separately from "other cash income" which will be the catch-all term for all other components. On average, the largest components of other cash income are pension/retirement income and Social Security income. Table 2.2 presents the complete accounting of the cash income resources used in SPM resource definition. The "total cash income" in Table 2.2 are the components used in OPM's definition of total resources.

As is clear in Table 2.2, resources for the SPM attempt to measure a more complete picture of the resources that a family actually has at their disposal, taking the OPM resources as its starting point. The goal of the SPM resource measure is to monetize all the resources that a family has to spend on food, clothing, shelter and utilities. Therefore, SPM resources add non-cash public benefit recipiency to the total cash income of the OPM.

[3] Arguably, even renters may be sharing resources, to an extent, with the rest of the household which is why many researchers choose to use households as the unit of analysis (McKernan and Ratcliffe, 2005).

The CPS collects information on the yearly receipt amount of the Supplemental Nutrition Assistance Program (SNAP), and the Low-Income Home Energy Assistance Program (LIHEAP). For Supplementary Nutrition Program for Women Infants and Children (WIC), National School Lunch Program and housing subsidies, the CPS asks about the receipt of the subsidy but imputes the amount.

From these resources, the SPM subtracts resources that are not available for a family's consumption on food, clothing, shelter and utilities. The first, and perhaps the most obvious, are taxes. The Census Bureau imputes various different tax components using a model that simulates the tax liability of each family. Tax liabilities include federal and state income taxes, property taxes, Federal Insurance Contributions Act taxes (FICA) and federal retirement payroll deductions. To these liabilities, tax credits are subtracted out: the Earned Income Tax Credit (EITC), Child Tax Credit (CTC) and Additional Child Tax Credit (ACTC). In other words, these credits are added into a family's total resources.

The second set of components that are subtracted from SPM total resources are necessary expenses. The SPM accounts for three in particular. The first set are work expenses. Using data from the SIPP, the SPM takes a national estimate of the reported expenses on commuting and work-related expenses (for example, uniform purchases).[4] Second, child care expenses are directly asked in the CPS and subtracted from the total resources. The third set of expenses are child support paid which are also collected in the CPS.

The last set of expenses that are subtracted from SPM total resources are medical expenditures. Medical expenditures have been acknowledged as a significant part of a family's budget and therefore were added as a question to the CPS. The CPS collects the amount a family pays for health insurance premiums and other medical expenditures not covered by insurance (i.e. prescription drugs and copayments). Medicare Part B premiums are subtracted for those reporting recipiency of Medicare. As Table 2.2 shows, there are nearly 40 components that constitute total resources under the SPM.

As a result of these definitional differences, the OPM and SPM exhibit different patterns of overall poverty in a cross-sectional context. The most obvious is that the SPM rate is typically about half a percentage point higher than the OPM rate. In 2015, the SPM rate was 14.3 percent while the OPM rate was 13.5 percent which amounts to around 2

[4] Specifically, the SPM uses a value of 85 percent of median weekly expenses multiplied by the number of weeks respondents in the CPS reported working in the year.

million more people in poverty under the SPM (Renwick and Fox, 2016). Children and elderly populations (65+) are the groups with the most pronounced difference across the definitions. SPM rates for children are about 4 percentage points lower than OPM rates, mainly due to the inclusion of non-cash benefits and tax credits that aid families with young children (Renwick and Fox, 2016). Meanwhile, SPM rates are higher for elders by about five percentage points due mainly to the inclusion of medical expenditures (Renwick and Fox, 2016). To date, these patterns have only been analyzed in a cross-sectional manner and it remains to be analyzed how patterns in poverty transitions differ across the two definitions.

2.3 Data

2.3.1 Linked CPS

The CPS has been the official source of employment statistics since 1940 and the source of the official poverty rate since the 1950s. While the poverty rate is published once a year, unemployment rates are released on a monthly basis. Underlying this difference is the CPS sampling methodology which is central to understanding the data used in this paper. The CPS is a rotating panel representative of the civilian, household-based population of the U.S. The monthly CPS samples, which are the basis of the employment statistics, are known as the Basic Monthly CPS. Various topical supplements are fielded throughout the year that encompass a diverse set of topics from child support and tobacco usage to food security and voting. For a more complete overview of these supplements, see Flood and Pacas (2016).

The most well-known of these supplements is the Annual Social and Economic Supplement. Fielded primarily in March, the CPS-ASEC includes a large set of income questions that provides the basis of the measure of family income used in developing the official poverty rate of the U.S. In order to construct the panels necessary for this analysis, linking keys are provided by the Census Bureau that longitudinally identify people and dwellings in the CPS. The procedure used here generally follows that outlined in Rivera Drew, Flood and Warren (2014). Since the linking process is well-documented, the process is not covered in depth here. More important for this analysis is that there are two main issues with linking that affect the analysis of poverty transitions. First, the rotation pattern of the CPS allows for a given respondent to be found in at most two CPS-ASECs. Respondents for the CPS are selected at the dwelling-level in a given month; more specifically, it is the physical dwelling that is selected and not the set of people living at that dwelling. Once a dwelling is selected for participation, the entire household roster is interviewed (typically one respondent responds for the entire household). The CPS is administered to this dwelling the next three months for a total of four interviews in a given year. After these four months, the dwelling is out of the CPS for eight months at which point the CPS is given to the dwelling another four times for the same calendar months. For example, a dwelling selected for the CPS in January of 2017 will be interviewed in February, March and April of 2017 and then January through April of 2018. Thus, a dwelling will only participate in the CPS-ASEC twice. Second, attrition in the CPS affects poverty analysis. The CPS follows a physical dwelling as opposed to a particular set of individuals therefore, when people move away from the dwelling, those people are lost from the sample. The common issue with attrition in the CPS is that it is non-random and is further explored below.

A simple approach to merging CPS-ASECs across years is to use the individual linking keys described above to find an individual across the two years. Once a link has been made, a researcher typically validates the link on typical time invariant characteristics such as sex and race/ethnicity[5] or characteristics that change in predictable ways such as age (i.e. age in t_2 = age + 1 in t_1). Two reasons lead a researcher to drop a linked observation. First, the characteristics of the person are not the same across the two time periods and therefore the link is likely a false-positive. Second, there is no data on the person for one of the two time periods so no corroboration can be made. By dropping these observations, a researcher ensures a valid linked-sample.

[5] In the CPS-ASEC, sex and race are asked in the first interview and typically do not change across subsequent interviews. Field representatives do ask about household roster changes but not about changes to an individual's sex or race. Survey redesign periods are the only instances where these questions are re-asked to respondents outside of their first interview.

2.3.2 Linking families

For person-level analysis, this methodology is sound. In order to properly identify family-level changes, however, one needs to refine this methodology. Family composition changes necessarily imply the arrival or departure of a person and thus there will necessarily be missing data from one time period. Typical linking validation techniques will result in the masking of real family composition changes. Consider the example in Table 2.3. The arrival of a newborn would be missing data in the first time period for the child born between time period t_1 and t_2. If the researcher remains unaware of the error then the poverty threshold will be assigned to an incorrect family type. Specifically, in the first time period, the family has 3 people with 1 child and a corresponding poverty threshold of $19,078. In the second time period, the addition of a newborn results in a family type of 4 people with 2 children with a corresponding threshold of $24,339. However, if the newborn is not included because it has been dropped in the linking process, the data will show a family of 3 people with 1 child that has been assigned a poverty threshold of $24,339 which corresponds to a family of 4 people with 2 children.

A slightly more complicated example is considered with the death of a family member. Consider an elderly couple where the elder male householder dies between t_1 and t_2. In the second time period, the widow is still in the dwelling (and now the householder) but now her adult child has moved into the dwelling. In this example, a simple validation rule of dropping any person with missing data in either time period would overlook the death of the elder householder and the arrival of the adult child. Indeed, that approach would treat the family as having only one person (the elder spouse turned widowed householder) in both time periods. Once again, the correct poverty thresholds should correspond to 2 people (with householder over 65) which is around $14,500 in both years. Instead the erroneously linked sample would show a 1 person household with the incorrect threshold.

The researcher at this point must choose one of two suboptimal solutions. The first option is to leave the data with the erroneous family sizes and poverty thresholds. In this case, analysis correlating family types to any outcome will be unreliable given the inclusion of these erroneous family types. The second option is to adjust the family types and thresholds; that is, give the families the correct poverty thresholds for the erroneous family sizes. This option is less optimal still. Consider the two examples in Table 2.3.

For the arrival of a newborn, the erroneous family size is 3 people with 1 child which would have a poverty threshold in 2016 of \$19,318. For the purpose of highlighting the error, consider that a family has a total family income of \$19,000 in t_1 and \$19,500 in t_2. Because of the addition of the newborn, that family should remain in poverty in both time periods. But, given the erroneous family size and an updated poverty threshold, that family exits poverty in t_2. Similarly, in the second example, the correct family types and thresholds would classify the family as poor in both time periods. On the other hand, the erroneous family type with an updated threshold would be classified as nonpoor in both time periods. Overall, these examples illustrate the potential errors that can be easily overlooked in linking CPS-ASECs across years. These errors are more pronounced when conducting analysis at the family-level as in the case of poverty measurement. In order to fully utilize the potential of a linked CPS, these family composition changes should be kept as part of the final sample.[6] In this paper, family composition changes are preserved.

2.3.3 Descriptive statistics - Poverty transitions reveal new disproportionalities

This paper links consecutive CPS-ASECs from 1996 through 2016, maintaining the family composition changes noted in the section above.[7] However, when studying the SPM, the analysis uses only 2010 through 2016. The simple reason for this is that, as noted above, the SPM was first made available in 2010. The resulting sample size is 127,146 families for the 2010-2016 SPM sample and 127,404 for the 2010-2016 OPM sample (see Table 2.4 and Appendix Table A.1).[8] The OPM analysis conducted for 1996-2016 includes 412,743 families.

To clarify terminology, the first year a family is observed in the data will be referred to as t_1 and the second year will be denoted by t_2. Thus, the sample ends with a t_1 of 2015

[6] The two papers that link CPS-ASECs to study poverty transitions (Feng (2013) and Hokayem and Heggeness (2014)) make no explicit reference to the issue of linking families. Indeed, most papers that link CPS-ASECs do not explicitly speak to this issue. However, most of these studies are not concerned with family-level transitions.

[7] As noted in Flood and Pacas (2016), the CPS-ASEC provides a continuous set of two-year panels from 1996 on. The CPS-ASEC 1995 is not possible to link to 1996 and thus an undesirable break in series occurs.

[8] Differences in the 2010-2016 SPM and OPM samples are due to the different family definitions used by each measure but the same number of individuals are in both samples.

and a t_2 of 2016. Poverty rates will refer to the poverty rate using t_2 observations such that a given year's poverty rate can be decomposed into families entering poverty plus families remaining in poverty. Poverty transitions will generally be referred to in the following manner. Those families not transitioning, those below or above the poverty threshold in both time periods, will be referred to as Poor-to-Poor (P-P) or "Always Poor" and NonPoor-to-NonPoor (NP-NP) or "Never Poor." Similarly, families transitioning will be denoted by Poor-to-NonPoor (P-NP) for those exiting poverty and NonPoor-to-Poor (NP-P) for those entering poverty.

Table 2.4 presents statistics on selected demographics for the SPM sample. These same set of statistics are available for the OPM in Appendix Table A.1.[9] Table 2.4 highlights how studying poverty transitions using the linked CPS-ASEC can reveal new insights into the most vulnerable populations. For example, it is known that African American and Hispanics are disproportionately poor under both SPM and OPM definitions.[10] This trend continues to be true in this longitudinal setting but a more interesting pattern emerges. African Americans and Hispanics disproportionately become poor but they are even more disproportionately represented in the group that remains poor. Specifically, African Americans represent about 8 percent of the total SPM sample and about 12 percent of the poverty transitioners (NP-P and P-NP). This rate is about double for those remaining in poverty at around 15 percent. This pattern is similar for Hispanics where the full sample is about 10 percent Hispanic but the P-P population is about 21 percent Hispanic. The non-citizen population exhibits this same pattern.

These statistics suggest that the underlying reason for being in poverty for these groups are different than for other groups. Though this paper does not focus on understanding these racial disparities, these statistics show that the linked CPS-ASEC can identify more economically vulnerable populations. The remainder of this paper unravels the ways in which the linked CPS-ASEC can be used to give insight into the year-to-year changes in family resources.

[9] Overall, the patterns noticed between SPM and OPM are the same and therefore does not warrant separate discussion.

[10] Note that the race/ethnicity assigned to a family unit is typically that of the householder, as is customary in official Census reports. For complex households with more than one family, the designation is typically made with the eldest member of the family.

2.4 The Churn of Poverty Rates

This paper is the first to document the churn of the official poverty rate. Figure 2.1 plots the OPM flows between 1996 and 2016 as a percentage of total U.S. population.[11] The top line in Figure 2.1 graphs the overall OPM rate in t_2 for the linked CPS-ASEC sample. The most striking pattern in Figure 2.1 is that churn accounts for about 50 percent of the total poor population at any given time while 50 percent of the poor population are families who remained in poverty in both time periods.

Through this analysis, it is also possible to visually inspect what drives the overall poverty rate. For example, in the pre-2000 time period, one notices an overall poverty rate drop from around 11 percent to about 9.5 percent. As seen in the poverty transitions, this was driven by an overall drop in the P-P group as well as an increase in the poverty exits relative to the poverty entries. A clear shift is seen in 2000, at the beginning of the recession, where the poverty entries exceed the exits. The Great Recession (2007 through 2009) period can be explained by a rise in the P-P group and poverty entries, relative to the poverty exits until 2014 where the poverty exits surpassed the poverty entries for the first time since 2007. In sum, taking advantage of a linked-CPS allows for the study of poverty churn and highlights that the rate of churn is about equal to the rate of people who remain in poverty. The analogous analysis for SPM is presented in Figure 2.2 for 2010 through 2016. As opposed to the OPM, the SPM shows that poverty transitions are higher than those who remain in poverty.

To more clearly make the comparison, Figure 2.3 compares the flows of poverty transitions using the OPM and SPM. The first trend to notice from these figures is that the SPM poverty rates are higher than the OPM rates, which is in line with the published rates (for example, the 3-year OPM average from 2011 to 2013 is nearly one percentage point lower than the SPM rate for the same time period (14.9 versus 15.9)). The second trend is that the churn under the SPM definition is higher than that of the OPM. Under the SPM, the P-P stays around 7 percent but the transitioners are closer to 10 percent of the total population, or closer to 60 percent of the poor population.

What drives the higher churn rate under the SPM? Two potential sources are obvious starting points from the differences between OPM and SPM definitions. First, family units

[11] The denominator for all groups is the total U.S. population in t_2.

under the SPM are more expansive and thus may lead to more family composition changes between years. Second, SPM resource components are more comprehensive than under the OPM definition. In particular, the inclusion of expenditures that may be more volatile, such as medical expenditures, may drive the higher churn rate. These explanations are investigated in more details in subsequent sections.

2.4.1 Who becomes poor? Even the really rich and the middle class.

A different way of looking at poverty transitions is to look at the relationship between resources in one year and the likelihood of being in poverty in the next year. While the overall churn rate reveals the number of people who move in and out of poverty, this analysis reveals which families are most likely to be in poverty. The main finding of this section is that, under both the OPM and the SPM, the likelihood of being in poverty from one year to the next is non-zero, regardless of a family's resource level in a given year. In other words, no family is immune to negative economic shocks.

Figure 2.4 plots the percent of families in poverty in t_2 given their resources in t_1. Resources are measured relative to each family's poverty line (threshold). For example, a family of four people with $40,000 in resources will be about 1.5 times over their poverty line while an individual with $40,000 will three time over their poverty line.[12] Thus, by using the ratio of resources to poverty line, families with similar resources per person are grouped together. The top panel of Figure 2.4 plots the percent of families that are in poverty in t_2 for each percentile of ratio of resources to poverty line. The bottom panel gives the distribution for the ratios of resources to poverty line. For example, under the OPM definition, four percent of families at the 50th percentile of resources in t_1 (which is a ratio of about 2.3 times over the poverty line) are in poverty in t_2.

What does this analysis imply for poverty transitions? In general, it shows that all families are vulnerable to falling into poverty though the relatively wealthy families are much less likely to do so. More importantly, even families who might be considered economically

[12] Resources are measured as total cash income for the OPM and total resources for the SPM (as described in depth above). The thresholds refer to the respective poverty threshold for each family under each definition.

secure transition into poverty at a relatively high rate. Consider the lower end of middle-income families, or families who are two times over their poverty line.[13] These families are found around the 30th percentile of resources which implies that about 9 percent of these families fall into poverty in t_2. Similarly, under the SPM, families around two times their poverty line (43rd percentile) are about 10 percent likely to fall into poverty. Thus, even those families who are able to make their way into the middle-class still face a one-in-ten chance of falling into poverty. Of course, the probability of falling into poverty drops as resources increase but even two to three percent (OPM and SPM, respectively) of families at the top percentile of resources (about 35 times and 20 times over the poverty threshold for OPM and SPM, respectively) fall into poverty.

On a more methodological note, Figure 2.4 also highlights important differences between the OPM and SPM. The relatively flatter pattern of the SPM line relative to the OPM line shows that the risk of being in poverty is more evenly distributed under the SPM than under the OPM. Consider those in poverty in t_1 which is denoted by the vertical line which corresponds to those with a resource-to-threshold ratio below one). For the OPM, for all those in poverty in t_1, the likelihood of being in poverty in t_2 is over 40 percent. Indeed, the likelihood for those in the bottom 5 percentiles is over 50 percent. In contrast, the likelihood for those in poverty in t_1 under SPM is closer to 35 percent for those percentiles right below the poverty line (between the 10th and 14th percentile) and around 45 percent for those below the 10th percentile. The implication is that the OPM gives a different picture of the economic mobility of those in poverty; that is, families in poverty are less likely to exit poverty. However, including other resources with the SPM shows that the OPM overstates the level at which families stay in poverty.

Above the respective poverty lines, the likelihood of poverty under the OPM is continuously lower than the SPM. This finding reinforces the idea that the SPM shows more people moving in and out of poverty and, more importantly, that these people are found across the entire resource distribution. Indeed, for the OPM, the likelihood for families two times over the poverty threshold (30th percentile) is about 10 percent but about 20

[13] This analysis uses the definition for middle-income from Pew Research Center (2016). Specifically, these categories are $24,042-$72,126 for a family of one, $34,000-$102,001 for a family of two, $41,641-$124,925 for a family of three, $48,083-$144,251 for a family of four, and $53,759-$161,277 for a family of five. On average, the OPM thresholds for each of these family sizes is about two times below the lower limit and about 6 times below the upper limit.

percent for families in the 30th percentile of the SPM. Around the 50th percentile of the ratio distribution, both the SPM and OPM flatten to around 5 percent and 2 percent, respectively.

The bottom panel of Figure 2.4 highlights that, on average, the OPM and SPM ratios are similar up to about the 13th percentile at which point the OPM ratios exceed the SPM. At the higher percentiles, the OPM ratios exceed the SPM ratios by larger amounts with the 100th percentile of OPM exceeding the SPM ratio by about 75 percent (an OPM ratio of 35 versus an SPM ratio of 20). The reason for this difference is due mainly to the definitional differences of resources (see Table 2.2) as opposed to the different threshold definitions.[14] One pattern that emerges from the resource-to-threshold graph is the role that taxes play across the SPM and OPM definitions. Intuitively, the SPM captures the resources that families have at their disposal; at the higher percentiles of income, the tax liability of a family reduces their cash income significantly. While there are many other SPM components that contribute to this difference, taxes are the component that most affects the higher percentiles of income.[16] Overall, these findings highlight the long tails of resources under both definitions and shows that the OPM resources are increasingly larger than the SPM resources at the higher percentiles of income.

2.4.2 Changing picture of poverty with alternate thresholds

One question that continuously arises when discussing poverty is the sensitivity of the poverty rate to alternate thresholds. Researchers have called this the "near poor" population (Hokayem and Heggeness, 2014). A simple exercise is to look at alternative poverty thresholds and consider how the pattern of transitions change across the income distribution. Figure 2.5 and Figure 2.6 duplicate the top panel of Figure 2.4 but also add the likelihood of being in poverty if the poverty thresholds were redefined at 125% above their current levels. Figure 2.5 shows that the OPM exhibits a nearly parallel shift up to the

[14] Though defined differently, the SPM thresholds are on average rather close to the OPM thresholds. For example, the 2015 SPM thresholds had a national average of about $24,440 for a family of 2 adults with 2 children.[15] Meanwhile, the OPM threshold for a family of that type in 2015 was $24,339.

[16] One final pattern to notice is that the bottom two percentiles under the SPM are negative while the bottom percentile for the OPM is negative. Though this is not an important finding, it is worth noting that it is possible for families to show negative resources under both definitions. While the SPM has numerous resource components that can drive this negative resource (see Table 2.2), a negative cash income under OPM is typically driven by farm/business income.

40th percentile at which point this new line converges with the original OPM line from Figure 2.4. More specifically, around the 10th percentile, the likelihood of being in poverty is about 40 percent under the normal OPM thresholds but about 55 percent for 125% of the OPM poverty threshold. On the other hand, Figure 2.6 shows a much larger jump below the 40th percentile for the SPM. At the lowest 5 percentiles, the shift is about 10 percentage points but swells to about 30 percentage points at the 100% poverty line (14th percentile). The lines then begin to converge around the 125% poverty line (22nd percentile) and ultimately converge around the 50th percentile. These findings demonstrate that a more complete resource definition, such as the SPM, leads to a higher percentage of families living just above the poverty line.

Ultimately, this section gives insight into why the SPM exhibits higher churn but also how the definitions give a different look at the economic vulnerability of families. The question to be answered now is why families are likely to fall in and out of poverty? What role does family composition play and what role do changes in resources play? Most importantly, how do real-life events push people in and out of poverty? The rest of this paper proposes new methods for answering these questions.

For simplicity of exposition, the paper now focuses solely on the SPM as the definition of poverty. It is important to note that the SPM includes total cash income as one of its resource components which is the only resource component in the OPM definition. Thus, in looking at resource changes, this paper will treat cash income separately to highlight the patterns of the OPM.

2.5 Characterizing the relationship between family composition changes and poverty

2.5.1 Descriptive statistics

Holding all resource changes constant between one year and the next, two factors can affect a family's poverty status in the subsequent year. First, the poverty thresholds are adjusted each year for inflation and thus, all else remaining equal, the inflation-adjusted threshold could be responsible for a change in poverty status. For example, in 2015, the poverty threshold for a family of 4 (2 adults and 2 children) was $24,036. In 2016, this threshold was \$24,339. Therefore, if a family of this composition had an income of \$24,300 in both 2015 and 2016, the family would fall into poverty in 2016 based solely on the inflation adjustment. As later sections will show, changes in income are rarely zero and thus changing thresholds to adjust for inflation is typically not a source of poverty entries.

The second factor that can affect poverty transitions when all resources are held constant is a change in family composition. As Table 2.3 shows, a family of 3 people in t_1 that has a newborn (and thus 4 people in t_2) moves their poverty threshold from \$19,078 to \$24,339 (in 2015 and 2016, respectively). Thus, the addition of a baby can be responsible for a transition into poverty because of the increased need for resources as indicated by the higher poverty threshold for the now-larger family. Conversely, a loss of a family member will typically lead to a lower poverty threshold and thus can be responsible for a poverty exit. Overall, any family composition change could lead to a different poverty threshold.

Table 2.5 demonstrates that about 13 percent of all families experience a change in family composition between one year and the next (see Appendix Table A.2 for the comparable OPM statistics). In this analysis, family composition change is defined by the arrival/departure of a family member.[17] About 7 percent of families experience a departure of a family member, about 5 percent of families have a family member arrive, and a little under 1 percent of families have both an arrival and departure. A few patterns are worth mentioning. Those families entering poverty have a proportionally higher rate of family member departures (9.67 percent) than those never in poverty (6.85 percent). Second, these departures are mostly adult departures. Families entering poverty have nearly two percentage points higher adults departing than those never in poverty (7.80 percent versus 6.00 percent). Elderly departures for families entering poverty are twice as likely than they are for those never in poverty (2.44 percent versus 1.02 percent). To the extent that can be noted in the data, divorce/separation is about 1 percentage point higher for families entering poverty than those never in poverty (1.73 percent versus 0.67 percent).[18] Finally, the percent of families losing a spouse (widowed) is about 1 percentage point higher for those entering poverty than those never in poverty (1.50 percent versus 0.66 percent). In

[17] In practical terms, an arrival is defined as the presence of a person in t_2 not in the data in t_1 while a departure is the presence of a person in t_1 but not found in the data in t_2.

[18] Divorce/separation is noted by the departure of a family member whose marital status was married and then the remaining spouse changes their marital status to divorced/separated. For widows, the marital status changes to widowed.

general, the arrival of adults is more common for families exiting poverty than those never in poverty (5.19 percent versus 3.43 percent). In sum, these patterns suggest that family composition changes are more common occurrences for families transitioning in and out of poverty than those not transitioning and thus could be an important factor in explaining poverty transitions.

2.5.2 Counterfactuals

In order to separate the role of family composition in poverty transitions, it is necessary to tease out the effect of concurrent resource changes. A simple intuitive counterfactual can be developed. Consider that a family is in poverty when family resources are less than their corresponding poverty threshold. Adjusting resources and thresholds so that there is no effect of inflation (i.e. CPI-deflated), then it is possible to isolate the effect of family changes from resource changes. The counterfactuals can be constructed as follows. Hold the resources of a family at the t_1 level and then only allow their threshold to change to their corresponding t_2 level:

$$(\text{Resources } t_1/ \text{ Threshold } t_2) \leq 1 \rightarrow \text{Poverty Rate from Family Composition Changes.}$$

For the role of resource changes, the inverse counterfactual can be calculated:

$$(\text{Resources } t_2/ \text{ Threshold } t_1) \leq 1 \rightarrow \text{Poverty Rate from Resource Changes,}$$

which captures the poverty rate in t_2 if only resources had changed and family composition stayed at its t_1 state.

Table 2.6 presents the results of this counterfactual analysis and shows that the family composition changes result in less families falling into poverty than resource component changes. The overall SPM rate for t_1 is 13.94 percent. If only family composition changed, the SPM rate would not be statistically different at 14.07 percent. If only resources changed, the SPM rate would by about 0.8 percentage points higher and statistically different at 14.74 percent. Therefore, despite family composition changes being rather common (recall that about 13 percent of families experience a composition change), these changes alone do not lead to a higher poverty rate. Resource changes have a larger effect on the poverty rate. The overlap of family composition changes and resource changes is clear and even logical.

When an adult family member leaves the household, they also take their resources with them. Thus, family changes alone, through their changes on the relevant poverty threshold, are not a largely influential source of poverty transitions. However, the concurrent change in resources associated with family composition changes are more influential. In order to more fully understand the role of resource changes, more structure must be given to the analysis.

2.6 Decomposing poverty transitions into resource components

This section unravels the changes in resources from one year to the next. To simplify the analysis, this section focuses only on those families entering or exiting poverty. The overarching questions posed here is: What are the resource changes that drive poverty transitions and which resource components are the most influential in driving these transitions? Figure 2.7 graphs the distribution of family resource-to-threshold deficit (i.e. total SPM resources minus their corresponding SPM poverty threshold value). The overall SPM poverty rate is noted at 14.2 percent (noted with the vertical line (a)) and over nine percent of families are $5,000 or less above their threshold. The implication here is that, if these families were to receive a shock to their resources of $5,000 and all else remained the same, the poverty rate would increase to about 23.5 percent. Yet the national poverty rate has never increased by more than about a percentage point in any given year. Logically, resources fluctuate for families in both directions and, as the churn of poverty rates has already shown, families move in and out of poverty at about the same rate in any given year between 2010 and 2016.

2.6.1 Descriptive statistics

Table 2.7 presents the average change in SPM resource components by poverty transition type. SPM total resources are broken down by the categories in Table 2.2 though total cash income is further broken down into income from wages/salaries and all other cash income (see Appendix Table A.3 for the t_2 averages).[19] On average, a family entering poverty loses about $29,000 in total resources relative to their t_1 total resources. This loss comes mainly from losses in family wages/salaries (about $18,000) as well as other cash income (about $15,000). By definition, tax liability is also decreased when income drops; specifically, taxes drop by nearly $6,000. For expenditures, medical expenditures increase by about $2,000 for families entering poverty while necessary expenses actually decrease on average by about $300. Necessary expenses are, by construction, a function of child care and work expenses. The drop in wages/salaries signals that there is likely a loss of a job or of work hours and hence there is likely a decrease in the child care and work expenses. Finally, government subsidies increase nearly $150 on average for families entering poverty which would be expected for these newly poor families.

Families exiting poverty exhibit a symmetrical set of changes in resource components. On average, the gain in total SPM resources is about $28,000. Families exiting poverty gain about $32,000 in cash income, on average, while families entering poverty lose about $33,500. The symmetry of these changes at the average is surprising given that, a priori, one would expect the changes pushing families into poverty may be different from those pushing families out of poverty. However, averages provide a limited picture of these changes.

In order to best understand the full distribution of these changes, Figures 2.8 - 2.14 plot the distribution of resources changes for these poverty transitions. The symmetry between total SPM resource changes is noted in the distribution in Figure 2.8. 99 percent of poverty entries experience a decrease in total SPM resources while 99 percent of poverty exits see an increase in total SPM resources. This finding is key. Figure 2.8 shows that only 1 percent of families lose resources and still exit poverty while only 1 percent of families gain resources and still enter poverty. The magnitudes of these changes also matter. For these 1 percent of families who enter poverty despite gaining resources, the average change in total resources is only about $3,000. For those exiting poverty, the families lose about $6,000 in total resources. Relative to the changes noted in the entire distribution (ranging from a loss of over $250,000 to gains of over $250,000), these changes are small.

Figure 2.8 also demonstrates that the change in total resources for about 30 percent of families exiting poverty is less than $10,000 while around 55 percent of families exiting

[19] Note that total cash income is the only resource component included in the OPM and so patterns noticed from this resource is largely reflective of OPM patterns.

poverty experience increases of over $20,000. There is a significant proportion of families who exited poverty but are still close to the poverty line. For example, families that increased their resources by only $5,000 amount to about 12 percent of all families exiting poverty. The symmetry between those exiting poverty and those entering poverty implies that a similar story is true for those families entering poverty. In sum, the distribution of change in total resources shows that a majority of families exiting poverty are not experiencing large increases in resources but are arguably gaining an amount that places them in a vulnerable position but better off than before. From Figure 2.4, it is known that the lower the ratio of resource to poverty line, the higher the likelihood of falling into poverty. Therefore, the smaller the change in total resources for those families exiting poverty, the higher the probability there is of these families falling back into poverty.

The symmetry of these changes is also noticed in the distribution of individual SPM resource component changes. The distributions of the changes in family wages/salaries for families exiting poverty show that families gain, lose and see no change in their wages/salaries (see Figure 2.9). Nearly 40 percent of families experience no change in the wages/salaries. About half of all families exiting poverty had an increase in their wages/salaries and about 10 percent of families that exited poverty experienced a decrease in their wages/salaries. Families entering poverty exhibit an almost identical symmetrical pattern not just in the distribution but also in magnitude.

Figure 2.9 shows that about 28 percent of families exiting poverty gain the equivalent of a full-time year-round minimum wage job or less (48th percentile to the 67th percentile of the change in wages/salaries), implying that over a quarter of families exiting poverty are likely not adding full-time jobs.[20] The 10 percent of families losing wages/salaries highlight how the more complete accounting of resources included in the SPM lead to different conclusions about poverty transitions. Under the OPM, except in vary rare cases where families experience a loss from business/farms, all families exiting poverty experience increases in wages/salaries. Thus, by including expenses, the SPM shows that there are about 10 percent of families exiting poverty that lose wages/salaries but gain enough in other resources to exit poverty.

[20] Minimum wage jobs year-round full-time: $7.25 x 40 hours x 52 weeks = $15,080.

The distribution of changes for other cash income (see Figure 2.10) is smaller in magnitude to that of wages/salaries and shows that no families have a zero change in other cash income. 75 percent of families exiting poverty experience an increase in other cash income while about the same percent of families entering poverty see a decrease. The simple reason for why families experience a change of some sort is that there are about 19 different sources comprising other cash income.

Figure 2.11 and Figure 2.12 plot the distributions of taxes paid and necessary expenses, respectively (note that the axes on these are different than Figures 2.8 - 2.10). As these resource components are a function of cash income, it follows that the distribution of these changes are similar to that of wages/salaries.[21] Specifically, about 63 percent of families entering poverty pay less taxes while about 20 percent have no change in their taxes (Figure 2.11). For necessary expenses, about 45 percent of families exiting poverty see an increase while about 35 percent see no change (Figure 2.12). The symmetry between poverty entries and exits is again noted.

The change in government subsidies merits more discussion as poverty typically increases the likelihood of being eligible for government subsidies. Surprisingly, nearly half of all families entering poverty and half exiting poverty see no change in government subsidies at all (see Figure 2.13). The lack of a change implies that many of these families were not receiving government subsidies in either time period. Another surprising finding is that families entering poverty experience an increase in government subsidies in nearly the same percentage and magnitude as families exiting poverty. That is, about 20 percent of both families exiting and entering poverty see an increase in government subsidies. The same is true for the decrease in government subsidies. One potential explanation for this pattern is that government subsidies are helping families out of poverty. It is possible that the 20 percent of families that see increases in government subsidies and exit poverty are exiting poverty because of these subsidies. Similarly, the 20 percent of families that lose government subsidies and enter poverty could be entering because of the loss of these subsidies. The following section looks into this matter in more depth.

The distributions of changes in medical expenditures show that medical expenditures increase and decrease for both exiting and entering families (see Figure 2.14). However,

[21] Taxes change as total cash received changes. Necessary expenses change as wages/salaries change inasmuch wages/salaries are correlated with hours worked.

the largest decreases in medical expenditures occur for families exiting poverty while the largest increases occur for those entering poverty. Indeed, the top 1 percent of medical expenditures changes for families entering poverty is an increase of about \$100,000 while the top 1 percent of those exiting poverty is about \$100,000. While it is common to hear about catastrophic medical issues leading to poverty, these distributions show that only about 3 percent of families entering poverty experience an increase in medical expenditures of over \$20,000 and only 7 percent experience an increase of over \$10,000.

The main takeaway from this analysis is that resources do not uniformly increase for families entering poverty or decrease for families exiting poverty. Resources change concurrently but, when summed to total resources, almost all families entering poverty see resources decrease and almost all families exiting poverty see resources increase. Given all these changes, it is necessary to develop a method for determining which of these components are the most influential in poverty transitions.

2.6.2 Importance of resource component changes on poverty transitions

With so many moving parts to the SPM, it is difficult to ascertain which components are in fact the most relevant in predicting poverty transitions. From the averages and overall distributions of changes, one would naturally suspect that the most relevant factors would be those that have the largest changes in total magnitude (i.e. income from wages/salaries and other cash income). The full analysis requires more structure. This section proposes a simple method for formalizing the relative importance of resource changes to poverty transitions. The intuition behind the approach developed here is captured by the questions: (1) For how many poverty entries is the change in a particular resource sufficient? (2) For how many poverty entries is the change expected? Sufficient change occurs when the change in a particular resource is large enough to push that family into poverty by itself. Expected changes are those where resources change in the expected direction (i.e. wages/salaries drop for families that fall into poverty while medical expenses increase for families that fall into poverty). Expected changes determine the upper bound of relevance of a resource change. Of the percent of expected changes, a proportion will be sufficient. The higher the proportion of sufficient changes, the more influential that resource is in determining poverty (i.e. a lower bound of relevance).

Consider the following example. A family of four people has total SPM resources of \$60,000 in t_1 where \$80,000 come from wages/salaries and medical expenses are \$20,000. In t_2 wages/salaries drop by \$55,000 to \$25,000 while medical expenses drop by \$20,000 to \$5,000. The total SPM resources in t_2 are \$20,000. This pushes the family into poverty since the poverty threshold is about \$24,000. In this example, the change in wages/salaries and the change in medical expenses are expected. Neither one of these resource changes are sufficient, however, since the change of a single resource alone does not push that family into poverty.

Now, consider the same family with the \$60,000 in t_1. This time, the change in wages/salaries drops by \$60,000 to \$20,000 and medical expenses still drop to \$5,000. Total SPM resources are \$15,000. In this case, the change in wages/salaries is sufficient since the family would fall into poverty regardless of the change in medical expenses (i.e. \$20,000 in wages/salaries already puts the family into poverty). The change in medical expenses is still expected because it moves in the expected direction.

This example uses only two resources whereas total SPM resources consist of six major resource categories. To formalize this analysis consider families that are not poor in t_1 but enter poverty in t_2. By definition, the following is true:

$$TotalResources_{t_1} > Threshold_{t_1} \Rightarrow \text{Not poor in } t_1,$$

$$TotalResources_{t_2} < Threshold_{t_2} \Rightarrow \text{Poor in } t_2.$$

With some algebraic manipulation, the following must be true:

$$TotalResources_{t_2} - TotalResources_{t_1} < Threshold_{t_2} - Threshold_{t_1}.$$

For all practical purposes, the right-hand side is zero. Section 2.5.2 established that, in practice, the changes in thresholds alone (i.e. family composition changes alone) have a negligible impact on poverty transitions in comparison to the change in total resources. Thus, to simplify analysis, it is defensible to study the following:

$$TotalResources_{t_2} - TotalResources_{t_1} < 0.$$

By definition, total SPM resources in any time period are defined as:

$$TotalResources_t =$$

$$\Big(WorkIncome_t + OtherIncome_t + GovSubs_t\Big)$$
$$-\Big(MedicalExp_t + Taxes_t + NecExpenses_t\Big),$$

where $WorkIncome$ is the total family cash income from wages/salaries, $OtherIncome$ is all other cash income, $GovSubs$ is the total non-cash benefits, $MedicalExp$ is the total family medical out-of-pocket expenditures, $Taxes$ are total taxes paid, and $NecExpenses$ are a family's necessary work and child care expenses. The change in total SPM resources between two time periods is simply:

$$\Delta TotalResources = TotalResources_{t_2} - TotalResources_{t_1} =$$

$$(WorkIncome_{t_2} - WorkIncome_{t_1}) + (OtherIncome_{t_2} - OtherIncome_{t_1}) + (GovSubs_{t_2} - GovSubs_{t_1})$$
$$-\Big((MedicalExp_{t_2} - MedicalExp_{t_1}) + (Taxes_{t_2} - Taxes_{t_1}) + (NecExpenses_{t_2} - NecExpenses_{t_1})\Big)$$

This can then be decomposed into specific SPM resource components and, for any given component, the equation can be arranged to isolate a particular change in one resource component. For example, income from wages/salaries can be described as follows:

$$(WorkIncome_{t_2} - WorkIncome_{t_1})$$
$$< \Big((MedicalExp_{t_2} - MedicalExp_{t_1}) + (Taxes_{t_2} - Taxes_{t_1}) + (NecExpenses_{t_2} - NecExpenses_{t_1})\Big)$$
$$-\Big((OtherIncome_{t_2} - OtherIncome_{t_1}) + (GovSubs_{t_2} - GovSubs_{t_1})\Big)$$

This relationship states that for families who enter poverty, the loss in wages/salaries must be less than the change of expenses of a family net of their other income.

Using this relationship, one could establish a counterfactual framework by setting all other changes in resources to zero. That is, if none of the other components changed, would this family still be in poverty? But, this approach does not advantage of the actual changes in resources experienced by a family. Allowing resources to change, it is possible to define sufficient and expected changes. A sufficient change is defined as a change in resources that is large enough to push a family into poverty even if other resources increase. Then it must be true for a change to be sufficient that a family would not enter poverty without the change in this resource. Then, setting $(WorkIncome_{t_1} - WorkIncome_{t_1}) = 0$, a family would not enter poverty if the following holds:

$$0 > \Big((MedicalExp_{t_2} - MedicalExp_{t_1}) + (Taxes_{t_2} - Taxes_{t_1}) + (NecExpenses_{t_2} - NecExpenses_{t_1})\Big)$$
$$-\Big((OtherIncome_{t_2} - OtherIncome_{t_1}) + (GovSubs_{t_2} - GovSubs_{t_1})\Big).$$

In the case where this relationship holds true, the change in the work income is sufficient for exiting poverty, once all other components have changed. Using this final equation, the percent of families who enter poverty for whom the change in work income is sufficient to enter poverty can be calculated. First, define $poor_{sufficient}$=1 if the equation above holds. Then, the percent is calculated as:

$$SufficientWorkIncome = \frac{\sum poor_{sufficient}}{\sum Poor - to - NonPoor}.$$

The expected condition can be constructed in a similar fashion. Since we know that the change in total resources must be negative in order for a family to fall into poverty, then the expected condition occurs when work income decreases regardless of the change in the other resources. The expected condition for families that enter poverty occurs when:

$$(WorkIncome_{t_2} - WorkIncome_{t_1}) < 0.$$

Intuitively, this percentage captures the number of families that enter poverty where the change in resources happen in the expected direction. That is, positive resources (e.g. work income) decrease for families entering poverty while negative resource (e.g. medical expenses) increase. Thus the expected conditions capture all the families that experience the expected change in resources while the sufficient condition captures the subset of these families where the change in one resource is enough to push the family into poverty.

For families exiting poverty, the algebraic exercise is analogous to that just shown but where $TotalResources_{t_2} - TotalResources_{t_1} > 0$. That is, all families that exit poverty experience a positive change in total SPM resources overall. The results for these sets of calculations are shown in Table 2.8.

Changes in income from wages/salaries are the most influential factor in poverty transitions. As Table 2.8 shows, for families entering poverty, about 54 percent experience a decrease in wages/salaries (the expected condition). Moreover, about 31 percent of families experience a sufficient change in wages/salaries. That is, of the 54 percent of families who experience a decrease, about 60 percent (31.2/53.5) of families experience a decrease in wages/salaries large enough to push them into poverty. Other cash income is another influential resource. Though 28 percent of families experience a sufficient change in other cash income, this change represents only 36 percent of all expected changes in other cash income (28/76.1). Most families entering poverty experience changes in other cash income but a smaller proportion of those changes are sufficient relative to changes in wages/salaries.

Changes in medical expenditures prove not to be as influential as changes in wages/ salaries. Only about 4 percent of families entering poverty experience a sufficient increase in medical expenditures though more than half of families (54 percent) experience an increase in medical expenditures. This proportion amounts to only 7 percent (3.8/54.0) of increases in medical expenditures that are sufficient relative to the total expected changes. Changes in taxes paid, necessary expenses and government subsidies are much less likely to be expected and rarely sufficient.

The overall changes presented in Figures 2.8 - 2.14 demonstrated an almost symmetrical distribution between families entering poverty and those exiting poverty. However, is there reason to suspect that the sufficient and expected changes would not reflect such a symmetry? As Table 2.8 shows, the symmetry holds even in this analysis. Changes in wages/salaries tend to push families out of poverty at a higher sufficient rate than all other resources. Most families also experience changes in other cash income that are expected but a lower proportion of these are sufficient changes. Finally, all other resource changes do not prove to be as influential though, of all these other resources, medical expenditures tends to have more of an impact.

A final set of analysis is conducted to ascertain the level at which changes in two resources are sufficient at the same time. Both poverty exits and entries exhibit similar patterns. About 25 percent of families exiting poverty experience the expected changes in wages/salaries and the expected changes in medical expenditures. However, less than 1% of families exiting poverty experience changes in these two resource that are both sufficient. This pattern holds for the combination of changes in wages/salaries with changes in other cash income as well as changes in other cash income with medical expenditures. This finding suggests that the movements out of poverty tend to be strongly driven by one major component change (either family wages/salaries or other cash income). Indeed, nearly 60 percent of all poverty exits and entries can be explained by the changes in these two resource components alone.

2.7 Poverty transitions and real-life events

The sufficient and expected framework can be used to describe events in addition to changes in resource categories. To highlight the potential of this framework, this section analyzes four different life-events that are commonly associated with poverty transitions. Specifically, this section looks at (1) reaching retirement age and social security receipt, (2) large increases in medical expenditures, (3) the loss of jobs and the role of social safety nets and (4) departures of family members.

2.7.1 Social security

Social Security income has lifted millions of elderly Americans out of poverty. Specifically, in 2016, it was calculated that "without Social Security benefits, 40.5 percent of elderly Americans would have incomes below the official poverty line, all else being equal" (Romig and Sherman, 2016). This estimate uses the CPS-ASEC, relies only on cross-sectional data and is conducted by subtracting the amount of social security income from total income and comparing it to the OPM threshold. Because of the cross-sectional nature of the data, the estimate cannot calculate the number of families that actually exit poverty because of Social Security income.

The sufficient framework is applied to poverty exits based on Social Security income by the age of householder.[22] Figure 2.15 presents the percent of families that exit poverty where the increase in Social Security income is sufficient conditional on the age of the householder. The figure marks three key ages: (1) 61 for early retirement benefits which means that the householder turns 62 in t_2, (2), 65 to denote full benefit age which begins at 66 and (3) 69 to denote the maximum retirement age.[23]

[22] Recall that the age of householder denotes the person on the household roster who was identified as the householder during the CPS-ASEC interviews. When the householder changes across two time periods, it is typically the remaining spouse or the next oldest person that is assigned as householder. In linking the samples, householder is assigned to the person that is in the household in both years.

[23] To expand on the Social Security retirement age:

> Social Security's full-benefit retirement age is increasing gradually because of legislation passed by Congress in 1983. Traditionally, the full benefit age was 65, and early retirement benefits were first available at age 62, with a permanent reduction to 80 percent of the full benefit amount. Currently, the full benefit age is 66 for people born in 1943-1954, and it will gradually rise to 67 for those born in 1960 or later. Early retirement benefits will continue to be available at age 62, but they will be reduced more. When the full-benefit age reaches 67, benefits taken at age 62 will be reduced to 70 percent of the full benefit and benefits first taken at age 65 will be reduced to 86.7 percent of the full benefit. There is a financial bonus for delayed retirement. An individual reaching the full-benefit age in 2015 (66 years old) receives an additional 8 percent benefit for each year he or she delays collecting benefits. If he or she delays taking benefits until age 70, the benefit will be 32 percent higher because of that delay. The maximum retirement benefit for someone who waits until age 70 to collect benefits is

The role of increases in Social Security income in pushing families out of poverty is clearly more influential as the householder reaches retirement age (as seen in Figure 2.15). Below age 45, the percent of families exiting poverty where Social Security income is sufficient is less than 10 percent and about 10 percent for ages 60 and below. This percent increases sharply between early retirement age and full benefit age, from around 20 percent to 40 percent. By the maximum retirement age, this percent is at or over 50 percent. To put this in more concrete terms, of families exiting poverty that have a householder aged 69 or over, over 50 percent of these families exit poverty solely because of the increase in Social Security.

2.7.2 Medical expenditures

It is typical to read stories of families who suffer a catastrophic medical emergency, rack up expensive medical bills, potentially lose a job due to disability and end up in poverty.[24] The goal of the subsequent analysis is to measure the percent of families that enter poverty because of medical expenses, thus giving insight into the prevalence of these catastrophic events. Rather than defining a particular cutoff for what constitutes a large increase in medical expenses, Figure 2.16 plots the percent of families that entered poverty that experience a sufficient increase in medical expenses by the level of increase in medical expenses. Only families that experienced an increase in medical expenses are included in the analysis.

Figure 2.16 shows that the level at which sufficient changes in medical expenditures occur grows as the absolute value of medical expenditures increases. Of the families entering poverty that experience an increase of $10,000 in medical expenditures, only for about 10 percent or less of families is this increase sufficient. This level increases to about 20 percent for families entering poverty that experience an increase in medical expenditures by $20,000. It is only around the $45,000 mark that the percent of families entering poverty for which the increases are sufficient increases to over 40 percent. These results are surprising since one would expect that increases in medical expenditures of over $50,000 would be large enough

$3,501 a month in 2015(Schreur and Arnold, 2015).

[24] Disentangling the cause and effect of job loss and disability is difficult and beyond the scope of this paper though the framework demonstrated in this paper has the potential to give insight into this relationship.

to push most families into poverty. The analysis here reinforces the idea shown in Table 2.8 above that large medical expenditures that push families into poverty are not as prevalent as expected. Moreover, conditional on the increases being large, this analysis suggests that catastrophic medical events that push families into poverty are not commonplace.

2.7.3 Job loss and social safety nets

It has been shown that changes in income from wages/salaries are the most influential factor pushing families into poverty. However, changes in wages/salaries can occur for a variety of reasons that are not limited to job loss (i.e. work hours are reduced or wages/salaries fall). Figure 2.17 breaks down the work schedule changes for a householder and highlights the growing influence of losing work hours for falling into poverty (see Appendix Table A.4 for a tabular representation of this figure). The sample for this figure are all those nonpoor families in t_2 where the householder worked 52-50 full-time equivalent (F.T.E.) weeks in t_1. When the householders maintains 52-50 weeks in t_2, only about 3 percent of those families fall into poverty. As the number of F.T.E. weeks fall, the percent of families in poverty increases. When householders drop down to 49-40 weeks, about 5 percent of families fall into poverty. This percent more than doubles once the householder drops to working only 13-1 weeks. Moreover, the proportion of these families that fall into poverty solely because of the change in wages/salaries increases as the weeks worked drop. For those dropping to 49-40 weeks, a little over half (56 percent) fall into poverty because of the loss in wages/salaries. For those dropping to 13-1 weeks, about 80 percent fall into poverty because of the loss in wages/salaries. When a householder becomes unemployed after having worked full-time year-round, over one-third of those families end up in poverty and nearly two-thirds of them fall into poverty because of the loss in wages/salaries. In contrast, when householders retire, about 15 percent of these families fall into poverty.

On the flip side, there are families that experience job loss but do not fall into poverty. Government subsidies are intended to provide a safety net for events such as job loss.[25] Since the sufficient framework only works for the poverty transitioners, this analysis employs a different methodology. Using a family's t_2 resources, the total government subsidies

[25] Government subsidies include SNAP, lunch subsidies, WIC, housing subsidies and energy subsidies *as well as* EITC.

received by a family are subtracted from their total resources. If that family's total resources net of government subsidies are below their poverty line then the government subsidies are considered to be keeping the family out of poverty. This analysis is also conducted for unemployment compensation.

Figure 2.18 shows that unemployment compensation helps mitigate job loss resulting in poverty for more families than other government subsidies (see Appendix Table A.5 for a tabular representation of this figure). As the number of F.T.E. weeks worked in t_2 falls, the more government subsidies and unemployment compensation help keep families out of poverty. As expected, this role is most pronounced for families facing unemployment rather than just a reduction of weeks worked. That is, of all nonpoor families in t_1 whose householder goes from working full-time year-round to unemployed, about 8 percent of them do not fall into poverty because of government subsidies while 12 percent of them avoid poverty because of unemployment compensation. This result implies that there are another 80 percent of families not falling into poverty because of some other mechanism.

A couple explanations could be explored. First, families may be avoiding poverty due to sources that are excluded from government subsidies, namely workers' compensation, Supplemental Security Income, public assistance (welfare), and disability benefits (these are lumped into other cash income). Second, other adults in the family may increase their work hours to make up for this loss of income. In sum, the role of government subsidies and unemployment help about a fifth of families losing employment mitigate poverty. Alternate explanations should be explored for the other four-fifths of families losing employment but not entering poverty.

2.7.4 Family departures

As shown in Section 2.5, the role of family composition change alone, net of the resource changes associated with them, plays a small role in explaining poverty transitions. This section uses the sufficient framework to highlight how the loss of resources associated with family departures affects poverty transitions. To simplify analysis, Figure 2.19 focuses on the departures of householder and spouses broken down by sex for families that enter poverty in t_2 (see Appendix Table A.6 for a tabular representation).[26] The figure does

[26] Departures are defined as noted in Section 2.5. Specifically, adults depart when they are present in the family in t_1 but not in t_2. Moreover, divorce and death requires a change in the marital status of the

not distinguish between who is listed as the householder or the spouse but rather focuses on households where both a householder and spouse exist as opposed to families consisting of just one individual or a single parent. The departure of males leads to a higher proportion of families entering poverty than when females depart. Regardless of the reason for the departure, when a male adult leaves the household, about 20 percent of those families will enter poverty. For female departures, only about 10 percent of families enter poverty while 5 percent enter poverty when a female passes away. These results likely reflect the higher propensity of males being the primary earner in U.S. families but the explanation should be further explored.

Finally, using the sufficient framework, it is noted that the change in family wages/ salaries alone pushes a large proportion of families into poverty. When a male departs because of divorce, 55 percent of these families enter poverty because of the loss in wages/ salaries associated with his departure. This percent is nearly the same for families where the female departs because of divorce, though female divorces in general are associated with fewer poverty entries. For deaths, the loss in wages/salaries is only responsible for about 10 percent of poverty entries. It is likely that deaths tend to occur at an older age when income is likely to be in the form of retirement rather than wages/salaries. In sum, it is clear that the role of adult departures in poverty entries is in fact an important one through its loss of resources.

2.8 Discussion

2.8.1 Data limitations - imputations

There are two major concerns in using linked CPS-ASECs that require discussion. The first issue, and in no particular order of importance, is that of imputations. Imputations in the CPS-ASEC have long been studied and have been pointed out to affect the results regarding income and earnings (Lillard, Smith and Welch, 1986). The prevalence of imputations in the CPS-ASEC has increased over time. In relation to poverty, it has been shown that the observations with imputed income have the lowest poverty rates while those without imputation have higher poverty rates. That is, those who respond to the CPS-ASEC tend

remaining spouse (i.e. from married to divorced/widowed).

to have lower incomes than those who do not respond. Indeed, between 1981 and 2007, the imputed sample exhibited a poverty rate that was about 3 percentage points lower than those without imputations (Turek et al., 2009). While the role of imputations has been studied in a cross-sectional setting, no study has looked at how these imputations affect poverty transitions.

The approach employed here is to first look at how imputations relate to the overall poverty churn and then to look at how the main patterns found in this paper change, if at all, because of these imputations. In order to clearly disentangle the different combinations of imputations, families are grouped into four mutually exclusive and exhaustive types of imputations : (1) no imputes in either t_1 or t_2, (2) imputes in both t_1 and t_2, (3) imputes in t_1 only and (4) imputes in t_2 only. Figure 2.20 decomposes the SPM churn into these four groups. Of the total poverty exits and entries (including imputes), which each amount to about 8 percent of the total population, about 30 percent of those are non-imputes. Of the 6 percent of those in poverty in both time periods, about 40 percent are non-imputes. Recalculating the churn based solely on the total population that does not have imputations (these percentages are presented in parentheses in Figure 2.20) shows that fewer families transition in and out of poverty (6.9 percent and 7.2 percent) than those remaining in poverty (7.6 percent). This result would seemingly reverse the finding of the total population graphed in Figure 2.20, where more families transition in and out of poverty than remain in poverty. It also begs the question of whether the higher churn rate of the SPM is being driven by the imputed values. That is, are these poverty transitions artifacts of imputations?

Imputations in both time periods are more prevalent in the poverty transitioners than those remaining in poverty. Indeed, the families with imputations in both time periods comprise the largest proportion of the churners (36.8 percent and 35.2 percent) while the most common group for families remaining in poverty are no imputes (41.1 percent). Recalculating the churn for this group of both imputes shows the percent of poverty transitions is about three to four percentage points higher than those remaining in poverty (9 percent and 8.5 percent versus 4.7 percent).

A similar pattern is noticed with the imputations that occur in one time period alone. Key to understanding the pattern are two facts. First, as mentioned before, imputations lead to, on average, a higher income than non-imputations. The immediate implication is that a family with imputed income in t_1 and no imputes in t_2 will more likely move into poverty, on average. Second, imputations are more likely to occur in t_2. This fact occurs simply because non-response increases as the number of CPS interviews increases. Recall the rotation pattern of the CPS. By the time a family is in their t_2 period, they will be potentially responding for a fifth or even eighth time. Thus, respondents tend to refuse to respond more in t_2.

The proper question to ask is: Are imputations necessarily wrong? Research has shown that, in a regression setting where the coefficient of interest is a variable not included in the imputation process, the results of the regression may be biased (see Hirsch and Schumacher (2004)). However, since the analysis in this paper is not regression-based, dealing with imputations is not as easy as dropping all imputed observations. The real need is to assess whether the imputed values are systematically lower or higher than their "true" values. To this end, the relevant research uses administrative tax data to assess the accuracy of imputations. Using data from W-2 forms, Hokayem, Bollinger and Ziliak (2014) replace all imputed earnings values with administrative data and calculate the overall poverty rate. They find that replacing the imputed data with administrative data lowers the OPM rate by about 1 percentage point. The implication is that the imputed data leads to higher incomes than their "true" values but it does not seem to lower poverty enough to negate the patterns seen here. In a similar paper, Bollinger et al. (2014) focus on the earnings of individuals and find that individuals who respond in one year but not the next (i.e. imputed in the second year) are likely to be those whose wages/salaries drop significantly. This result is in line with the patterns noticed above where imputations tend to be higher than the real values, implying that the results in this paper are understated.[27]

For completeness, all the analysis in this paper was rerun for all these imputation groups. The overall conclusions remain true. The main areas where imputations seem to have an effect are on the overall magnitudes of changes in resources and the level of

[27] There is not much known about the accuracy of public benefit reporting in the CPS-ASEC. In order to assess the reliability of these imputations, administrative data needs to be linked to the CPS responses. This requires agreements with specific states and thus is less common. One current project has linked the administrative data from Maryland SNAP to the CPS-ASEC and found that of all SNAP recipients found in the CPS-ASEC for Maryland, about 60 percent report not receiving SNAP in the past calendar year. Of those that do not misreport, the underreporting is non-negligible.

influence different changes have in the sufficient and expected framework. The distribution of changes are slightly changed but the overall patterns remain the same. For the sufficient and expected framework, the influence of resources are more pronounced when imputes occur in one time period only. For those with imputes in t_1, those exiting poverty have a larger percent of sufficient and expected changes. This occurs because the income in t_1 is higher and thus families are closer to their poverty threshold on average. This then results in more change being sufficient. For those with imputes in t_2, those entering poverty have a larger percent of sufficient changes. In sum, imputations do change the magnitude of change but the overall patterns remain the same.

2.8.2 Data limitations - attrition

The second concern in using the linked CPS is that of attrition. As demonstrated by Hokayem and Heggeness (2014) and Bollinger et al. (2014), the linked-CPS sample has slightly different characteristics than the full ASEC sample. They find that the poverty rate is about 3 percentage points lower in the linked sample than in the full ASEC (10.5 percent v 13.7 percent) and that the demographic characteristics tend to be different in both samples. Table 2.9 shows that this is the case with the linked sample used here. Of most importance is the fact that incomes tend to be lower in the full CPS-ASEC than in the linked sample. This is to be expected since attrition in the CPS-ASEC mainly occurs from families moving out of a household and families with less income tend to be less stable.

The implication this has for the present analysis is unclear. Assume that the families who are not linked are all poor in t_1. Do these families move from their home because they can no longer afford it and must move into public housing (i.e. remain poor)? Or do they move because they received a job offer in a different locale and thus move out of poverty? The missing information is what percent of these families continue to be in poverty in t_2 and what percent exit poverty. Moreover, it remains unclear what their overall distribution of resources will be in t_2. In order to answer this question, research must be conducted on the t_2 characteristics of those families that fall out of the CPS-ASEC and, to date, this research has not been conducted. A possible research design that could be used would be to leverage administrative data. One could link administrative tax data to get a complete earnings history of a person. By doing so, one would have data on t_2 earnings and even beyond.

To give some insight into the poverty rate of movers, Figure 2.21 and Figure 2.22 plot the poverty rates of those families in the sample for only one time period. Each figure shows the cross-sectional poverty rate for t_2 which uses the entire non-linked CPS for a given year. This sample includes families who are present in just t_1 and those present in only t_2. The linked-CPS poverty rate is the poverty rate using just the families who are present in both time periods. Two other rates are presented: 1) the poverty rate where the families who are present only in t_1 are added to the linked families, and 2) the poverty rate where the families who are present only in t_2 are added to the linked families. Figure 2.21 presents the results for the OPM (1996-2016 CPS-ASEC) while Figure 2.22 presents the results using the SPM (2010-2016 CPS-ASEC). As can be seen, the poverty rates that add in the moving families are extremely close to the cross-sectional poverty rates (in the case of the OPM, are almost exactly the same as the cross-sectional). Moreover, the poverty rates of these moving families are nearly identical to each other. These results suggest that the overall results are likely unaffected since the poor families moving out tend to be replaced in the sample by an equal amount of poor families.

2.9 Conclusion and future work

The analysis presented in this paper has important policy implications. First, it is clear that the definition of poverty matters in terms of understanding the underlying factors that push families in and out of poverty. The SPM is a robust definition and should be leveraged in more research settings. Of most importance is the finding that more families move in and out of poverty under the SPM than under the OPM. This finding should not be taken as just an artifact of definition but rather a better reflection of what is actually happening to low-income families in the U.S. A high proportion of families in the U.S. are economically vulnerable and the SPM can better capture this fact. Second, it is shown that access to jobs is the most important vehicle through which families avoid poverty. This finding supports the main result found in the poverty transition literature. Job training programs or even job relocation programs are an important anti-poverty mechanism. Moreover, higher levels of resources are associated with a lower probability of entering poverty. Thus, whether it be through educational programs or higher mandated wages/salaries, it is worth promoting better wages/salaries for families. These implications are not novel but are important to reiterate in light of these results. Third, Social Security receipt was shown to be an effective method of escaping poverty. Nearly 50 percent of the families with householder turning 70 or more exited poverty because of Social Security.

From a research standpoint, the potential of this framework is vast. This paper focused mainly on describing the overall patterns in resource changes. The next step would be to apply this framework with more focus on causality, policy and specific life-events. For example, a simple analysis could be conducted to look at the effect of the Affordable Care Act on poverty transitions or medical expenditures. Since the CPS-ASEC includes data on health insurance, it would be straightforward to look at the uptake in insurance after the implementation of the ACA and the subsequent medical expenditures of a family.

Another vein of research are the role of social programs. The research presented here gives initial evidence of how these programs can mitigate poverty transitions. These programs can be studied in more depth. The other cash income category includes sources such as disability and public assistance. Unraveling these sources would be an insightful project. Moreover, in the vein of family composition changes, how many families mitigate poverty by having other family members move in? All these questions can be studied through the framework established in this paper.

A final more descriptive sort of research would be to understand more fully the characteristics of those families who remain in poverty. In this paper it was highlighted that African American and Hispanics are disproportionately represented in the poor-to-poor population, even more so than those transitioning into poverty. What are the distinguishing characteristics of this always poor population? Understanding the characteristics of the population remaining in poverty can reveal important patterns that can inform policy.

From a data standpoint, there are various research topics that would improve the reliability and representativeness of this study. First, as mentioned in the data limitation section, it would be worthwhile to use administrative data to see how imputations affect the distribution of resource changes, as opposed to the overall distribution of resource in one time period. Second, administrative data could be leveraged to understand what happens to movers. Indeed, research on movers would be valuable for using the linked-CPS in general, even beyond poverty transitions. Third, the data presented here does not use population weights. This is particularly important for subgroup analysis and would also be a more accurate way of making nationally-representative claims. The typical method for dealing with attrition is inverse probability weights readjusted in order to account for non-random attrition. Finally, there are a couple of sources of data that would be worth looking analyzing. As mentioned earlier, the SPM takes root in the 1990s through the National Academy of Sciences (NAS). Building up to the SPM, the Census Bureau published the NAS experimental poverty files that include many of the components used in the SPM. Similarly, the Center on Poverty & Social Policy at Columbia University has recently released historical SPM datasets that impute the SPM back to 1967. These data sources could be included in this framework and would allow for a longer time frame to study these poverty transitions.

The goal of this paper was to put poverty statistics in more tangible terms. To be able to say something like: "When a male adult leaves the household, about 20 percent of those families will enter poverty but, for female departures, only about 10 percent of families enter poverty." In order to get to this point, this paper used the linked-CPS to understand poverty transitions and established the necessary building blocks and framework for analyzing poverty transitions with this dataset. By doing so, various insights were evident. First, the definition of poverty matters. The SPM has a higher churn rate than the OPM, implying that the SPM capture more families falling into poverty. Second, family composition changes are an important factor in explaining poverty transitions inasmuch as these departures result in lost resources for a family. Third, the paper documents how resources change by component for families entering and exiting poverty. To begin to show how one may use this framework to its full potential, a simple methodology is created to isolate the effects of SPM resource components. From this analysis, it is clear that employment is the most important factor for SPM transitions and that most transitions can be explained by the change in one resource alone. And lastly, above all, the framework is leveraged to understand poverty statistics in more human terms.

Table 2.1: OPM thresholds for 2016 by size of family and number of related children under 18 years

	Related children under 18 years								
Size of family unit	None	One	Two	Three	Four	Five	Six	Seven	Eight or more
One person:									
<Age 65	12,486								
Aged 65+	11,511								
Two people:									
Householder <Age 65	16,072	16,543							
Householder Aged 65+	14,507	16,480							
Three people	18,774	19,318	19,337						
Four people	24,755	25,160	24,339	24,424					
Five people	29,854	30,288	29,360	28,643	28,205				
Six people	34,337	34,473	33,763	33,082	32,070	31,470			
Seven people	39,509	39,756	38,905	38,313	37,208	35,920	34,507		
Eight people	44,188	44,578	43,776	43,072	42,075	40,809	39,491	39,156	
Nine people or more	53,155	53,413	52,702	52,106	51,127	49,779	48,561	48,259	46,400

Source: U.S. Census Bureau.

Table 2.2: Breakdown of SPM resource components

Total Cash Income[a]	
1. Wages and Salary	11. Disability benefits
Other Cash Income	12. Pension or retirement income
2. Business	13. Interest
3. Farm	14. Dividends
4. Unemployment compensation	15. Rents, royalties, and estates and trusts
5. Workers' compensation	16. Educational assistance
6. Social security	17. Alimony
7. Supplemental Security Income (SSI)	18. Child support
8. Public assistance (welfare)	19. Financial assistance from outside of the household
9. Veterans' payments	20. All other income
10. Survivor benefits	
	Plus:
Government Subsidies	
21. Supplemental Nutritional Assistance Program (SNAP)	
22. National School Lunch Program	
23. Supplementary Nutrition Program for Women Infants and Children (WIC)	
24. Housing subsidies	
25. Low Income Home Energy Assistance Program (LIHEAP)	
	Minus:
Taxes Paid[b]	
26. Federal income tax (-)	32. Child Tax Credit (CTC) (+)
27. State income tax (-)	31. Earned Income Tax Credit (+)
28. Annual property taxes (-)	33. Additional Child Tax Credit (+)
29. Federal Insurance Contributions Act (FICA) (-)	
30. Federal retirement payroll deduction (-)	
	Minus:
Necessary Expenses	
34. Expenses related to work	36. Child support paid
35. Child care expenses	
	Minus:
Medical Expenditures	
37. Medical Out-of-Pocket (MOOP) expenses and Medicare B subsidy	

Source: Proctor, Semega and Kollar, 2016. Notes: (a) Total cash income encompasses all income components used for OPM total family income. (b) Tax components are subtracted (-) while credits are added (+) such that total taxes are the difference between taxes (-) and credits (+).

Table 2.3: Identifying family composition changes using the CPS-ASEC

Panel A: Household Roster - Arrival of Newborn

Relationship		Correct Family Type & Threshold		Erroneous Family Type & Threshold	
t_1	t_2	t_1	t_2	t_1	t_2
Householder	Householder	\$ 19,078	\$ 24,339	\$ 19,078	\$ 24,339
Spouse	Spouse	\$ 19,078	\$ 24,339	\$ 19,078	\$ 24,339
Child	Child	\$ 19,078	\$ 24,339	\$ 19,078	\$ 24,339
—	Newborn	—	\$ 24,339	—	—
Family Type		3 people 1 child ↓	4 people 2 children ↓	3 people 1 child ↓	3 people 1 child ↓
Corresponding Poverty Threshold[a]		\$ 19,078	\$ 24,339	\$ 19,078	\$ 19,318
Total Family Income		\$ 19,000	\$ 19,500	\$ 19,000	\$ 19,500
Poverty Status		Poor	Poor	Poor	NonPoor

Panel B: Household Roster - Death of a family member and arrival of another family member

Relationship		Correct Family Type & Threshold		Erroneous Family Type & Threshold	
t_1	t_2	t_1	t_2	t_1	t_2
Elder h.h.	—	$ 14,326	—	—	—
Elder spouse	Widowed h.h.	$ 14,326	$ 14,507	$ 14,326	$ 14,507
—	Adult Child	—	$ 14,507	—	—
Family Type		2 people (65+) 0 children ↓	2 people (65+) 0 children ↓	1 person (65+) 0 children ↓	1 person (65+) 0 children ↓
Corresponding Poverty Threshold		$ 14,326	$ 14,507	$ 11,367	$ 11,511
Total Family Income		$ 12,000	$ 12,500	$ 12,000	$ 12,500
Poverty Status		Poor	Poor	NonPoor	NonPoor

Note: (a) Poverty threshold values correspond to calendar year 2015 for t_1 and 2016 for t_2.

Table 2.4: Descriptive statistics of families by SPM poverty transitions

	Full Sample	Never Poor	NonPoor to Poor	Poor to NonPoor	Always Poor
Number of People - T1	2.24	2.29	2.08	2.07	2.06
	(0.004)	(0.004)	(0.014)	(0.014)	(0.017)
Number of People - T2	2.22	2.27	2.03	2.08	2.07
	(0.004)	(0.004)	(0.013)	(0.014)	(0.017)
Selected Demographics of Householder					
Age of Householder	54.61	54.56	56.30	55.13	52.18
	(0.047)	(0.051)	(0.177)	(0.185)	(0.211)
% White - Non. Hisp.	76.77	80.08	68.66	68.47	56.89
	(0.118)	(0.127)	(0.457)	(0.461)	(0.570)
% Afr. Amer. - Non. Hisp.	8.25	6.79	12.69	12.31	15.39
	(0.077)	(0.080)	(0.328)	(0.326)	(0.415)
% Amer. Ind. - Non. Hisp.	0.45	0.40	0.54	0.64	0.77
	(0.019)	(0.020)	(0.072)	(0.079)	(0.100)
% Asian - Non. Hisp.	3.35	3.24	3.21	3.67	4.55
	(0.050)	(0.056)	(0.174)	(0.187)	(0.240)
% Other - Non. Hisp.	1.11	1.07	1.28	1.17	1.41
	(0.029)	(0.033)	(0.111)	(0.107)	(0.136)
% Hispanic	10.05	8.42	13.62	13.74	21.00
	(0.084)	(0.088)	(0.338)	(0.341)	(0.469)
% Non-Citizen	6.54	5.12	9.33	10.39	15.56
	(0.069)	(0.070)	(0.287)	(0.303)	(0.417)
% Naturalized Citizen	7.18	6.88	7.57	8.33	9.00
	(0.072)	(0.080)	(0.261)	(0.274)	(0.329)
Observations	127,146	99,116	10,308	10,170	7,552

Source: CPS-ASEC 2010-2016. Note: Standard errors in parentheses.

Table 2.5: Family composition changes by SPM poverty transitions

	Full Sample	Never Poor	NonPoor to Poor	Poor to NonPoor	Always Poor
Summary of Changes[a]					
% Departure of Any Family Member	7.04 (0.072)	6.85 (0.080)	9.67 (0.291)	6.46 (0.244)	6.54 (0.284)
% Arrival of Any Family Member	5.35 (0.063)	5.15 (0.070)	5.83 (0.231)	6.36 (0.242)	5.84 (0.270)
% Both Departure and Arrival	0.80 (0.025)	0.73 (0.027)	1.15 (0.105)	1.06 (0.101)	0.90 (0.109)
% No Change in Family Composition	86.82 (0.095)	87.28 (0.106)	83.36 (0.367)	86.12 (0.343)	86.72 (0.390)
Arrivals[a]					
% Newborn	1.86 (0.038)	1.87 (0.043)	1.83 (0.132)	1.47 (0.119)	2.33 (0.173)
% Arrival Child (2-6)	0.47 (0.019)	0.41 (0.020)	0.72 (0.083)	0.56 (0.074)	0.72 (0.097)
% Arrival Child (7-17)	0.75 (0.024)	0.66 (0.026)	1.08 (0.102)	0.97 (0.097)	1.23 (0.127)
% Arrival Adult (18-64)	3.64 (0.053)	3.43 (0.058)	4.17 (0.197)	5.19 (0.220)	3.51 (0.212)
% Arrival Elder (65+)	0.36 (0.017)	0.33 (0.018)	0.41 (0.063)	0.59 (0.076)	0.42 (0.074)
% Marriage	0.90 (0.027)	0.88 (0.030)	0.82 (0.089)	1.37 (0.115)	0.71 (0.097)
Departures[a]					
% Departure Child (0-6)	0.70 (0.023)	0.62 (0.025)	0.87 (0.092)	1.16 (0.106)	0.90 (0.109)
% Departure Child (7-17)	1.04 (0.028)	0.99 (0.032)	1.02 (0.099)	1.25 (0.110)	1.37 (0.134)
% Departure Adult (18-64)	6.12 (0.067)	6.00 (0.075)	7.80 (0.264)	5.94 (0.234)	5.54 (0.263)
% Departure Elder (65+)	1.12 (0.030)	1.02 (0.032)	2.44 (0.152)	1.00 (0.099)	0.86 (0.106)

% Departure Spouse	1.06	0.96	2.25	1.03	0.73
	(0.029)	(0.031)	(0.146)	(0.100)	(0.098)
% Departure Head	1.37	1.17	3.04	1.31	1.68
	(0.033)	(0.034)	(0.169)	(0.113)	(0.148)
% Divorce/Separation	0.75	0.67	1.73	0.66	0.62
	(0.024)	(0.026)	(0.128)	(0.080)	(0.090)
% Widowed	0.71	0.66	1.50	0.56	0.43
	(0.024)	(0.026)	(0.120)	(0.074)	(0.075)
Observations	127,146	99,116	10,308	10,170	7,552

Source: CPS-ASEC 2010-2016. Notes: Standard errors in parentheses. (a) The summary of changes are mutually exclusive but the breakdown of arrivals and departures are not mutually exclusive. Changes in family composition can be concurrent.

Table 2.6: Counterfactual poverty rates for family composition and resource changes - SPM

	Poor t_1	Poor t_2	Family Change	Resource Change
	Resource t_1 / Threshold t_1	Resource t_2 / Threshold t_2	Resource t_1 / Threshold t_2	Resource t_2 / Threshold t_1
SPM Counterfactuals				
Percent Poor	13.94	14.05	14.07	14.74***
	(0.097)	(0.097)	(0.098)	(0.099)
Observations		127,416		

Source: CPS-ASEC 2010-2016. Notes: Standard errors in parentheses. Difference between Poor t_2 (14.05%) and Resource Change (14.74%) is significant at *** p<0.01.

Table 2.7: Average change in SPM resource components by poverty transition type

	All		NonPoor Poor		Poor NonPoor	
	Mean t_1	$\overline{\Delta}$ $t_2 - t_1$	Mean t_1	$\overline{\Delta}$ $t_2 - t_1$	Mean t_1	$\overline{\Delta}$ $t_2 - t_1$
SPM Total Resources	54,206	-509**	38,543	-29,291***	9,604	27,775***
	(55,976)		(36,098)		(14,940)	
Total Cash Income	74,505	78	48,523	-33,506***	15,427	32,007***
	(84,684)		(52,888)		(13,863)	
Wages/Salaries	53,641	-288	25,673	-18,444***	7,666	17,550***
	(80,362)		(45,793)		(13,222)	
Other Cash Income	20,864	366**	22,850	-15,062***	7,761	14,457***
	(36,919)		(35,858)		(9,270)	
Taxes Paid	14,159	488***	6,264	-5,884***	281	5,972***
	(28,462)		(16,457)		(2,973)	
Necessary Expenses	2,075	-26***	1,329	-306***	1,092	261***
	(2,770)		(1,875)		(2,788)	
Government Subsidies	454	-12**	858	143***	1,038	-163***
	(1,762)		(2,465)		(2,589)	
Medical Expenditures	4,538	94***	3,269	2,093***	5,505	-2,183***
	(6,969)		(4,293)		(15,061)	
Observations	127,146		10,308		10,170	

Source: CPS-ASEC 2010-2016. All values are adjusted to reflect 2015 dollars. Adjustment made using annual average Consumer Price Index Research Series (CPI-U-RS). Standard deviations in parentheses. Difference ($t_2 - t_1$) significant at * p<0.1, ** p<0.05, *** p<0.01.

Table 2.8: Measuring the relative importance of SPM components to predict SPM poverty

	Poor to NonPoor		NonPoor to Poor	
	Percent that Exit Poverty		Percent that Enter Poverty	
	Sufficient	Expected	Sufficient	Expected
Wages/Salaries	31.4	51.9	31.2	53.5
	(0.005)	(0.005)	(0.005)	(0.005)
Other Cash Income	27.7	76.1	28.0	76.1
	(0.004)	(0.004)	(0.004)	(0.004)
Taxes Paid	0.4	12.2	0.1	13.2
	(0.001)	(0.003)	(0.001)	(0.003)
Necessary Expenses	0.3	18.2	0.4	22.3
	(0.001)	(0.004)	(0.001)	(0.004)
Government Subsidies	1.3	20.0	1.1	19.6
	(0.001)	(0.004)	(0.001)	(0.004)
Medical Expenditures	4.2	52.6	3.8	54.0
	(0.002)	(0.005)	(0.002)	(0.005)
Both components are sufficient				
Both Wages and Medical	0.14	25.4	0.26	27.7
Both Wages and Other Cash Income	0.17	32.3	0.12	33.0
Both Other Cash Income and Medical	0.14	38.3	0.28	39.3
	10,028		10,204	

Source: CPS-ASEC 2010-2016. Notes: One percent of observations from each sample has been dropped. For the NonPoor to Poor, 104 observations that had increases in total resources were dropped. For the Poor to NonPoor, 142 observations that had decreases in total resources were dropped. Standard errors in parentheses.

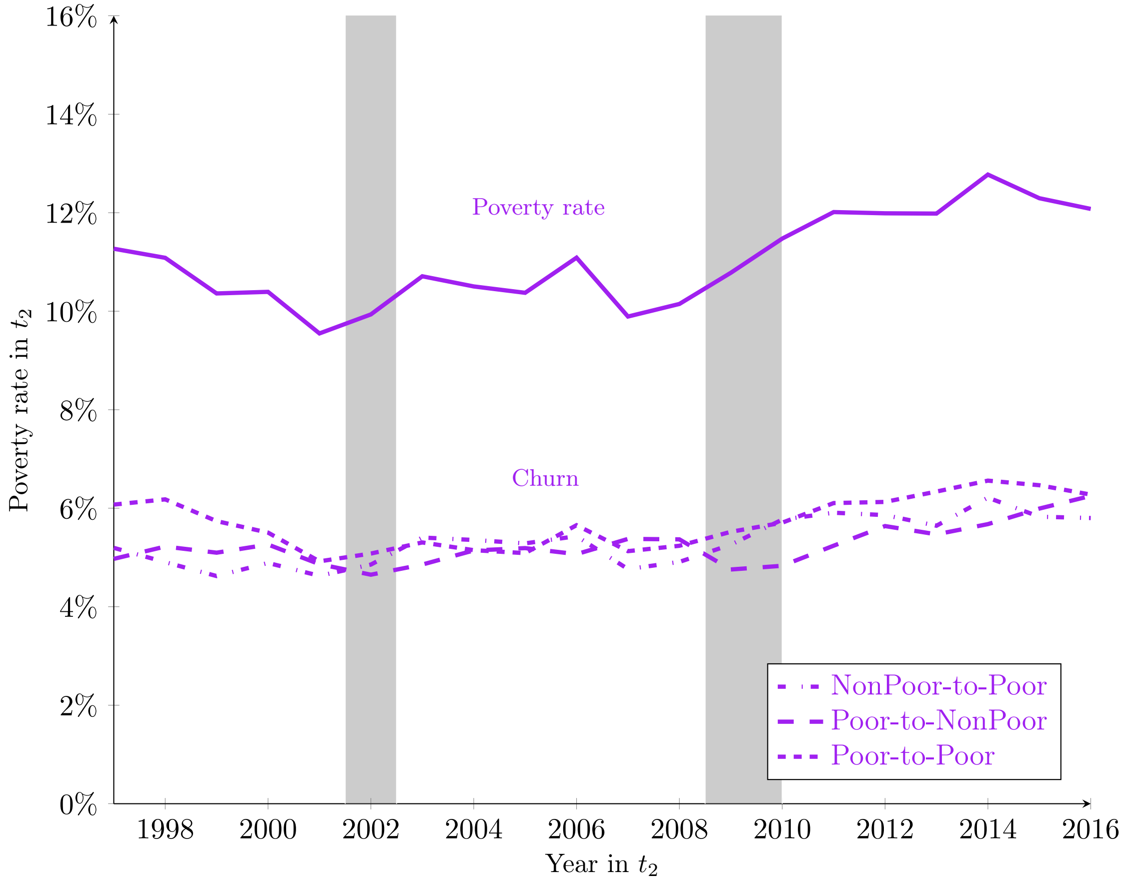
Table 2.9: Attrition in the linked CPS-ASECs

	Full Sample	Linked Sample
Number of People	2.38	2.24
	(0.002)	(0.004)
Selected Demographics of Householder		
Age of Householder	43.76	54.61
	(0.031)	(0.047)
% White - Non. Hisp.	67.17	76.77
	(0.064)	(0.118)
% Afr. Amer. - Non. Hisp.	12.20	8.25
	(0.044)	(0.077)
% Amer. Ind. - Non. Hisp.	0.65	0.45
	(0.011)	(0.019)
% Asian - Non. Hisp.	4.92	3.35
	(0.029)	(0.05)
% Other - Non. Hisp.	1.59	1.11
	(0.017)	(0.029)
% Hispanic	13.47	10.05
	(0.046)	(0.084)
% Non-Citizen	6.33	6.54
	(0.033)	(0.069)
% Naturalized Citizen	5.74	7.18
	(0.032)	(0.072)
Resource Components		
SPM Total Resources	51,966	54,206
	(55,246)	(55,976)
Total Cash Income	71,654	74,505
	(84,895)	(84,684)
Wages/Salaries	54,003	53,641
	(79,905)	(80,362)
Other Cash Income	17,651	20,864
	(35,283)	(36,919)
Taxes Paid	13,912	14,159

	(29,180)	(28,462)
Necessary Expenses	2,221	2,075
	(2,941)	(2,770)
Government Subsidies	606	454
	(2,052)	(1,762)
Medical Expenditures	4,174	4,538
	(6,754)	(6,969)
Observations	541,850	127,146

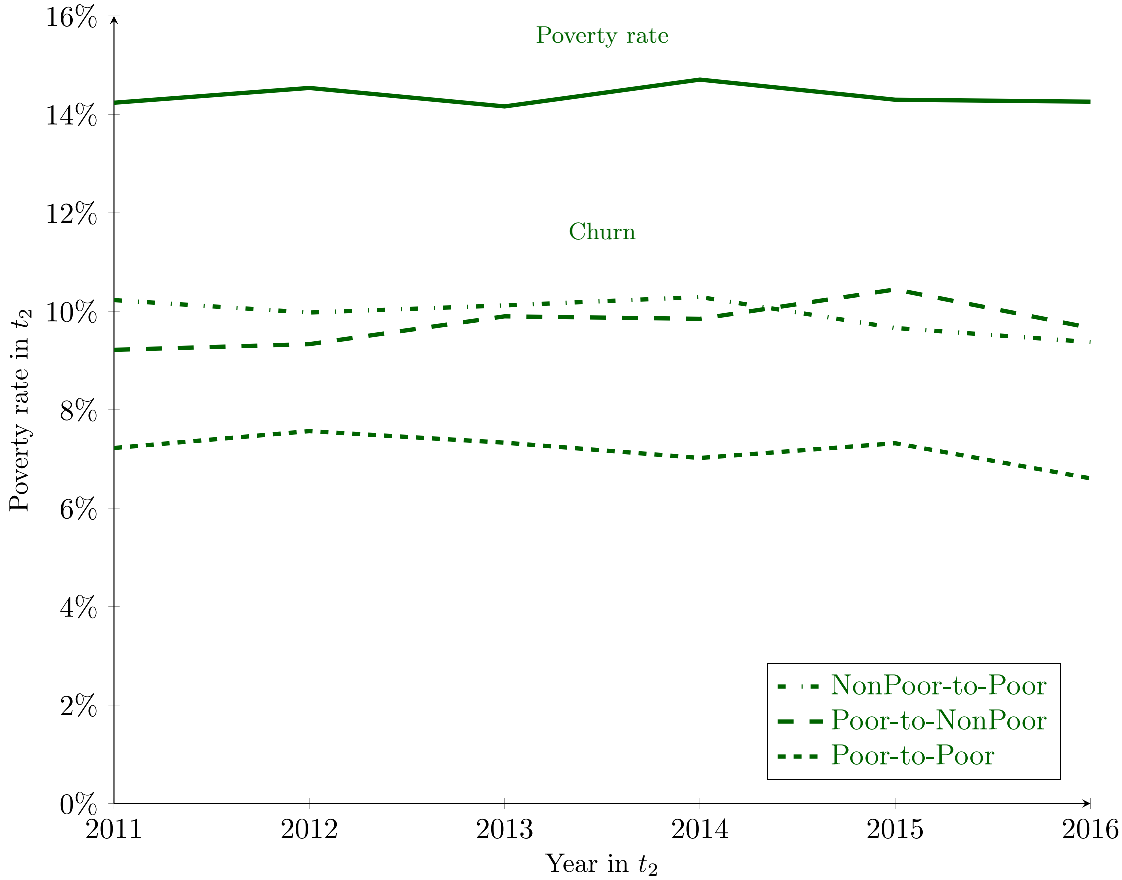
Source: CPS-ASEC 2010-2016. Notes: Standard errors in parentheses. Standard deviations reported for Resource Components.

Figure 2.1: Poverty churn as percent of total U.S. population, OPM



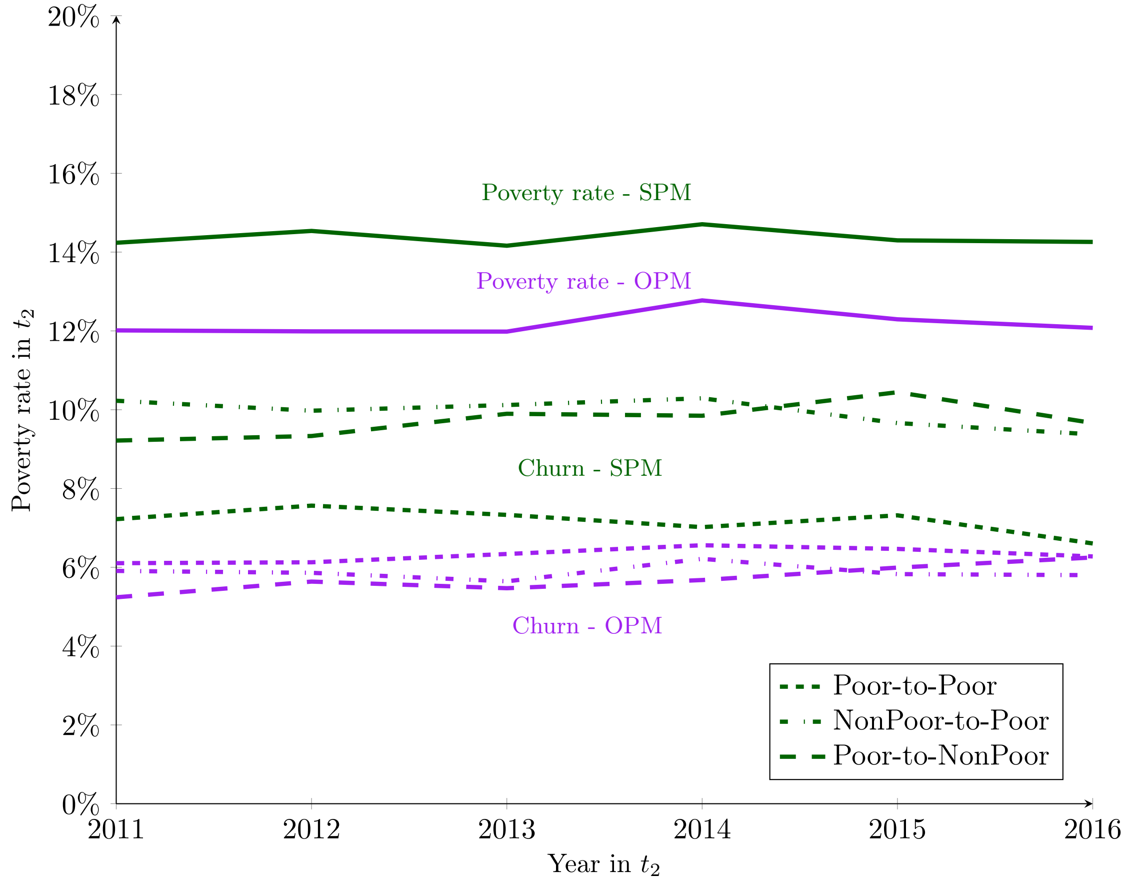
Source: CPS-ASEC 1996-2016. Notes: Shaded areas denote recession periods. Because the sample consists of two-year panels, the poverty rate uses t_2 observations such that the poverty rate is the sum of those entering poverty (NonPoor to Poor) and those staying in poverty (Poor to Poor).

Figure 2.2: Poverty churn as percent of total U.S. population, SPM



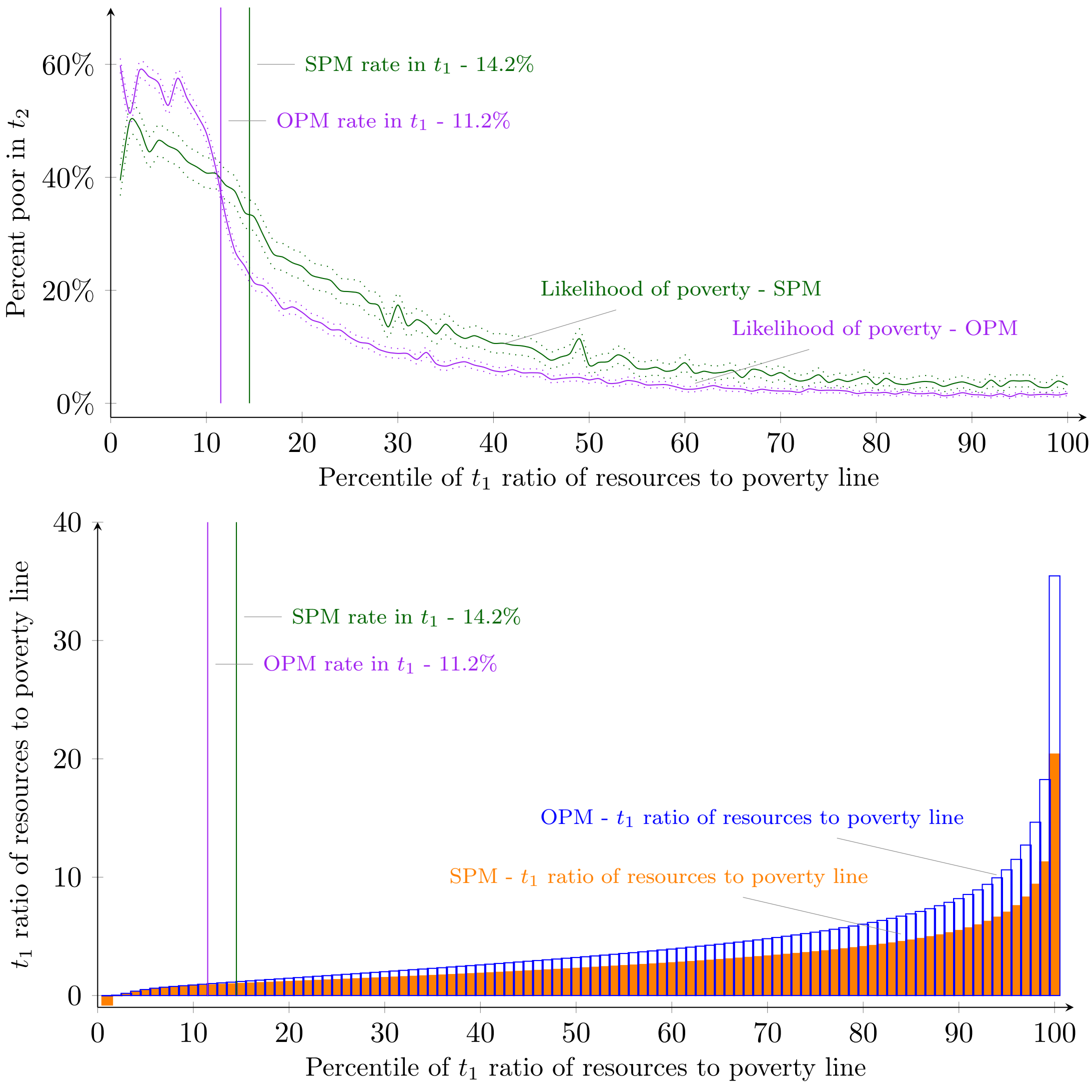
Source: CPS-ASEC 2010-2016. Notes: Because the sample consists of two-year panels, the poverty rate uses t_2 observations such that the poverty rate is the sum of those entering poverty (NonPoor to Poor) and those staying in poverty (Poor to Poor).

Figure 2.3: Poverty churn as percent of total U.S. population, OPM v. SPM



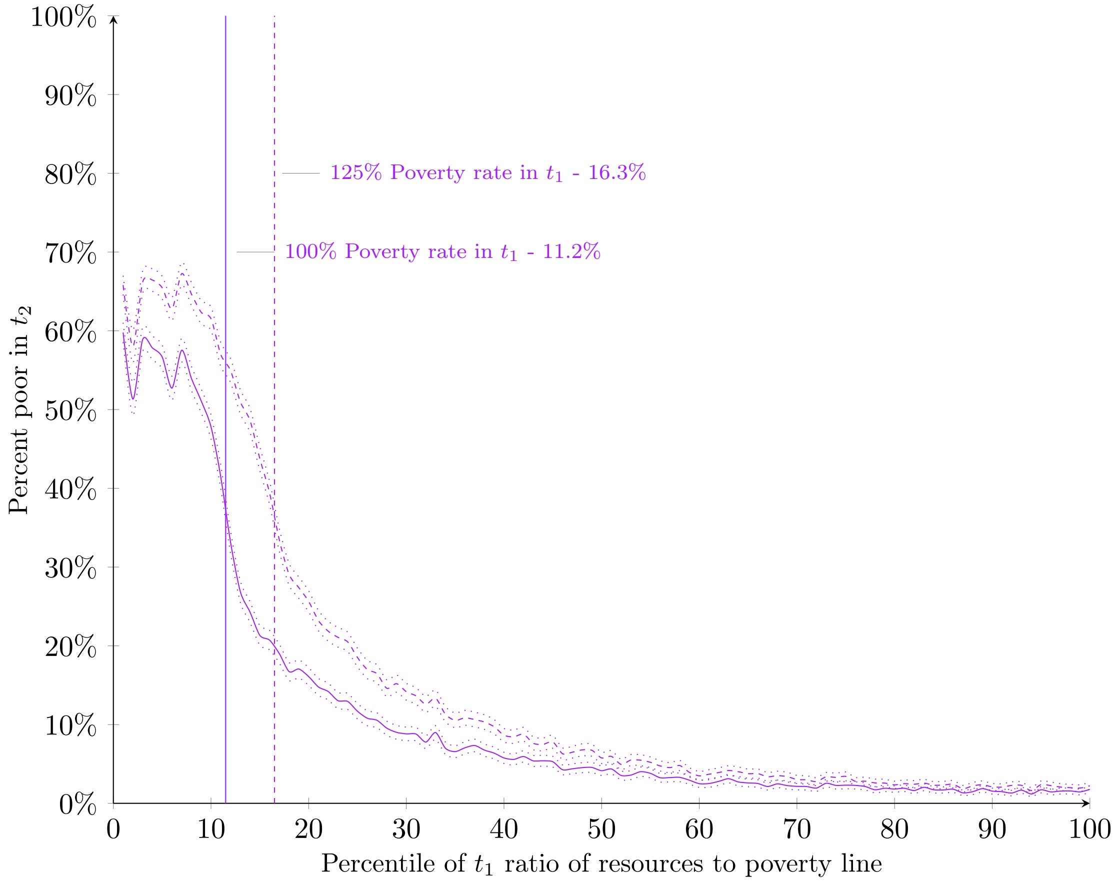
Source: CPS-ASEC 2010-2016. Notes: Because the sample consists of two-year panels, the poverty rate uses t_2 observations such that the poverty rate is the sum of those entering poverty (NonPoor to Poor) and those staying in poverty (Poor to Poor).

Figure 2.4: Percent of families in poverty in t_2 by t_1 ratio of resources to poverty line, OPM v. SPM



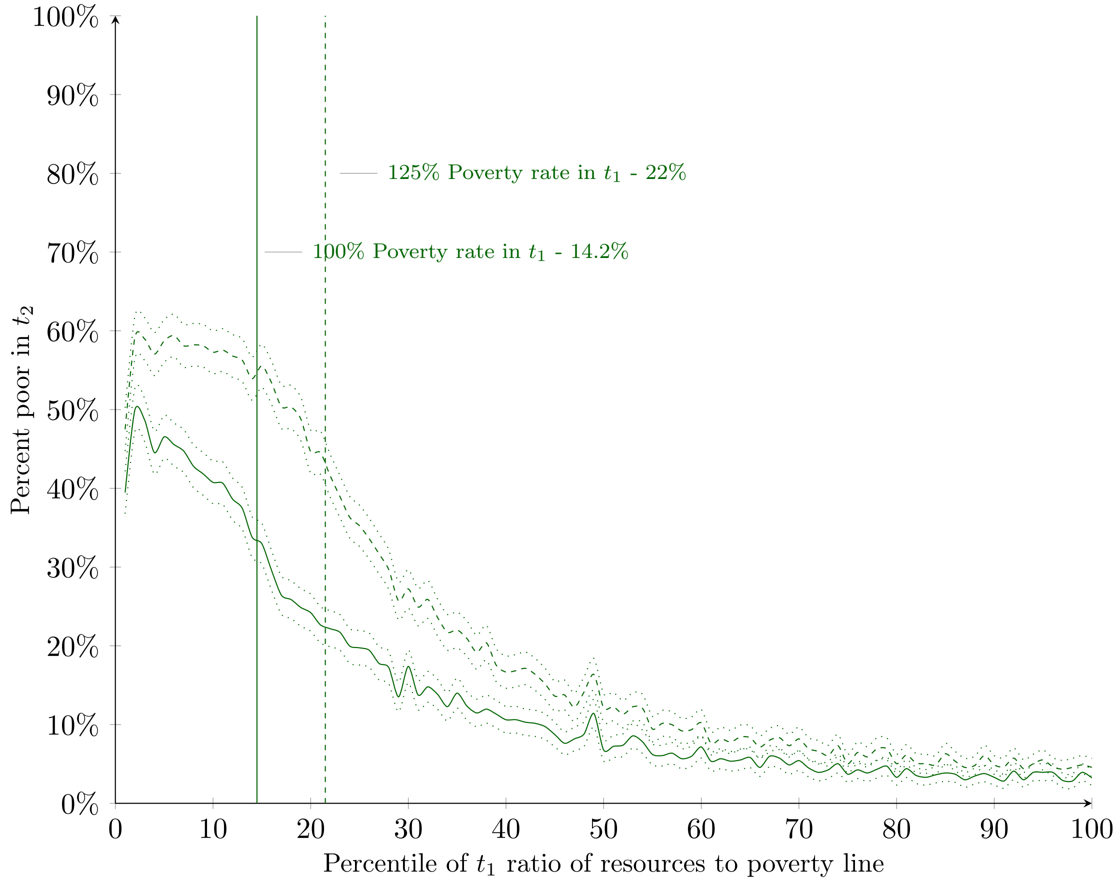
Source: CPS-ASEC 2010-2016. Notes: The probability of being in poverty in t_2 is higher for families with lower t_1 resources. For example, the probability of being in poverty for families two times over the poverty line (43rd percentile of t_1 ratio of resources to poverty line) is about 10 percent under SPM definitions. The definitions of OPM and SPM differ on various levels including resource components and poverty lines. For a complete overview, see Section 2.2.

Figure 2.5: Percentile of t_1 ratio of resources to alternate poverty lines, OPM



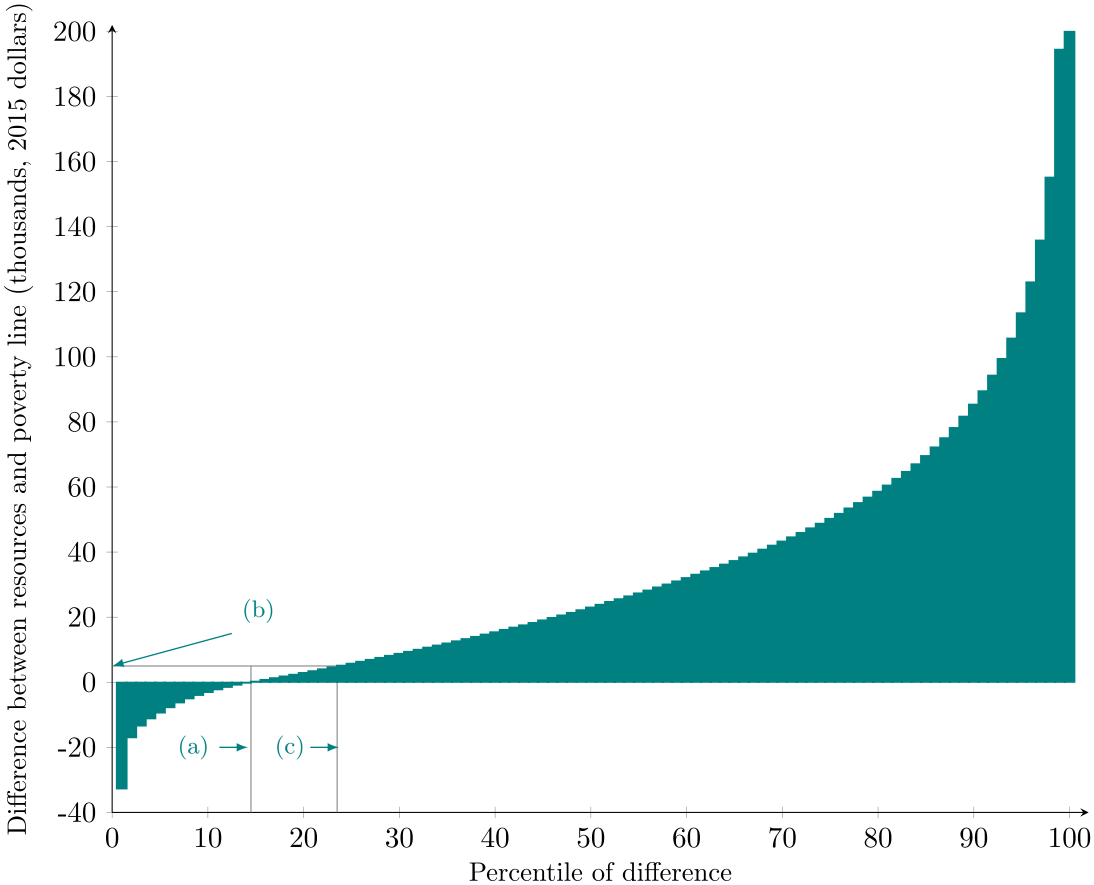
Source: CPS-ASEC 2010-2016. Notes: The alternate poverty lines and rates are calculated by multiplying the poverty lines in both t_1 and t_2 by 1.25 and then recalculating the t_1 ratio of resources to poverty line and each family's t_2 poverty status.

Figure 2.6: Percentile of t_1 ratio of resources to alternate poverty lines, SPM



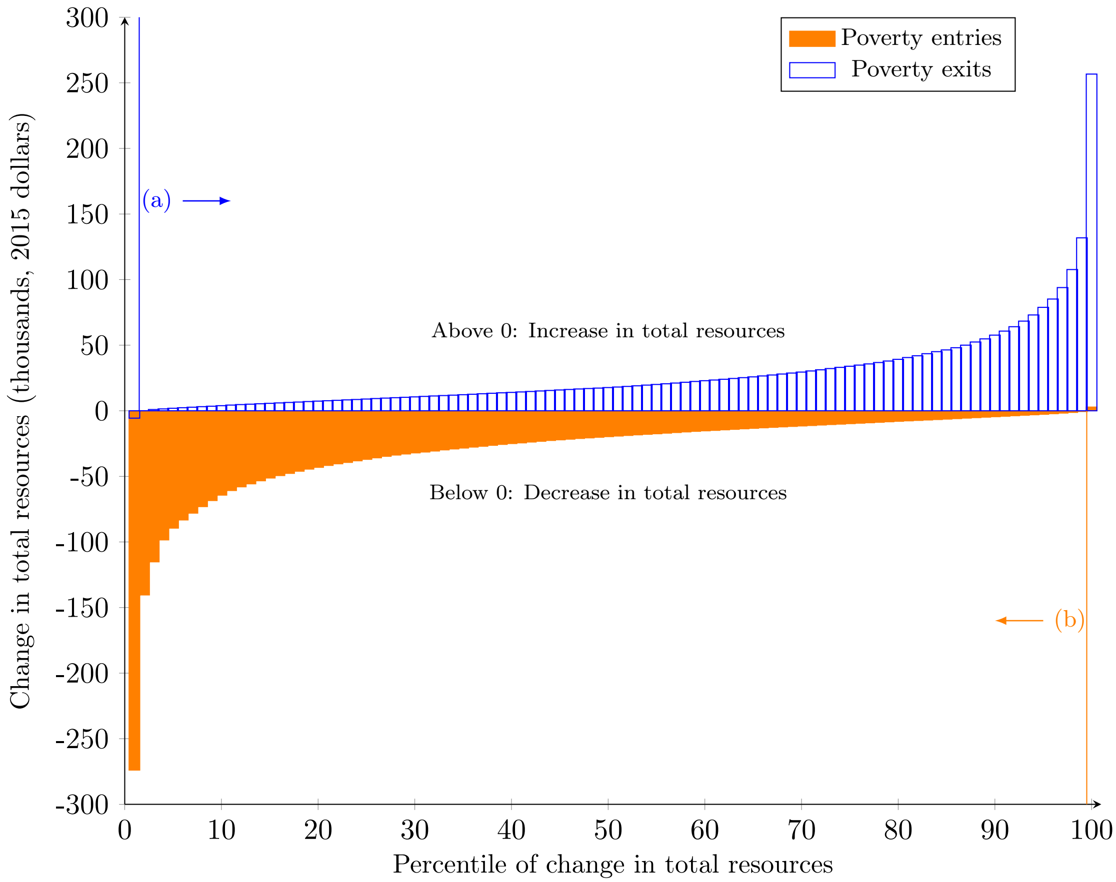
Source: CPS-ASEC 2010-2016. Notes: The alternate poverty lines and rates are calculated by multiplying the poverty lines in both t_1 and t_2 by 1.25 and then recalculating the t_1 ratio of resources to poverty line and each family's t_2 poverty status.

Figure 2.7: The depth of poverty - Resources relative to poverty line in t_1 by percentile, SPM



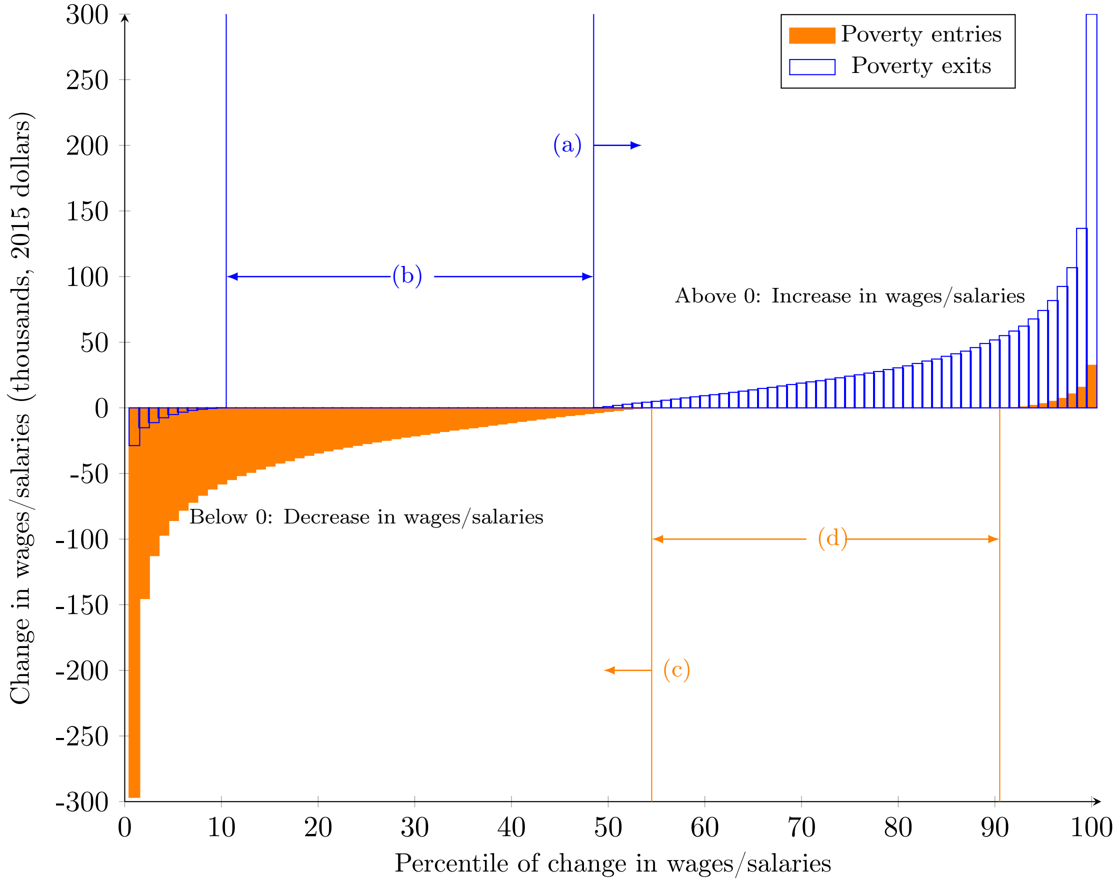
Source: CPS-ASEC 2010-2016. Note: 100th percentile truncated for presentation purposes. Real value is about \$370,000. (a) 14.2% of the families in the 2010-2016 sample were in poverty. (b) All else equal, if all families received a shock in income that reduced their resource by 5,000 dollars then (c) the poverty rate would increase 9 percentage points from 14.2% to 23.5%.

Figure 2.8: Change in total resources by percentile, SPM



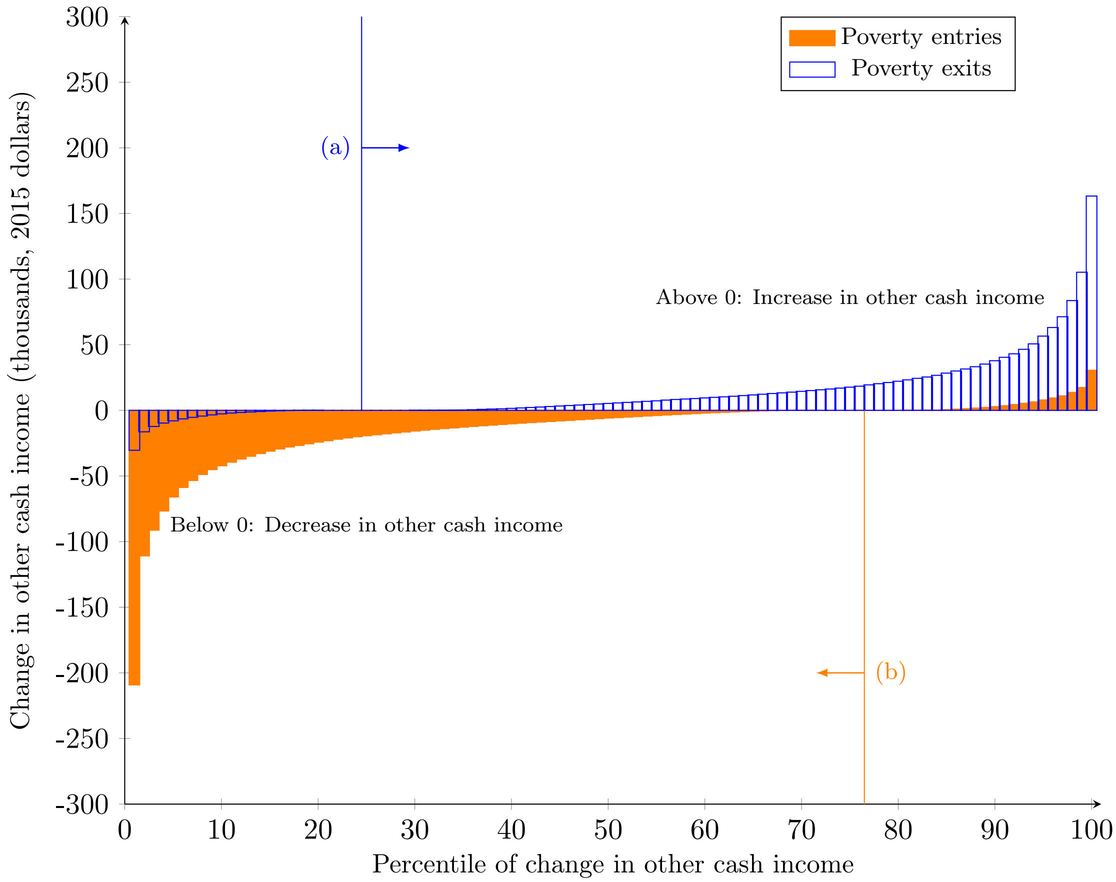
Source: CPS-ASEC 2010-2016. Notes: (a) 99% of families exiting poverty gain resources while (b) 99% of families entering poverty lose resources.

Figure 2.9: Change in wages/salaries by percentile, SPM



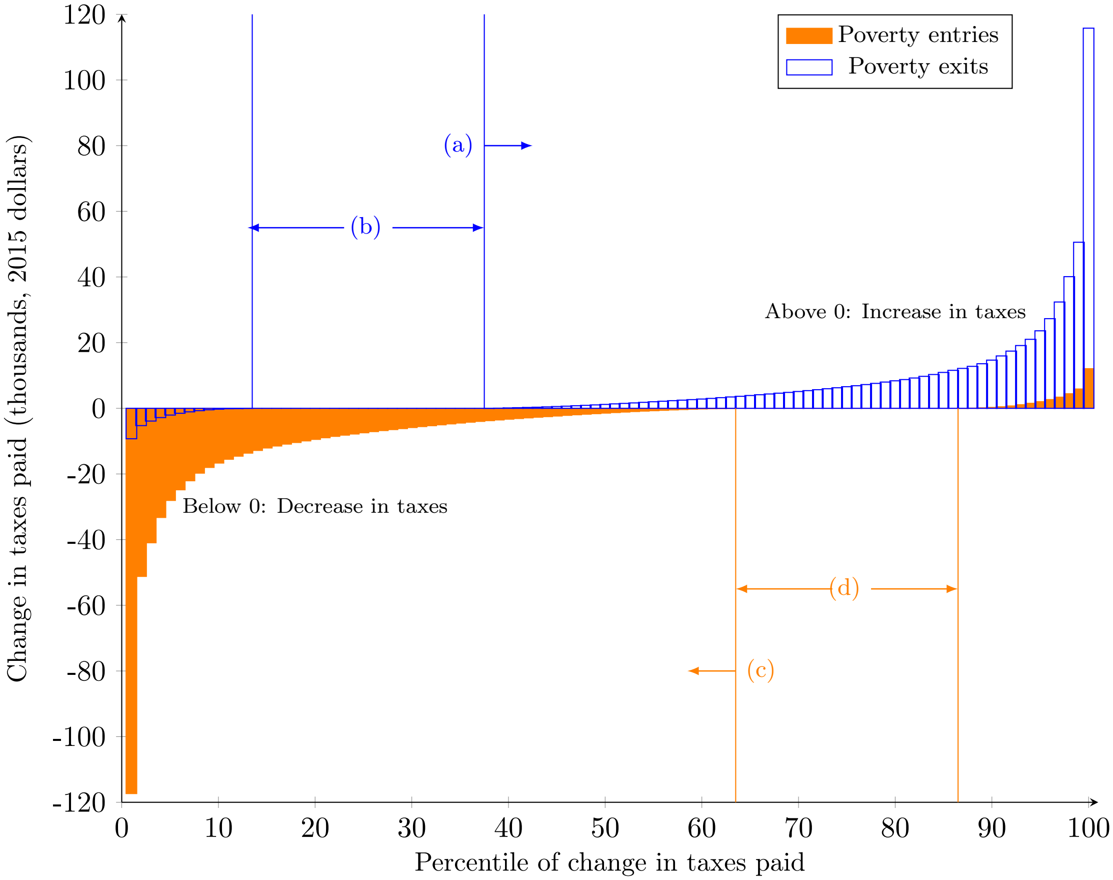
Source: CPS-ASEC 2010-2016. Notes: (a) 51.5 percent of families exiting poverty experience an increase in their wages/salaries while (b) 38 percent of families exiting poverty experience no change. 10.5 percent (not marked) of families exiting experience a *decrease* in wages/salaries. Similarly, (c) 54.5 percent of families entering poverty experience a decrease in wages/salaries while (d) 36 percent of families entering poverty experience no change in their wages/salaries. 9.5 percent (not marked) of families entering poverty experience an *increase* in wages/salaries.

Figure 2.10: Change in other cash income by percentile, SPM



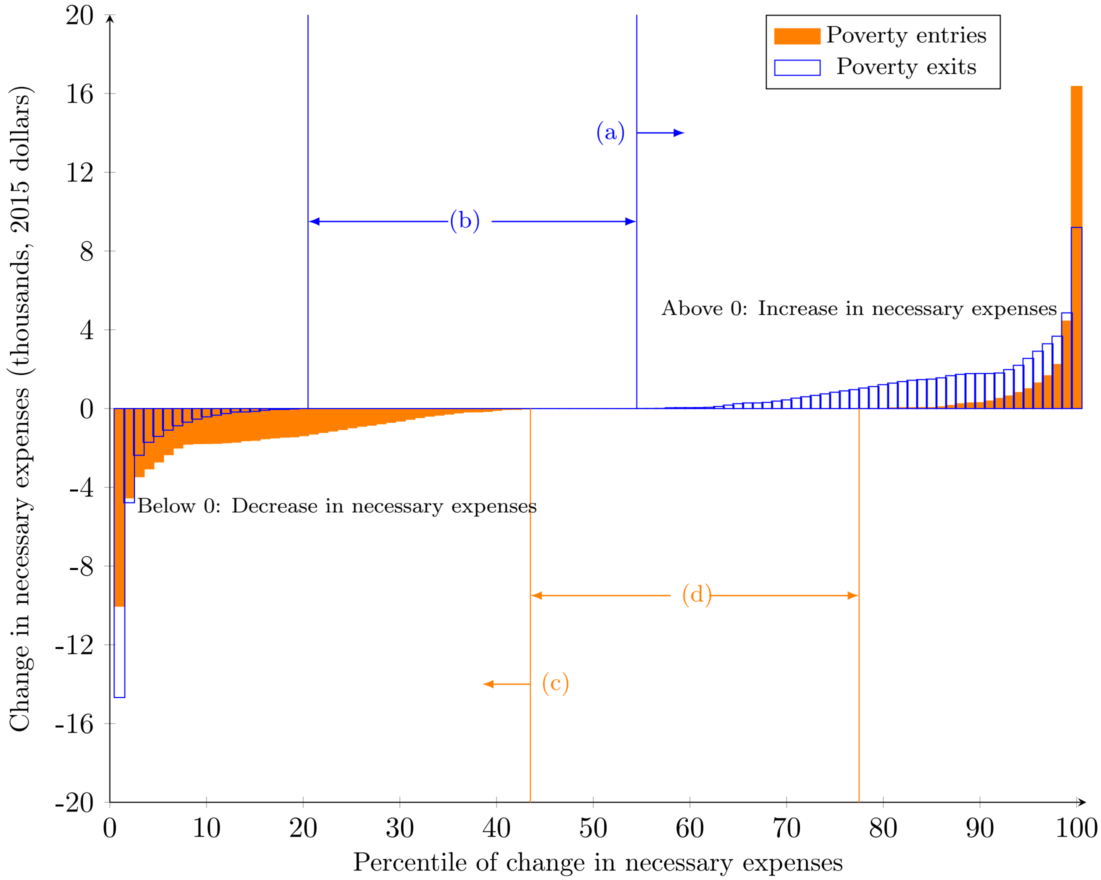
Source: CPS-ASEC 2010-2016. Notes: (a) 75.5% of families exiting poverty experience an increase in other cash income while (b) 76.5% of families entering poverty experience a decrease in other cash income.

Figure 2.11: Change in taxes paid by percentile, SPM



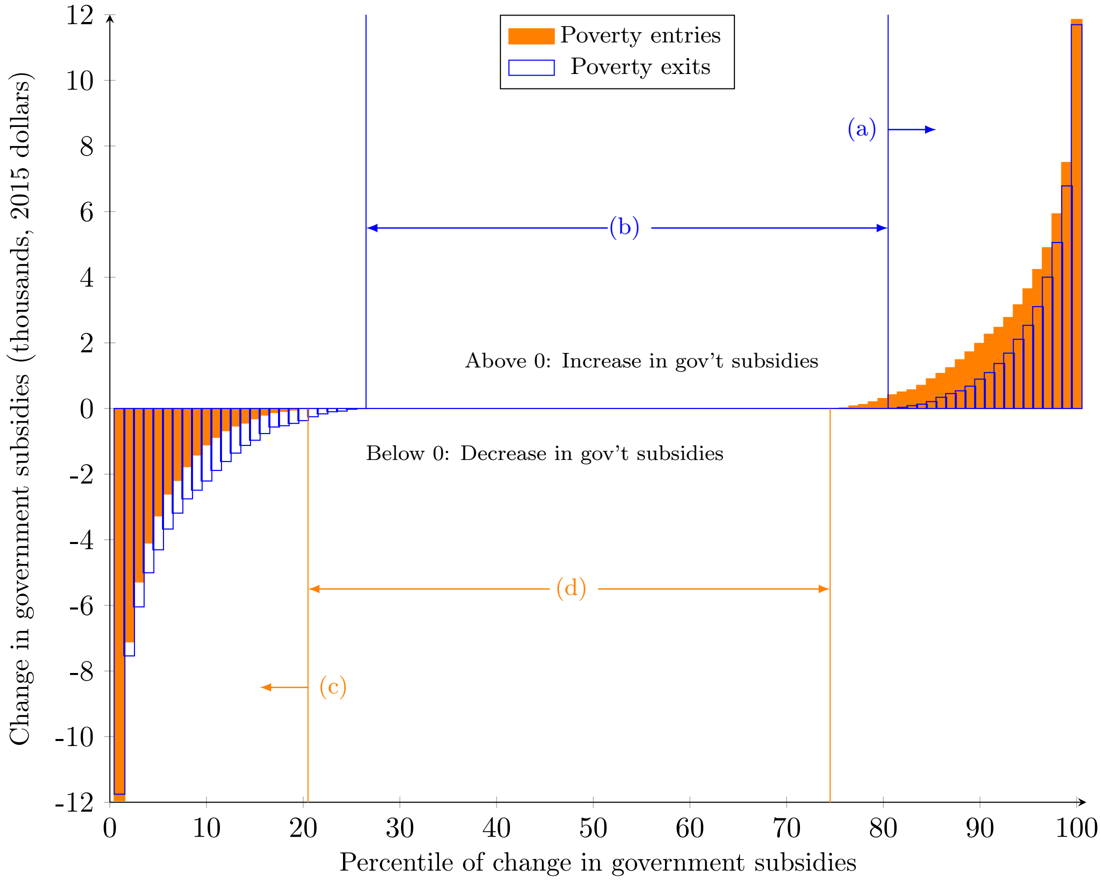
Source: CPS-ASEC 2010-2016. Notes: (a) 62.5 percent of families exiting poverty experience an increase in taxes paid while (b) 24 percent of families exiting poverty experience no change. 13.5 percent (not marked) of families exiting experience a *decrease* in taxes paid. Similarly, (c) 63.5 percent of families entering poverty experience a decrease in taxes paid and (d) 23 percent of families entering poverty experience no change in their taxes paid. 13.5 percent (not marked) of families entering poverty experience an *increase* in taxes paid.

Figure 2.12: Change in necessary expenses by percentile, SPM



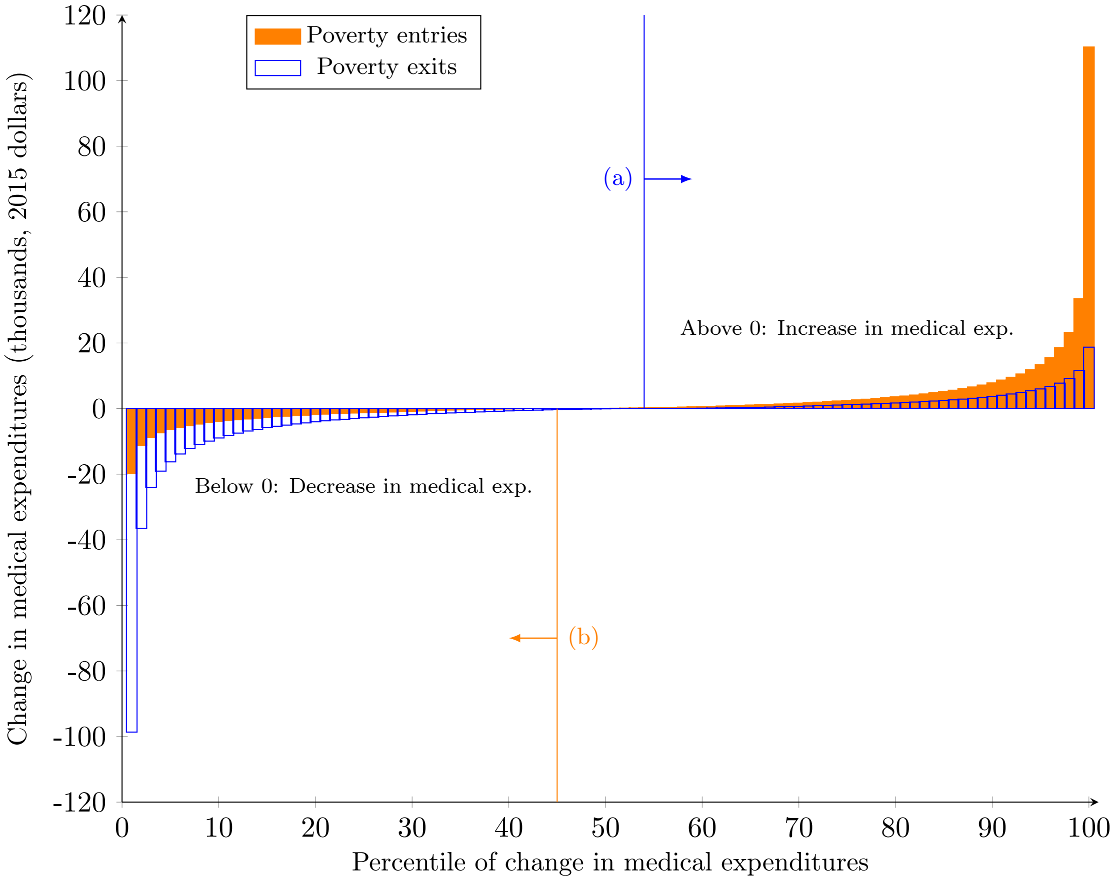
Source: CPS-ASEC 2010-2016. Notes: (a) 45.5 percent of families exiting poverty experience an increase in necessary expenses while (b) 34 percent of families exiting poverty experience no change. 20.5 percent (not marked) of families exiting experience a *decrease* in necessary expenses. Similarly (c) 43.5 percent of families entering poverty experience a decrease in necessary expenses and (d) 34 percent of families entering poverty experience no change in their necessary expenses. 22.5 percent (not marked) of families entering poverty experience an *increase* in necessary expenses.

Figure 2.13: Change in government subsidies by percentile, SPM



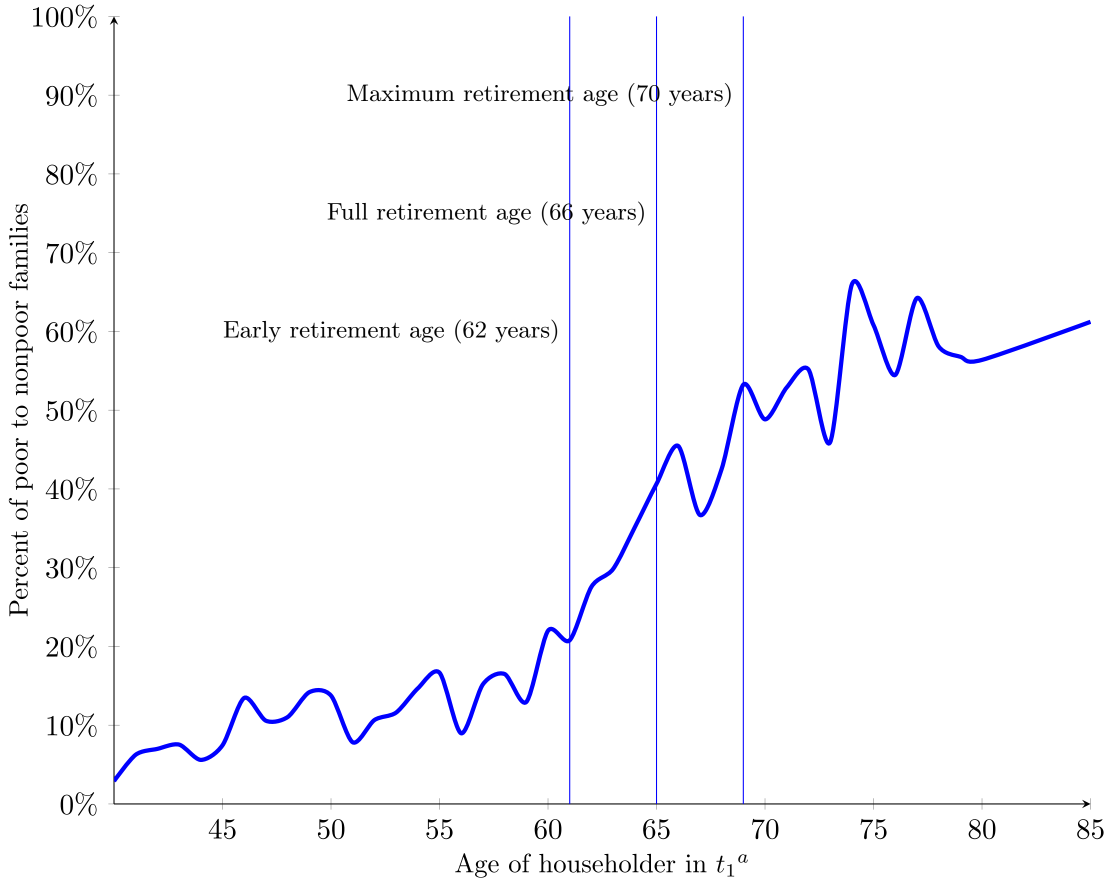
Source: CPS-ASEC 2010-2016. Notes: (a) 19.5 percent of families exiting poverty experience an increase in government subsidies while (b) 54 percent of families exiting poverty experience no change. 26.5 percent (not marked) of families exiting experience a *decrease* in government subsidies. Similarly (c) 20.5 percent of families entering poverty experience a decrease in government subsidies and (d) 54.5 percent of families entering poverty experience no change in their government subsidies. 25.5 percent (not marked) of families entering poverty experience an *increase* in government subsidies.

Figure 2.14: Change in medical expenditures by percentile, SPM



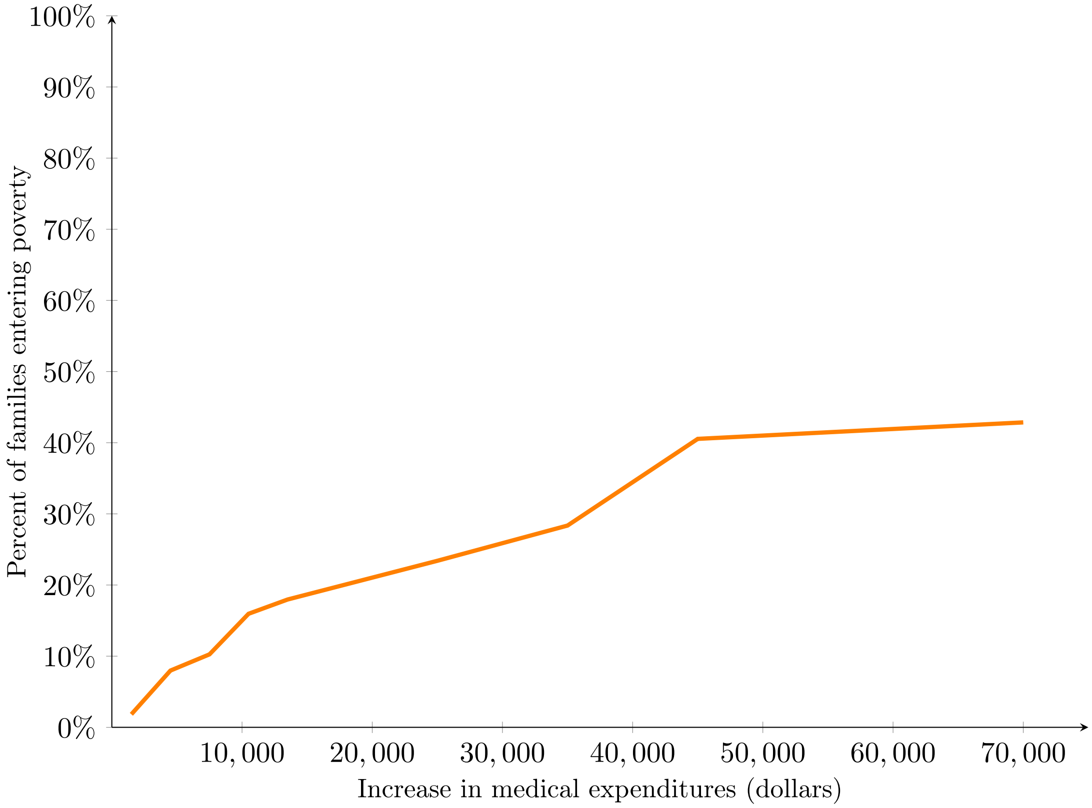
Source: CPS-ASEC 2010-2016. Notes: (a) 46% of families exiting poverty experience an increase in medical expenses while (b) 45% of families entering poverty experience a decrease in medical expenses.

Figure 2.15: Percent of poor to nonpoor families where change in Social Security receipt is sufficient to push families out of poverty in t_2 by age of householder, SPM



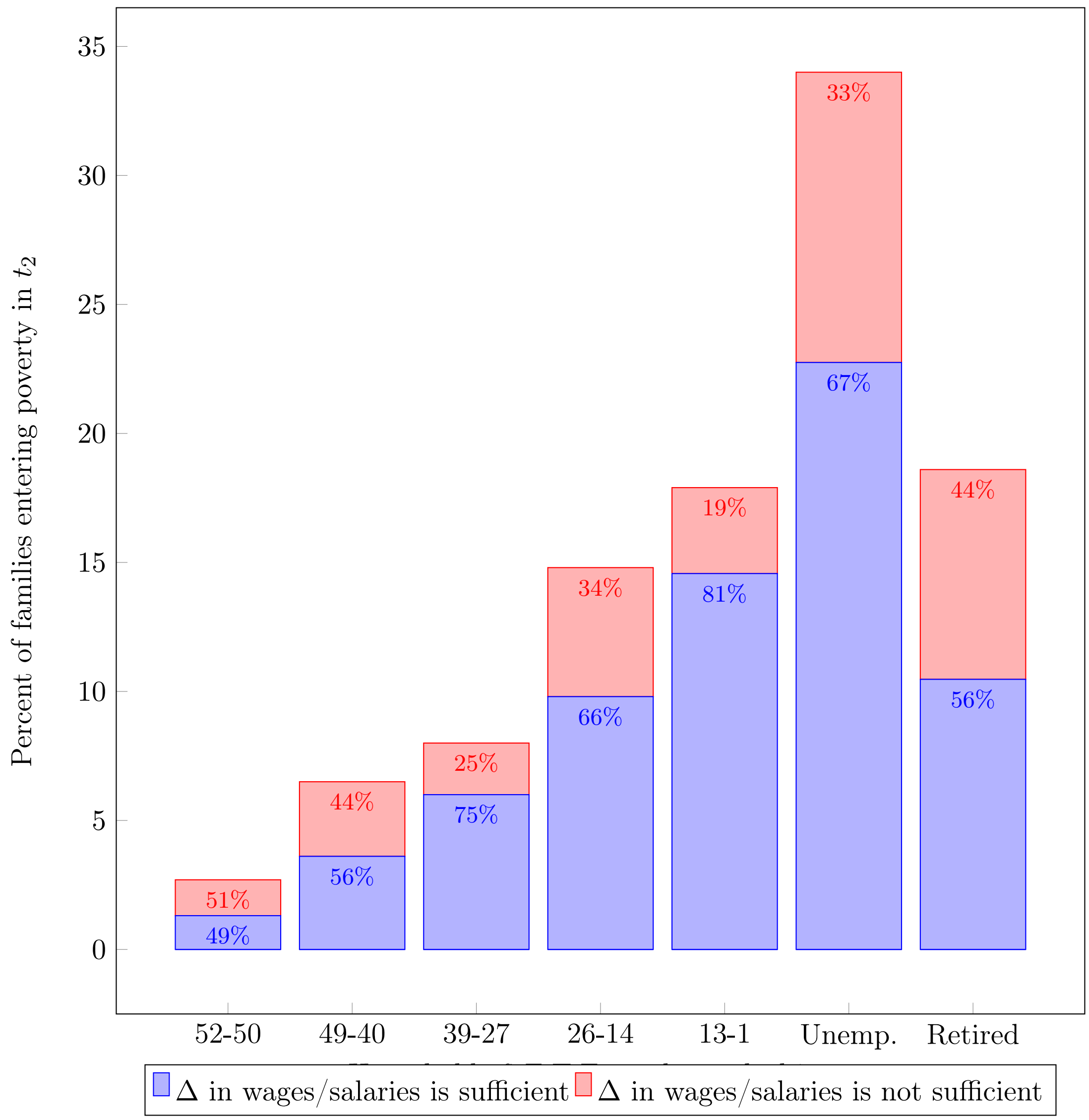
Source: CPS-ASEC 2010-2016. Notes: (a) In order to simplify analysis, the age of the householder is used to identify households that are likely to see an increase in Social Security. The CPS-ASEC refers to a householder as the person in the family who owns/rents the unit. In the case of married couples, the householder may be either one of the spouses. Certainly, there are families where the spouse may be entering retirement age but these are not captured in this analysis. The percentages plotted are derived for families where total SPM resources increase between t_1 and t_2 in following the "sufficient" setup establish in Section 2.6.2. Each age group consists of more than 100 families.

Figure 2.16: Families where increases in medical expenditures are sufficient for poverty entries, SPM



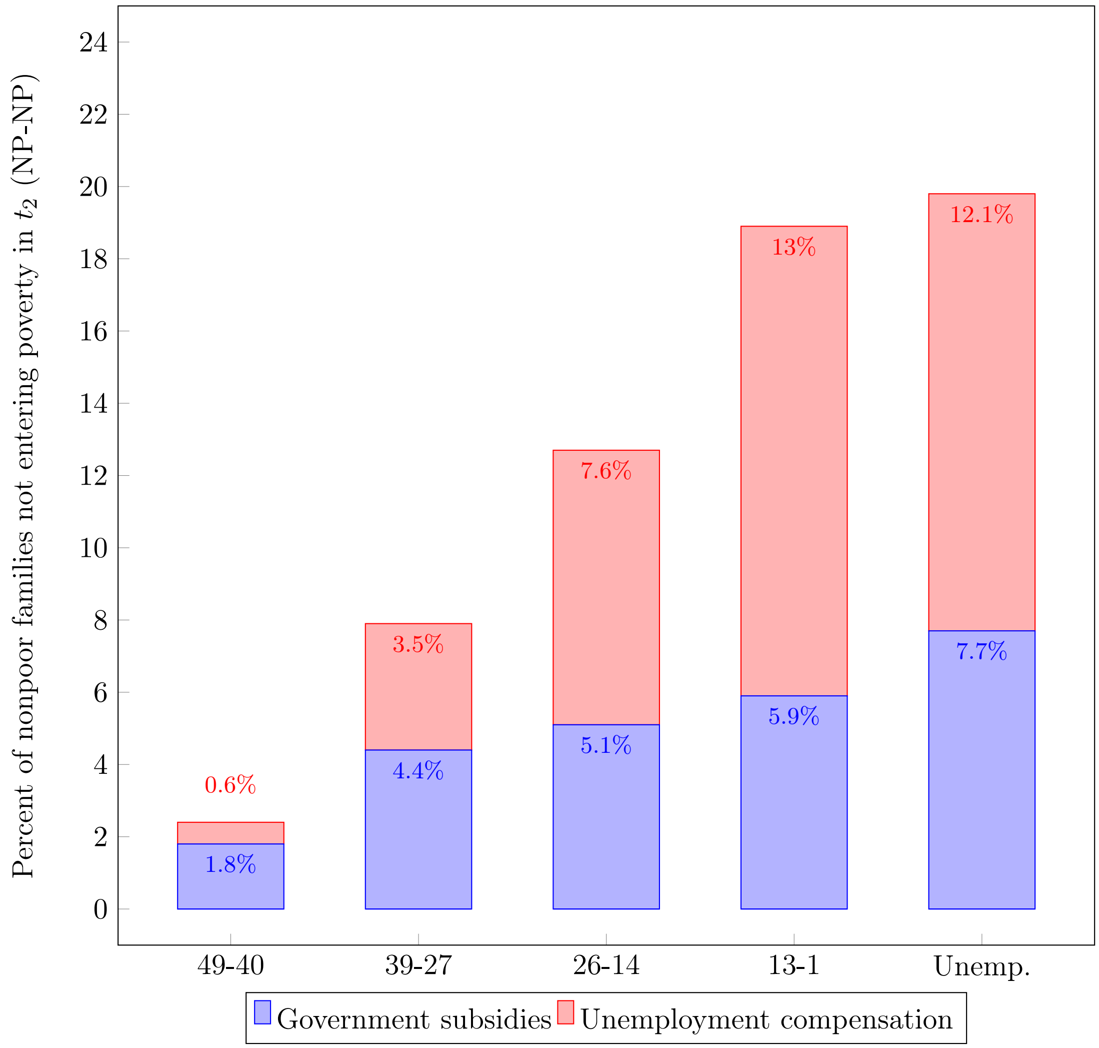
Source: CPS-ASEC 2010-2016. Notes: The probabilities plotted are derived from the number of families where total SPM resources decrease between t_1 and t_2 and medical expenses increase. Nine separate bins (observations) are created: 0-3,000 (3,387), 3,000-6,000 (928), 6,000-9,000 (468), 9,000-12,000 (282), 15,000-20,000 (167), 20,000-30,000 (141), 30,000-40,000 (67), 40,000-50,000 (37), 50,000+ (77).

Figure 2.17: Percentage of nonpoor families entering poverty for householders working 52-50 F.T.E. weeks in t_1 by F.T.E. weeks in t_2, SPM



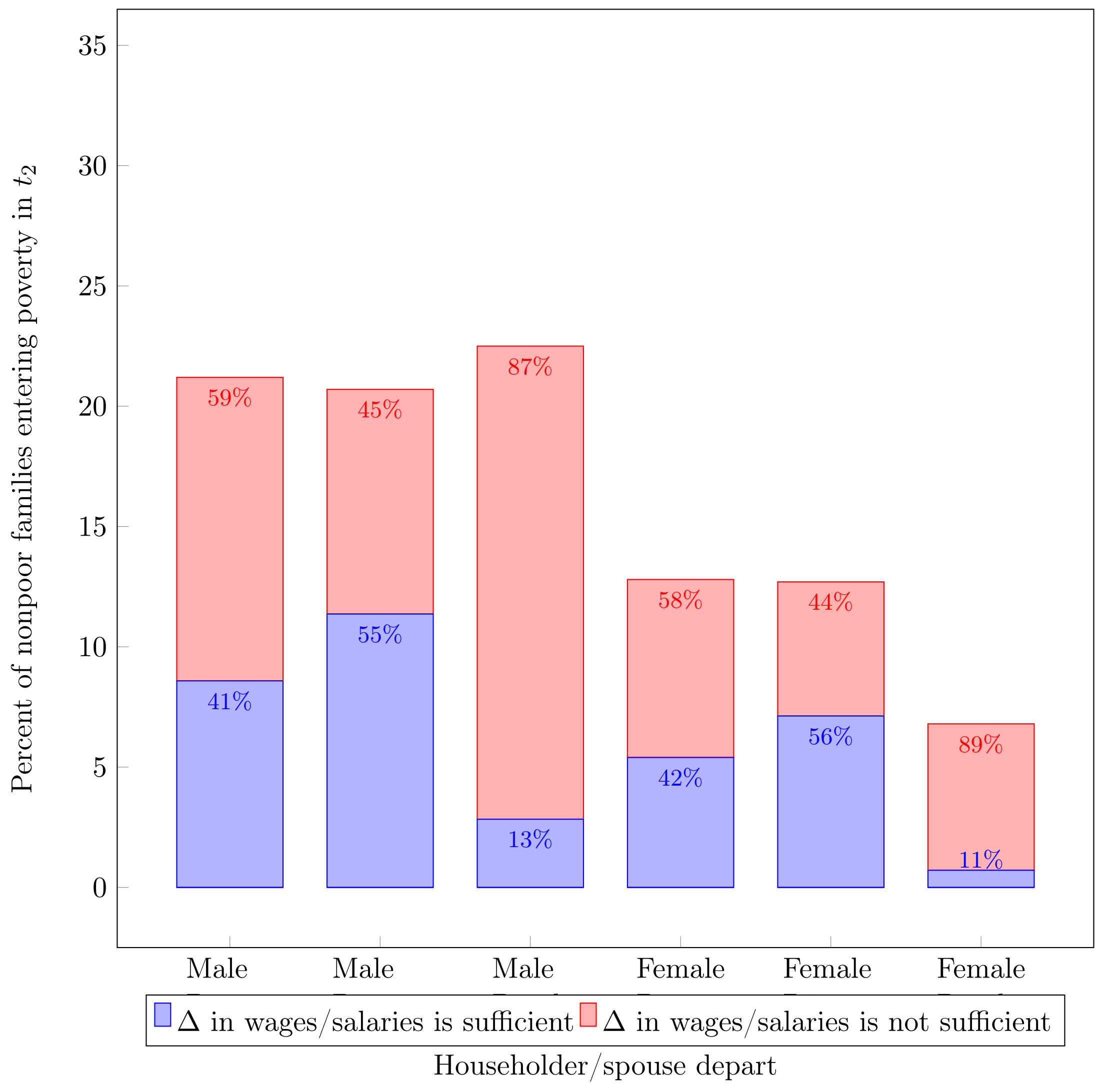
Source: CPS-ASEC 2010-2016. Notes: (a) The CPS-ASEC refers to a householder as the person in the family who owns/rents the unit. In the case of married couples, the householder may be either one of the spouses. Full-time equivalent weeks are calculated from householder responses to weeks worked in the past year who also reported being employed at a full-time basis (35 hours or more per week). The bars present the percentage of families for each work schedule change that fall into poverty in t_2 as well as the percent of those families for whom the change in wages (total family wages) is sufficient for pushing a family into poverty (see Section 2.6.2). For a tabular representation of this figure, see Appendix Table A.4.

Figure 2.18: Percentage of nonpoor families where government subsidies or unemployment compensation keep families out of poverty for householders working 52-50 F.T.E. weeks in t_1 by F.T.E. weeks in t_2, SPM



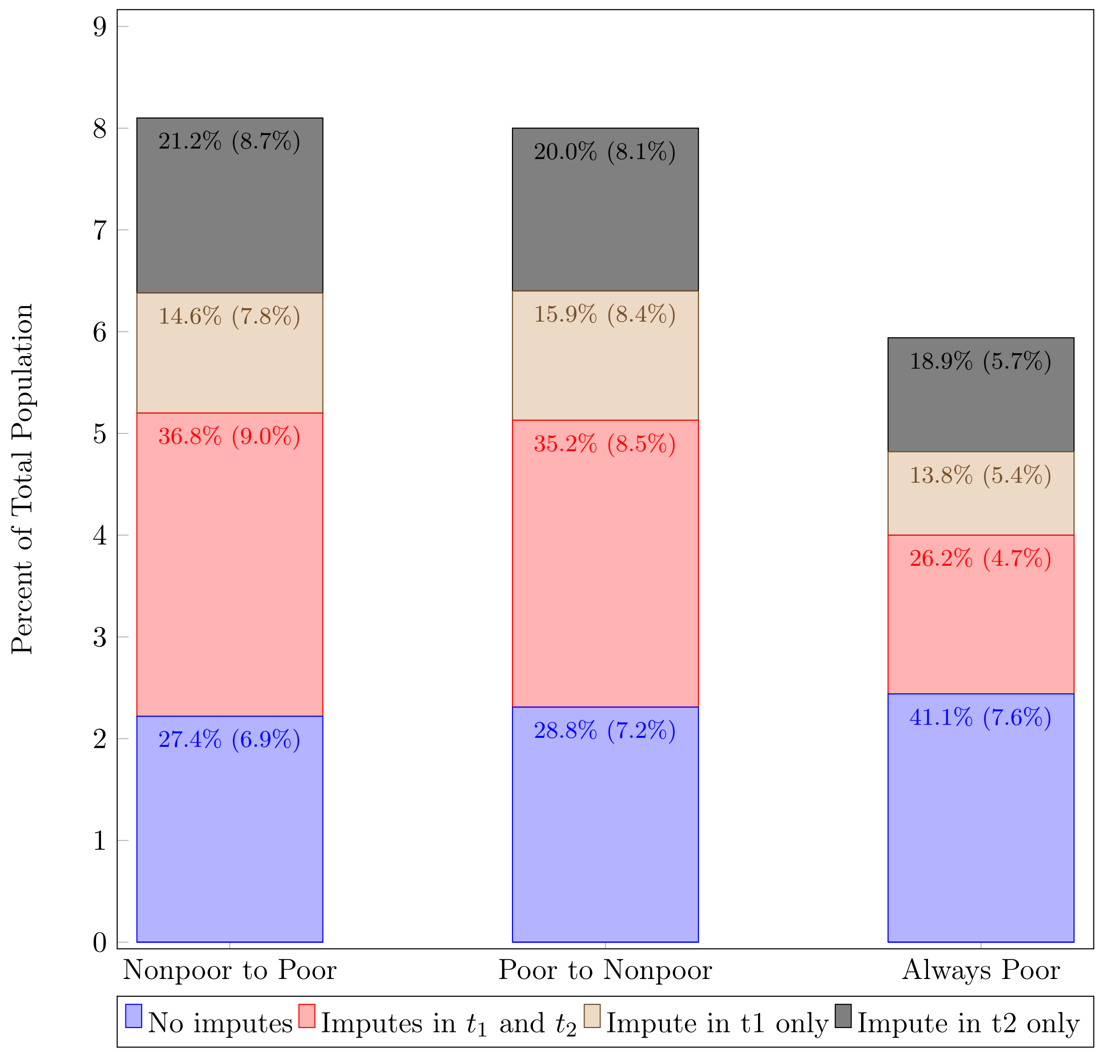
Source: CPS-ASEC 2010-2016. Notes: In order to calculate the percent of families staying out of poverty, government subsidies in t_2 are subtracted from SPM total resources in t_2. Families that fall below the poverty line without the government subsidies are considered to be those families for whom government subsidies keep families out of poverty. Each bar represents the percentage of nonpoor families in that work schedule change category that do not fall into poverty because of government subsidies and the percent of families that do no fall into poverty because of unemployment compensation. Government subsidies include SNAP, lunch subsidies, WIC, housing subsidies and energy subsidies *as well as* EITC. householders working 52-50 F.T.E. weeks in t_1 and 52-50 F.T.E. weeks in t_2 or retiring in t_2 are not presented as the number of families are essentially zero. For a tabular representation of this figure, see Appendix Table A.5.

Figure 2.19: Percent of nonpoor families in t_1 who fall into poverty in t_2 as a result of a householder or spouse departure, SPM



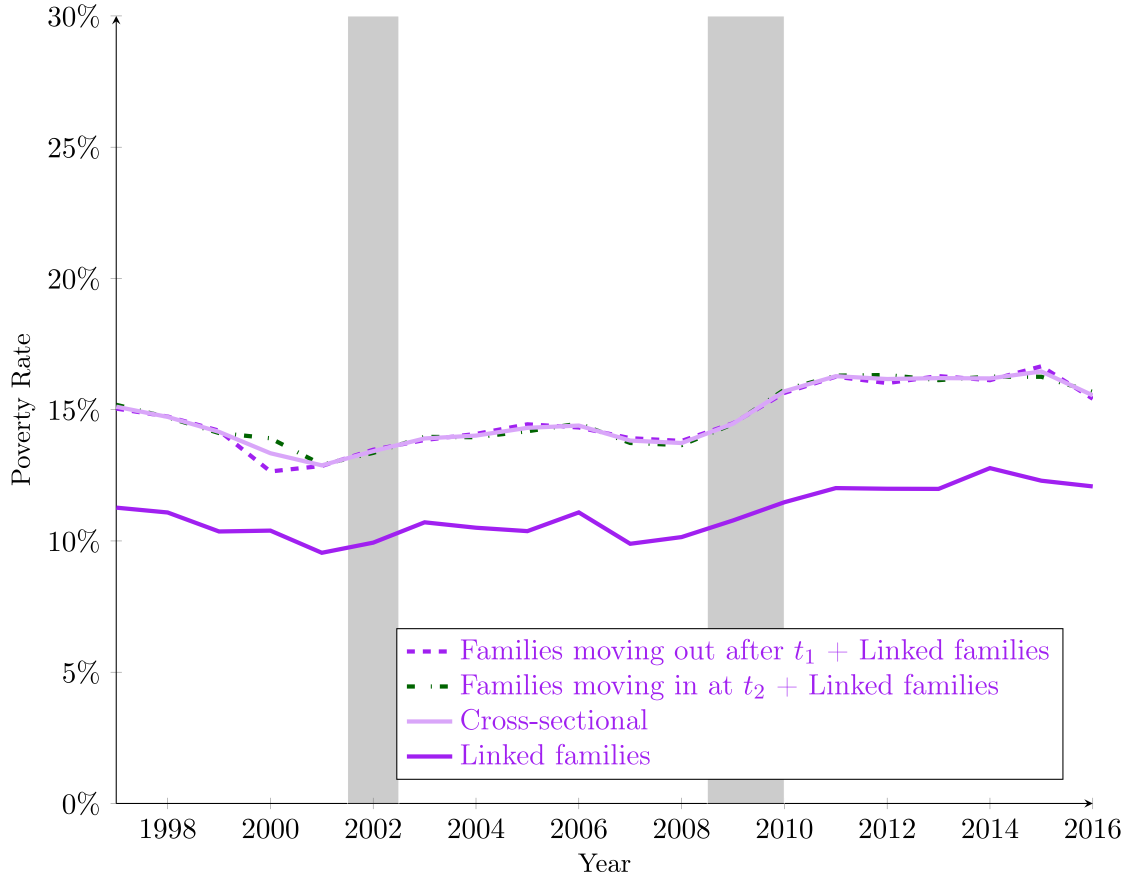
Source: CPS-ASEC 2010-2016. Notes: The bars represent the percent of a particular departure that fall into poverty as well as the percent of those families where the change in wages/salaries are sufficient to push those families into poverty. The sample includes only nonpoor families. For a tabular representation of this figure, see Appendix Table A.6.

Figure 2.20: Decomposing poverty rates by imputation type, SPM



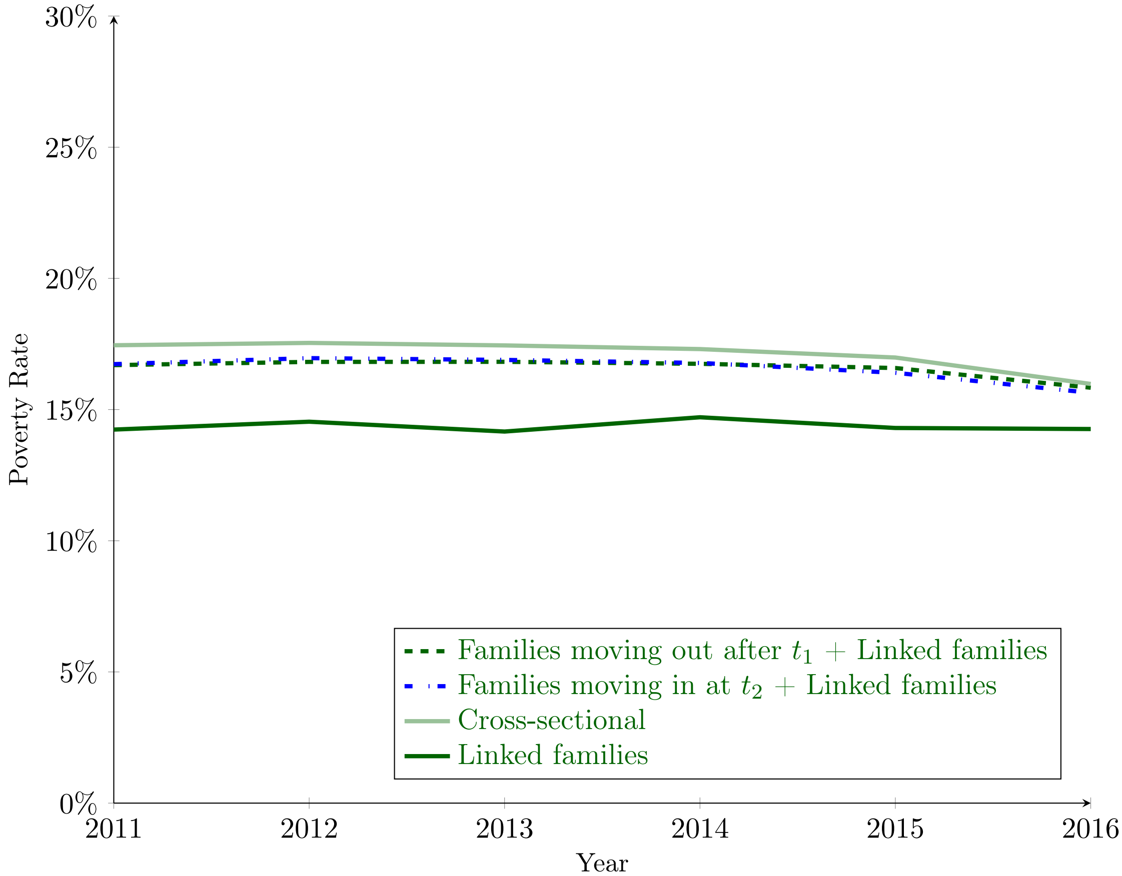
Source: CPS-ASEC 2010-2016. Notes: Two percentages are displayed in each block. The first set presents the percent that particular block represents of each type of poverty transition. The second, in parentheses, is the recalculated percent of total population where the total population includes only that particular type of imputation.

Figure 2.21: Gross flows as percent of total U.S. population, OPM



Source: CPS-ASEC 1996-2016. Notes: Shaded areas denote recession periods. The linked-CPS sample keeps families that are present in both t_1 and t_2. The linked poverty rate denotes the poverty rate for t_1. For example, the linked poverty rate in 2002 represents the linked sample for families where t_1 is 2002 and t_2 is 2003. Families moving out in t_1 are those families present in the CPS-ASEC only in t_1. These families are then replaced by families moving into the sample in t_2. The cross-sectional rate uses the cross-sectional CPS-ASEC, which includes all families regardless of their being in sample in t_2, to calculate the poverty rate.

Figure 2.22: Comparing poverty rates for linked-CPS samples, SPM



Source: CPS-ASEC 2010-2016. Notes: The linked-CPS sample keeps families that are present in both t_1 and t_2. The linked poverty rate denotes the poverty rate for t_1. For example, the linked poverty rate in 2002 represents the linked sample for families where t_1 is 2002 and t_2 is 2003. Families moving out in t_1 are those families present in the CPS-ASEC only in t_1. These families are then replaced by families moving into the sample in t_2. The cross-sectional rate uses the cross-sectional CPS-ASEC, which includes all families regardless of their being in sample in t_2, to calculate the poverty rate.

Chapter 3

Union card or welfare card? Evidence on the relationship between union membership and net fiscal impact at the individual-worker level[1]

3.1 Introduction

This paper offers the first evidence on whether union membership causes workers to use less public benefits and to pay more taxes. The prior literature's findings of union wage and benefit premiums gives reason to expect this, although prior work has neither tested nor measured this directly. The literature's evidence about effects of unionization on wages and benefits is insufficient to understand the effect of unionization on taxes paid or benefits received. Tax and public-benefit effects depend on interactions of workers' earnings with household characteristics and tax and benefit policy. For instance, a union-induced 10 percent wage increase will have different tax and benefit implications for a worker earning near the poverty line versus one earning at the median, for a childless worker versus with one with 3 children, and for a worker in California versus Mississippi.

[1] **This chapter is co-authored with Aaron Sojourner, Associate Professor, Department of Work and Organizations.**

Through unionization, many workers raise their labor compensation, both in earnings and employer-provided fringe benefits. The positive effect of unionization on labor earnings is especially pronounced for workers who would otherwise have very low earnings. Frandsen (2012) follows workers after close union elections and finds that unionization strongly raises post-election earnings for workers who were below the 25th percentile of the pre-election earnings distribution but has no effect for workers who were at higher percentiles. Frandsen's focus on earnings, rather than wages, accounts for any reduction in hours induced by higher hourly compensation. He also follows workers even if they leave the establishment and counts earnings as zero if they do not earn from any employer, so this also accounts for any reductions in employment driven by unionization. Union membership also raises workers' likelihood of having private, employer-provided health insurance and other benefits (Buchmueller, DiNardo and Valletta, 2002; Freeman and Medoff, 1984; Freeman, 1981). Employer expenditures on fringe benefits are 2.5 times higher per hour worked for unionized jobs than for nonunion jobs and, as with earnings, the effects of unions on benefits appear larger in lower-paying establishments (Budd, 2007).

Through these channels, unionization may have a positive net fiscal impact on public balance sheets by both (1) reducing public-benefit use and (2) increasing tax payments by workers. While media have speculated about this fact (Eidelson, 2013; Sanders, 2012), it has not received much direct attention from economists or social scientists. Economists have understandably focused most of our attention on the effects of unions on wages, employment and hours, and labor and organizational productivity. While these are the first-order, narrowly-economic questions, we have ignored closely-connected questions of social, policy, and economic import. For instance, what is the impact of unionization on household income? There is some work on labor earnings, the product of wages and hours, but little attention to other kinds of income or on contributions to and dependence on the public fisc. There has been extensive study of costs and benefits to the public of numerous, other economic phenomenon. Immigration (Auerbach and Oreopoulos, 1999; Blau, 1984; National Academies of Sciences, Engineering, and Medicine, 2016; Preston, 2014; Storesletten, 2000) and early childhood investments (Council of Economic Advisers, 2015; Elango et al., 2015) are two prominent examples. The analysis here is static, focusing on annual effects, rather than dynamic, computing a net present value of a point-in-time event (National Academies of Sciences, Engineering, and Medicine, 2016).

This paper estimates the average net fiscal impact of union members to observably-similar non-member workers using data from the Current Population Survey over 1994 to 2015. We measure individual net fiscal impact (NFI), which is taxes paid (T) less the cost of public benefits received (B): NFI = T - B. Theory tells us that the key mechanism by which individual unionization would affect these variables is through raising private income among low earners. The analysis yields evidence strongly consistent with this and provide the first estimates of the magnitude of these relationships. Additional analysis explores sensitivity to issues arising from the possibility that union membership affects transitions out of employment. Looking beyond this paper's main focus on worker-level analysis, the conclusion offers interpretation given evidence on other channels by which union membership might affect NFI, such as by affecting labor productivity, profit, and public policy.

3.2 Research design

We would ideally have an experiment among a representative sample of workers where some were randomly assigned to be union members and others to be nonunion. In that case, we could credibly interpret any observed union-nonunion differences in outcomes as causal effects of union membership. Unfortunately, randomization is not feasible.[2] Freeman (1984) describes many relevant issues in the study of union effects using CPS data arising from measurement error in the observed union-status variable. In particular, he discusses plausible conditions under which the true effect of union membership is bounded above by the cross-sectional estimator and below by the individual first-difference estimator. Following his lead, we will present both estimates and interpret our results in this framework.

To get a nationally-representative sample, we use the Current Population Survey (CPS), which includes detailed data on all key variables (Flood et al., 2015). The study period is 1994 to 2015, the longest over which the necessary variables are all available. Careful linking

[2] A more-credible, regression-discontinuity design (DiNardo and Lee, 2004; Frandsen, 2012; Sojourner et al., 2015) would require the ability to connect the population of individually-identified workers between the establishment where they worked during a NLRB unionization election and later, individually-identified measures of taxes paid and benefits received. This is not feasible.

is required to maximize sample size conditional on the necessary variables. Specifically, we focus on the subsample who were given both the Annual Social and Economic Supplement (ASEC) and the Outgoing Rotation Group (ORG) survey. The ASEC is necessary because it contains income, tax, and benefit data for outcomes. Annually, the ASEC fields only in March so only a third of the total CPS sample are eligible. The ORG is necessary to measure union-membership status. Two sets of individual, longitudinal identifiers recently produced by the Minnesota Population Center enable both linking of ASEC and ORG responses (MARBASECID, Flood and Pacas (2016)) and linking of ORG responses across consecutive years (CPSID, Rivera Drew, Flood and Warren (2014)). The Appendix gives details on the linking process and sample construction.

As is common in the study of union effects on wages, our primary sample screens in only non-student, employed, wage and salary workers age 18 or older.[3] The longitudinal nature of our analysis requires a few additional restrictions. First, the sample includes only observations that are present in both year t and t+1. Second, in linking the March ASEC to ORG questions from April through June, we keep only observations where age, sex and race match. More importantly, because we need union membership regarding the job referred to in March, we only keep observations whose April-June job characteristics match those job characteristics from March. In effect, these variables include, labor force status, employment status, class of workers, and broad classifications of industry and occupation. Third, for the links across years and months, we keep only observations that match on age, sex and race. For consistency between the cross-sectional and longitudinal analysis, we restrict attention to individuals linked across two waves of the CPS-ORG and meeting sample-screening conditions in both.

All cross-sectional analysis uses each observation's sample weights. Longitudinal analysis gives each individual the average sample weight of its two observations. Dollar amounts are inflated to 2015 dollars. Since we use a longitudinally linked-CPS, our sample consists of only those people who do not move from their household between two consecutive years. We assess robustness using an extended sample that includes those both in and out of the labor force across the two observations.

[3] Non-workers generally cannot belong to unions and plausibly have different unobserved characteristics than workers. If unionization impacts public balance sheets by reducing employment, our primary analysis will miss this channel. We return to this in the robustness section.

Nonresponse for various CPS variables are a potential source of bias for our analysis, as evidenced by the growing literature on imputations in CPS. Our general approach for dealing with nonresponse is to drop cases with nonresponse and then conduct robustness checks. More specifically, when a respondent refuses to respond to a particular survey question, rather than leaving the field blank, the Census Bureau allocates a value from a donor set comprised of respondents from that same sample. The process by which the allocation is conducted is known as the hot-deck imputation procedure and, in essence, takes a nonrespondent and matches based on a set of measured attributes. For earnings items, this set broadly consists of age, sex, race, employment status, and industry/occupation. As early as 1986, Lillard, Smith and Welch investigated the Census Bureau's approach to dealing with missing data and pointed out that the hot-deck procedure for imputing income likely affects results regarding income and earnings (Lillard, Smith and Welch, 1986). More recently, various studies (Hirsch and Schumacher, 2004; Bollinger and Hirsch, 2006, 2013) have warned about missing data in the CPS highlighting that "coefficient bias resulting from imputation of a dependent variable (earnings) can be of first-order importance" (Hirsch and Schumacher, 2004).

In our analysis, item nonresponse on earnings, union-status, and all sources of income raise familiar issues. The Census Bureau imputes values for missing data. However, relying on these imputed values has been shown to introduce bias in analysis like ours. Hirsch and Schumacher (2004) showed that using imputed earnings as an outcome, "if the attribute under study is not used as a census match criterion in selecting a donor, wage differential estimates (with or without controls) are biased towards zero." More importantly, "this bias is large and exists independent of any from the nonrandom determination of missing earnings" (p. 691). Bollinger and Hirsch (2006) estimated attenuation bias from missing union-status data to be about 5 percentage points for estimates between 1999 and 2001. The prevalence of imputations has only increased since 2001. We drop observations with imputed union status to reduce the attenuation bias introduced by the imputation itself. About 6 percent of the full sample have imputed union status. Secondly, the hot-deck procedure used for imputing earnings leads to attenuation that is roughly the size of the imputation rate (Bollinger and Hirsch, 2006). As they suggest, "the simplest approach to account for match bias is to omit imputed earners from wage equation (and other) analyses" (p. 517). Following their recommendation, we also drop individuals with any imputed earnings, who are about 45 percent of the sample.[4] Third, for respondents who answer the March Basic CPS but refuse to answer the longer ASEC questions, the Census Bureau performs a "full-line" impute for these cases, imputing answers to every income question. In other words, there are respondents in the March Basic for whom there is not enough income data collected. Rather than leaving these cases as non-responses, the Census Bureau uses a hot-deck procedure to impute the values of the missing income data (Stewart, 2002). Hokayem, Bollinger and Ziliak (2014) analyzes the role of nonresponse for the CPS-ASEC including the full-line impute and find evidence of bias from non-response. More importantly, in a longitudinal framework, dropping respondents with full-line imputes is preferred. In effect, comparing a full-line impute in one time period to an actual response in the second time period introduces unnecessary measurement error. We drop respondents with full-line imputes, who are about 14 percent of the sample. These three categories are not mutually exclusive. Dropping all observations with any type of imputation means dropping just over half (57 percent) of our otherwise-eligible sample. Table B.5 breaks down the resulting sample sizes from each imputation restriction imposed here.[5]

The primary analysis sample is 241,906 individuals, each observed in two consecutive years. Table 3.1 describes the frequency of union-status transitions in our sample. The sample includes 3,742 individuals moving from union in the first wave to non-union in the second, 3,986 moving from non-union to union, 14,185 who are union in both waves, and 99,040 who are non-union in both waves. We treat covered non-members as nonunion. This is a conservative assumption, as it diminishes the contrast between union and nonunion categories.

A different form of measurement error arises from any inaccurate reports of union status. While we present specifications for a balanced, pooled cross-sectional model, we also use an individual fixed-effect model. Doing so allows us to make a better causal claim of the effect of unionization. Indeed, Freeman (1984) points out two relevant facts: (1) there is substantial

[4] The analyses of Bollinger and Hirsch (2006) focuses primarily on a single earning variable from the ORG files while we focus on a larger set of variables from the CPS-ASEC (listed in Table B.3). But the imputation method is nearly the same for both surveys and, more importantly, imputation is more pronounced in the CPS-ASEC. These facts further warrant dropping the observations with imputed earnings and union status.

[5] We also look at the effect of dropping free/reduced price lunch and housing subsidies from our analysis. We find that there is no significant effect of doing so our final model includes these benefits. Results not presented.

measurement error in reported union status and (2) this can bias down estimates based on individual fixed effects. This measurement error comes from inaccurate responses, rather than the nonresponse discussed above. Freeman (1984) further argues that cross-sectional estimates can be interpreted as an upper bound on the causal effect of unionization, due to likely positive omitted-variable bias. That is, given high union wages, it is typically assumed that firms selected workers with higher unobserved ability which cannot be controlled for in a standard cross-sectional setup. Furthermore, the fixed effect estimate can be interpreted as a lower bound due to attenuation caused by the union-status measurement error.

3.3 Taxes

All tax variables are imputed in the CPS-ASEC using a Census Bureau created tax model. These variables include federal and state taxes, local property taxes, payments to social security and federal retirement. Also included are different credits such as the earned income tax credit, the child and additional child tax credit, and the Making Work Pay stimulus of 2009-2010 and the federal stimulus payments of 2008. The general approach the Census Bureau uses for imputing taxes is to statistically match CPS tax units to a Statistics of Income (SOI) public use file from the IRS (O'Hara, 2006). State and local taxes follow a similar procedure but includes different parameters as is relevant to specific state tax laws.

Wheaton and Stevens (2016) review different methods for calculating taxes in the CPS-ASEC and find that, on average, the Census Bureau's method produces roughly the same results as those using other tax models. However, no research has looked at whether the choice of tax model results in different results across union membership status. Future research would benefit from looking at the potential bias of tax models across different subgroups.

3.4 Outcomes

The primary outcome of interest is individual net fiscal impact (NFI) on public balance sheets, defined as taxes paid less the cost of public benefits received. The sample average (standard deviation) is \$8,862 (\$14,327) (Table 3.2), suggesting that the average worker pays

$8,862 more in tax liabilities than the value of public benefits and tax credits she collects. In the cross-section, union members average $11,505 in NFI and non-union workers average $8,399 implying a raw $3,106 or 37 percent difference that workers in unions contribute to the public purse over workers not in unions (Table 3.3).

To measure taxes paid by each individual, we add up reported annual federal and state income tax liabilities before credits, property tax, Social Security, and federal retirement plan payroll deductions. Income from tax credits - Earned Income, Make Work Pay, Child, Child Care, and Stimulus - are also included in this sum but enter with negative sign. The sample mean (SD) is $10,290 ($13,030), with union members paying $2,757, or 28 percent, more than non-union workers on average. Table 3.3 contains summary statistics for each component of taxes paid. Those that enter the sum negatively are denoted (-). Federal income tax and Social Security payroll deductions are the largest components.

To measure the public cost of public benefits received, we add up the reported value of benefits received through various programs. Following Bitler and Hoynes (2016), we look at the private-market value of three major public benefits.[6] Namely these are Food Stamps (SNAP), welfare in the form of Temporary Assistance for Needy Families (TANF) (Families with Dependent Children or AFDC prior to welfare reform), and Unemployment Insurance (UI).We further take advantage of the full list of programs for which the Census Bureau collects data following Sherman, Greenstein and Ruffing (2012). Admittedly, these programs are smaller in magnitude and cover a smaller portion of the population. These include the private-market value of supplemental Social Security Income, Medicaid, and Medicare benefits, and of school-lunch, housing, home heating subsidies, post-secondary educational assistance, Social Security, workers compensation, veteran's benefits, and survivor's benefits.[7] These benefits average $1,427 annually. Union members report $349 or 24 percent less in earned public benefits than non-union workers.

In our analysis, private income is the key mechanism by which unionization would affect

[6] Bitler and Hoynes (2016) look at fourth major program: the Earned Income Tax Credit. We include this in taxes paid.

[7] Most of these tax and benefit-income variables are reported by the individual respondent about him or herself individually. However, some of the benefits are supplied at the family-level: public housing, Medicare, Medicaid, food stamps, school lunch, and home heating. To match the individual-level sample-selection criteria and unionization measure, we construct an individual-level measure for each of those benefits. We allocate the total family's cost of the benefit equally to all adults in the family.

taxes paid and public benefits received.[8] To measure private income, we sum income from alimony, farm income, non-farm business income, child support, dividends, interest, rent, retirement, wage and salary income wages, assistance from friends and relatives, and income from other sources. For homeowners, we also include the flow value of housing services so that "income" from both housing and other investments is captured. Focusing on private income, in general, rather than labor income, in particular, makes sense for two reasons. First, nonunion workers may compensate for lower hourly compensation at their primary employer by devoting extra time to other income-generating strategies including self-employment and it does not make sense to ignore the available information on these channels, as these will affect outcomes. Second, if union members enjoy a long-term flow of higher income, this might allow them to accumulate greater assets, which would return additional income in interest, dividends, and the value of housing services, all of which would affect outcomes. The sample average (SD) is \$51,821 (\$36,378) in private income per year, with union members reporting \$10,113 or 20 percent more annual income than nonunion workers. By far, the largest component is wage and salary income with an overall average of \$47,904 and union workers earning \$7,817 or 17 percent more that nonunion workers on average.

3.5 Empirical methodology

To examine whether the raw mean differences by union status hold up in more homogeneous comparisons, we use mean regression analysis. The primary predictor of interest is an indicator of union membership. The excluded category is nonunion workers. Covered nonmembers, who work under a union contract without joining the union, are conservatively categorized as nonunion workers.

To isolate the relationship between outcomes and union status, we condition on other three types of observable determinants of the outcomes. First, we include a standard set of wage determinants ($\mathbf{X}$): potential experience in quartic form, indicators for educational attainment, marital status, race and ethnicity, sex, foreign-born, part-time work, size of

[8] Unionization may affect public balance sheets through the political economy as well, by encouraging political support for higher tax rates and more expansive public benefit programs. This channel is largely outside the scope of the current analysis. The concluding discussion explores this more fully.

metropolitan area, industry, occupation, employment by federal government, by state government, or by local government (private sector omitted) following Bollinger and Hirsch (2006). Second, we include measures of family structure ($\mathbf{F}$) because these govern tax liability and benefit eligibility. In addition to marital status, we condition on the number of adults in family, number of children aged birth to 5 in family, and number of children aged 6 to 18 in family. Table 3.2 presents summary statistics. Third, individuals' tax liabilities and income from public benefits will also depend on states' current economic and policy conditions. These may also be correlated with the likelihood of union membership. To mitigate this possible sources of omitted-variable bias, we include state-year fixed effects ($\mathbf{1_s 1_t}$) in all of our models, ensuring that all comparisons are made between individuals of different union status within the same state-year.

We estimate three models. The first specification is a pooled cross-section, regressing outcomes on an indicator for union membership, individual demographic covariates (all individual wage determinant and family structure variables), and state-year fixed effects. This is the Bollinger and Hirsch specification augmented with family structure and state-year fixed effects:

$$y_{it} = \beta 1(union)_{it} + \gamma_1 \mathbf{F_{it}} + \gamma_\mathbf{2} \mathbf{X_{it}} + \mathbf{1_s 1_t} + \varepsilon_\mathbf{it}. \tag{3.1}$$

In this setting, the identifying assumption is that, comparing across workers in the same state and year and controlling linearly for observed differences in family structure and standard wage determinants, the unobservable determinants of outcomes are not conditionally associated with union membership. β measures the mean difference in outcomes between union workers and otherwise-similar non-union workers.

To tighten the comparison further, we relax the assumption that linear controls are adequate and construct indicators for highly-interacted combinations of control variables. The first set of controls interacts the variables more-closely related to tax liability and benefit eligibility. Specifically, we interact number of kids 0-6, number of kids 6-18, total adults in family, marital status (6 categories: married spouse present, married spouse absent, separated, divorced, Widowed, and never married/single), sex, Hispanic origin, African American, Asian, Foreign-born status, state, and year. That is, we construct 101,249 indicators representing the observed combinations of these variables, denoted $\mathbf{1(F)_{it} 1_s 1_t}$.

In this specification, comparisons are only made between individuals in the same demographic cell-state-year. We also interact wage-determinant variables and denote the set of cell indicators as $\mathbf{1(W)_{it}}$. Specifically, we interact federal public sector, state public sector, local public sector, industry (13 categories), occupation (6 categories), part-time status, metropolitan size (7 categories), potential experience (in 5 year bins for a total of 10 groups), and education (4 categories: less than H.S., H.S or equivalent, some college or Associate's degree, and college degree or more) for a total of 26,545 indicators. Specification 2 is thus:

$$y_{it} = \beta 1(union)_{it} + \gamma_1 \mathbf{1(F)_{it} 1_s 1_t} + \gamma_\mathbf{2} \mathbf{1(W)_{it}} + \varepsilon_\mathbf{it}. \tag{3.2}$$

The third specification recognizes that union and non-union workers may differ in unobservable ways correlated with unionization status and NFI that are not credibly controlled for by cross-sectional comparisons, even with very flexible controls. To address this, we exploit the longitudinal nature of the data to estimate a specification with individual fixed effects. Ideally, this identifies the effect of unionization as the average change in NFI experienced by the workers' who switch between union and nonunion status, conditional on other changes in observables such as educational attainment, family structure, and state-year. More importantly, it allows us to control for unobserved qualities of people who choose to be in unions versus those who are not in unions by largely ignoring people are always union or never union and focusing on changes in outcomes coincident with changes in unionization status holding the worker fixed.

The nature of the outcomes studied here warrant a modification in the specification usually used to study union effects in longitudinal data. Hourly wages and weekly hours, the outcomes usually studied, adjust quickly when a person changes a job and, hence, union status. However, the outcomes studied here are stocks across a year (annual taxes due or benefits received) and, so, the timing of change matters. The union status and conditioning variables are defined at two points in time, twelve months apart. Each outcome is defined with respect to the prior twelve months. Identification comes from seeing how changes within person in union status relates to changes in outcomes reported at those two time periods. Consider someone who switches from non-union to union between the first and second survey. If the person was non-union for the whole year prior to the first survey, switches status immediately after the first survey and stays union for the whole intervening year. The estimated effect would be accurate. However, if the person switched only immediately prior to the second survey, the person would really be non-union the whole time and so though we would have the same measured change in union status. In this case, the estimated effect would be zero. Assuming that the timing of switches is distributed uniformly across the year, switches occur halfway between the first and second survey on average. So, the estimated effect is half of the true effect.[9] For this reason, a change in union status across a 12-month period represents an expected change for half the year. Including a 0.5 constant in the specification corrects for this, effectively doubling the estimate that would otherwise be obtained and letting β express the implicit effect of union status on annual outcomes. This issue does not arise with estimating wage effects because, like union status, wage is defined at a point in time. However, it would apply to any flow variables, including annual hours or annual earnings. Specification 3 gives the individual fixed-effect estimate:

$$\Delta_i(y_{it}) = (0.5)\beta\Delta_i(union_{it}) + \gamma_1\Delta_i(\mathbf{F_{it}}) + \gamma_2\Delta_i(\mathbf{X}_{it}) + \mathbf{1_s 1_t} + \Delta_i(\varepsilon_{it}). \tag{3.3}$$

[9] Ideally, we would measure the share of each year spent in each union status. Ignoring covered non-member status, suppose s_t measures the share of year $t = 1, 2$ a person spends working union in year-t and Y_u is the instantaneous flow of an outcome for each moment spent in union status u. The union effect is $\beta \equiv Y_1 - Y_0$. An observed outcome is $Y_t = Y_0(1 - s_t) + Y_1(s_t)$. Let u_t measure union status at the end of year-t. Our fixed effects analysis relates $\Delta Y \equiv (Y_2 - Y_1)$ to the observable $\Delta u \equiv (u_2 - u_1) \in \{-1, 0, 1\}$ but ΔY really depends on latent $\Delta s \equiv (s_2 - s_1) \in [-1, 1]$. Given persistence in jobs, Δu and Δs should be positively correlated. To take a simple case, if there is no change in the year prior to the first observation ($s_1 = u_1 \in \{0, 1\}$) and there is no more than a single change in u over the intervening year, then the sign of Δs equals the sign of Δu but the magnitude of the change in treatment is overstated: $\Delta s \in [-1, 0) \Leftrightarrow \Delta u = -1, \Delta s = 0 = \Delta$, and $\Delta s \in (0, 1] \Leftrightarrow \Delta u = 1$. An observed ΔY generated by a given true change in treatment Δs but is attributed to a change in measured treatment Δu with larger magnitude. The estimated effect will be attenuated to zero. Suppose that the switch occurs at a random, uniformly-distributed time during the intervening year independent of $(Y_0, Y_1), s \sim U[0, 1]$. Conditional on a change, the average magnitude of change is $E|s_2 - s_1| = 0.5$, although $E|u_2 - u_1| = 1$. Then, $\hat{\beta} = E[\Delta Y/\Delta u] = E[\Delta Y/2\Delta s] = 0.5(E[\Delta Y/\Delta s]) = 0.5\beta$. Are the assumptions of this case plausible? Uniform s is natural. The realism of the assumption that people make no more than one switch in status annually is difficult to evaluate. Just over 90 percent of individuals in the sample have the same status at the start and end of a year, consistent with a high degree of stability in status. Acknowledging that $s_1 = 0.9$ if $u_1 = 1$ and $s_1 = 0.1$ if $u_1 = 0$ would suggest amplifying the cross-sectional and longitudinal estimates by another 25 percent, as 1/(0.9-0.1)=1.25.

3.6 Results

We begin the regression analysis with NFI as the outcome. Specification 1 estimates that union membership is associated with a \$1,290 increase in NFI (Table 3.4: Top panel: Column 1). The controls account for 42 percent of the \$3,106 raw difference in union versus non-union sample means but 58 percent of the difference remains. In specification 2, which includes a much more flexible control set, the estimated association falls by less than 2 percent to \$1,264. Though the standard error increases, from \$92 in specification 1 to \$138 in specification 2 due to the large fall in degrees of freedom from the flexible controls, the association remain significant at the 1 percent level. Specification 3 gives the individual fixed-effect estimate. The estimated effect of union membership on NFI here is \$540, significant at 5 percent.

Next, the NFI result is decomposed between taxes paid and benefits received, as reported in the lower panels of Table 3.4. The logic of the analysis and the specifications used are the same. Only the outcomes differ. Union members pay about \$1,174 more (average for specifications (1) and (2)) in taxes each year, according to the cross-sectional regressions. This result is stable and highly significant statistically across both cross-sectional specifications. The individual fixed effect analysis yields an estimated union-membership effect of \$216 on annual taxes paid, though this is not statistically significant. Union members collect \$102 less (average for specifications (1) and (2)) in public benefits than observably-similar nonunion workers though the results for Specification 2 is not statistically significant. In the panel, the estimated effect is larger: union membership reduces benefit received by \$324 annually and this is significant at the 5-percent level. Whereas cross-sectional analysis suggests NFI effects are driven by more taxes paid, longitudinal analysis suggests a stronger role for reductions in benefits received.

Presumably, union members pay more taxes and collect less public benefits because they have higher incomes from private sources. Do we see evidence of this hypothesized channel in the data? In the cross-sectional analysis, union members earn about \$4,625 (average for specifications (1) and (2)) more than nonunion workers. In the longitudinal analysis, the estimated union effect on income is \$1,614. For this outcome, the fixed effect estimate is statistically significant at 1 percent, despite attenuation issues. Full estimates for all these models are not reported.

Because increased private income is the primary channel through which union membership increases tax payments and reduces public benefit receipt, our specifications for taxes, public benefits and NFI exclude income as a control. Including it would overcontrol (Wooldridge, 2005). This theory provides a testable implication: including private income as a control should soak up much of the estimated "effect" of union status. To investigate this, we run our 3 specifications including polynomials of income as regressors. The logic of this analysis is that the if income is indeed the channel taxes are increased and benefits reduced for union members, then the coefficient on union membership should decrease substantially when private income is included in the regression. As Table 3.5 shows, this is largely the case for NFI and income, but much less so for public-benefit receipt. The first panel reproduces the results of our main specifications for reference. As NFI is a function of taxes paid and benefits received, we begin by looking at the results for taxes paid before turning back to NFI. In all specifications, adding private income linearly reduces the magnitude of the estimated union effect and none of the estimates are significant. Tax schedules are not linear functions of income, so the lower panels show estimated union effect coefficients when higher-order polynomial terms of private income are added to allow more flexibility in the relationship. The estimates remain much smaller than those excluding income, although the cross-sectional specifications yield statistically-significant estimates. In all cases, the fixed effect estimate is not significant. These results support the fact that income is a main driver of higher tax payments, on average, for those in unions.

The results for benefits do not exhibit this pattern entirely. The specification-1 estimates drop in magnitude and significance, consistent with the theory. However, estimates in specifications 2 and 3 do not change substantively with the addition of private income terms, providing some evidence against the theory. Alternatively, it may be a spurious result driven by a long lag in public-benefit changes as income changes. Certainly, individuals have incentives to reduce tax liabilities immediately but to delay loss of benefits as long as possible.

The NFI results follow the tax results and largely confirm the theory. Differences in private income associated with union membership largely explain the association between union membership and NFI that is observed when private income is excluded.

3.7 Heterogeneous effects: sector and education

Union membership may have different effects for public-sector workers than private-sector workers for various reasons. Union membership rates differ dramatically between the sectors. Union members now comprise about 7 percent of private-sector workers but about 35 percent of public-sector workers (Hirsch and Macpherson, 2003).

To examine whether the relationship between union membership and outcomes are stronger in certain subgroups, we generalize specification 1 by interacting all of its coefficients with an indicator for public-sector. Our results (Table 3.6: top panel) show that the effect of union membership is statistically different between public and private sector workers for all outcomes. Union members are estimated to earn \$1,769 more in private income than similar nonunion workers in the public sector. Among workers in the private sector, union members enjoy a much larger advantage, earning \$6,192 more than similar non-union workers. The estimated difference in the union coefficient between sectors is a practically and statistically significant \$4,223. Consistent with this, union membership has a much larger association on taxes paid in the private sector than the public sector. Somewhat surprisingly, the reduction in public-benefits received associated with union membership is larger in the public sector than the private sector. It is statistically significant in each sector. Following the tax result, the positive association of union membership with NFI is larger in the private sector than the public sector.

We also look at different effects among workers with different education levels, in particular workers with a college degree versus those without a college degree. In theory, we would expect union-membership effects to be more pronounced for those with lower levels of education than those with higher education. We find evidence in support of this. Among college graduates, union members have \$2,724 more in private income than similar non-union workers. The union difference is \$5,541 among those with no college degree. Union membership does not relate to public-benefits received among college grads but it does among those without a college degree. The effects are also larger on taxes paid and NFI among those without a college degree than among those with.

3.8 Beyond workers

The propensity to remain employed may differ by union-membership status, which might bias the analysis towards the results we found. Suppose union companies were more likely to go out of business than other companies and, so, throw a higher share of employees into unemployment or out of the labor force, and onto public benefits and into lower tax liabilities. These kinds of workers would fall out of our main sample due to the sample-inclusion requirement that workers be employed in both periods. Unionization would, by this channel, have a negative impact on taxes paid, positive impact on benefits received, and negative impact on NFI but this channel would be hidden from our main analysis. This section develops relevant evidence.

First, we measure the relative propensity of union workers and non-union workers to transition into unemployment and out of the labor force, which is partitioned into either in school (and not working) or idle. Table 3.7 presents transition probabilities of the full sample. Contrary to the concern, transitions into unemployment and out of the labor force, into either idleness or school, are more likely for non-union workers than for union workers. For instance, union workers in one year have a 1.4 percent chance of being unemployed 12 months later. Nonunion workers have a 2.1 percent chance.

Second, we expand the sample to include all people older than 18 and estimate models that add indicators for unemployed, in school, and idle in addition to employed union, leaving employed nonunion as the omitted category. Summary statistics for all outcomes and predictors by status are presented in Appendix Table B.3. The regression results, presented in Table 3.8, corroborate our main findings and are, in most cases, stronger. The estimated coefficient of union membership on NFI is about $1,534 in specification 2 and $976 in the individual fixed effect model, higher than the original sample ($1,300 and $540, respectively). Estimated union effects on taxes paid are higher in this sample as well and here all are statistically significant. In the original sample, the individual fixed effect estimate was about $200 but not statistically significant. In this extended sample, the individual fixed effect estimate is nearly double ($400). Estimated effects on benefits received are also nearly twice as much in this extended sample. The individual fixed effect estimate was about -$325 in the main sample but is -$565 in the extended sample. Finally, for private income, we see similar union premiums in our main and extended samples.

The effects for those not working follow expected patterns. As compared to non-union workers, those who are unemployed, idle, or in school have negative NFI, pay less in taxes, receive more public benefits and earn less in private income. These results are robust to all 3 specifications.

3.9 Displaced Worker Survey

One of the advantages of using a linked-CPS sample is the ability to link to the Displaced Worker Survey (DWS). This sample focuses only on individuals who report being displaced from a prior job as a result of a plant or firm closure in the prior three years. Their current outcomes, current employment and union status, and union status at the job from which they were displaced, are observable. With only current outcomes measured, only a cross-sectional model can be estimated. Further, this sample is much smaller, containing only 2,823 workers. Despite these limitations, it offers a different cut at the problem. Specifically, we run:

$$y_{it} = \beta 1(union)_{it} + \beta_2 1(union\ in\ prior\ job)_{it} + \gamma_1 \mathbf{F_{it}} + \gamma_\mathbf{2} \mathbf{X_{it}} + \mathbf{1_s 1_t} + \varepsilon_\mathbf{it}. \tag{3.4}$$

This is similar to specification 1, except it adds a control for union status in a prior job and is limited to individuals who found a job in the prior 3 years after being laid off for reasons outside their control. For these reasons, the identifying assumption that unobservable influences are conditionally mean independent of current union status may be more credible here than in the main analysis. As Table 3.9 shows, the coefficients all have the same sign as in specification 1 of the main analysis and are all larger in magnitude. The sample size is almost 100 times smaller and the standard errors are much larger. Estimates on NFI, taxes paid, and private income are all still statistically significant but that on benefits received is not. Additional robustness analysis is discussed in the appendix.

3.10 Social Security

One of the tax components included in our analysis is payroll taxes (i.e. Federal Insurance Contributions Act (FICA)). However, this particular tax has a delayed benefit and thus poses a particular problem for this analysis. In particular, the amount of benefit received in retirement age is generally a function of the wages earned but this function is not a one-to-one relationship. Even though union workers tend to earn higher wages, the net value of retirement contributions and retirement benefits received is uncertain. Therefore, as it relates to the present study, it is unclear whether payroll taxes bias taxes paid upward. Ideally, we would have a net present value measure of future retirement benefits received and this value would be included in the benefits received.

As a simple test, we exclude FICA contributions from taxes paid and re-run the main set of regressions. Table 3.10 presents the results. The coefficients all have the same sign as in the main analysis but the change in NFI (individual fixed effect model) is no longer statistically significant at conventional levels. Moreover, the magnitudes of the coefficients for NFI and taxes are reduced. The coefficients of benefits received and private income do not change. Future research will more carefully consider the role of current FICA contributions and future benefit receipts.

3.11 Discussion

The analysis provides strong, though not completely robust, evidence that, at the employee level, union membership has a large, positive effect on net fiscal impact. Union members appear to pay more every year in federal, state, and local taxes than do similar non-union workers, which is connected to the fact that they earn thousands more dollars in annual private income on average. Furthermore, union members appear to receive less in public benefits. Aggregating across NFI components and measuring NFI at the individual level, we observe that union members contributed on average $1,300 more per year to the public balance sheet than similar non-union workers. The fixed-effect estimate is smaller in magnitude, union membership causes an additional $540 more per year in NFI, but points to the same substantive conclusion. If one accepts the conditions laid out in Freeman (1984), we have bounded a causal NFI estimate between these. Though the prevalence of unionization is declining, this evidence suggests that nearly 15 million American union members are contributing an average of between $900 and $1,300 more annually to the public balance sheet than they would otherwise be. To our knowledge, this is the first analysis focusing on or quantifying this effect of unions.

This worker-level analysis ignores other channels by which union membership might affect NFI. However, available evidence allows discussion and approximation of some other, potentially-important channels. To achieve a full accounting of the net fiscal impact of unionization, one must understand from where the higher, private compensation of union members derives. Lee and Mas (2012) estimate that unionization reduces firm equity by 10 percent, implying a 10 percent reduction in the stream of future profits or stream of payments to equity owners. As Lee and Mas (2012) discuss, this 10 percent reduction is composed of two parts: a change in the overall size of the pie and a change in the way the pie is split. The former is the reduction in organizational productivity (p). The latter is the change in labor's share of surplus (s). A 10 percent reduction in profits is consistent with any combination such that $-p - s$ = -10. Lee and Mas (2012) assume that unionization triggers an 8 percent wage premium for labor (s=8) and a negative 2 percent impact on productivity (p=2). However, their data is consistent with other (p, s) combinations. Consider the implications of these two channels separately.

For any given level of p, consider an increase in labor's share (s). Organizations are assemblages of workers and capital aimed at producing value. After consumer surplus is deducted and suppliers are paid, the enterprise's surplus must be divided among labor and capital. For a given level of productivity, unions shift the distribution of an organization's surplus towards workers and away from investors.[10] So, the overall net fiscal impact should account for the fact that each extra dollar in union members' earnings coming through this channel implies a dollar less in shareholder earnings. The question becomes what is the difference between the NFI of the marginal dollar in workers' pockets compared to the NFI of the marginal dollar in investors' and managers' pockets.

First, the effects of unionization on worker taxes paid and benefits received should be offset by changes in associated impacts among firm owners. For a back-of-the-envelope estimate, we turn to estimates of marginal effective tax rates and compliance rates. The marginal federal tax rate on capital income from large C-Corporations businesses is 35 percent. It seems reasonable to assume that the cost of public benefits used by shareholders will not be affected, as ownership of companies is concentrated among those unlikely to be on social safety programs. The average effective marginal federal tax rates on low- and moderate-income workers' income is 31 percent, including changes in both taxes paid and

[10] Our analysis accounts for effects of unionization via differences in the distribution of wage and salary income among employees.

benefits received (U.S. Congressional Budget Office, 2014).[11] These effective marginal tax rates should be adjusted for differential noncompliance. Only 1 percent of labor income is lost to noncompliance, while approximately 10 percent of business and corporate income goes untaxed due to noncompliance (U.S. Internal Revenue Service, 2012). It is worth noting that our estimates derived from the microdata over the study period are very consistent with this 31 percent estimates. In the fixed effect estimates (Table 3.2: specification 3), unionization caused a \$1,614 increase in private income and a \$540 increase in NFI, suggesting a 33 percent effective marginal rate. The cross-sectional estimate in specification 2 suggests a 28 percent rate. Marginal tax federal revenue approximately equal from both sources: marginal revenue from labor income $= 0.99*0.31 = 0.307 \approx 0.90*0.350 = 0.315 =$ marginal revenue from capital income. From these calculations, to the extent that unionization affects only income distribution within the firm, the net fiscal impact on workers of unionization appears approximately fully offset by reduced taxes paid by firm owners. The pie-splitting channel appears to be a wash.

Now, hold share fixed and consider the case where unionization changes productivity. Unionization may cause some ceteris paribus boost to labor productivity (Freeman and Medoff, 1984; Sojourner et al., 2015). On the other hand, if it lowered productivity, this would generate real economic cost with negative fiscal impact through many channels. Changes in on-the-job productivity are only partly reflected in the analysis. Changes in productivity that affect workers' earnings holding employer fixed are reflected.

Our comparisons between similar individuals in the same state-year considers only channels involving labor-management bargaining that changes the creation and distribution of value within organizations. However, unions have fiscal impacts through policy channels as well. For instance, organized labor often advocates for larger public budgets, higher tax rates on higher-income individuals and corporations, and more generous social safety nets. In addition to influence exerted through political action, working-class legislators have different policy preferences than other legislators (Carnes, 2012, 2013) and unionization increases the likelihood of working people holding elected office (Sojourner, 2013). Brady, Baker and Finnigan (2013) provide that states with higher levels of unionization have more generous public-benefit programs for the working poor and lower rates of working poor.

[11] Frandsen (2012) finds little effect of unionization on earnings above the 20th percentile of earnings.

Table 3.1: Union-status transition joint frequencies and probabilities

	Status in year $t+1$	
Status in year t	Union member	Non-Union
Union member	14,185	3,742
	79.1%	20.9%
Non-Union	3,986	99,040
	3.9%	96.1%

Source: CPS-ASEC 1994-2015, CPS-ORG 1994-2015. Notes: Cell frequencies and row percentages reported. Sample includes the 120,953 individuals linked across 2 outgoing rotation groups and meeting sample inclusion criteria in both waves: employed, non-student, wage and salary workers aged at least 18 years with non-imputed union status and no missing covariate or outcome data.

Table 3.2: Summary statistics for longitudinally-linked sample

	Mean	S.D.	Min	Max
Outcomes				
Net fiscal impact	$8,862.42	$14,326.50	-$46,140.16	$71,915.36
Taxes paid	$10,289.70	$13,030.14	-$9,295.38	$71,915.36
Benefits received	$1,427.28	$5,477.84	$0	$38,212.23
Private income	$51,821.23	$36,378.42	-$17,434.94	$206,800.70
Treatment				
1(union member)	0.15		0.00	1.00
Selected demographics[a]				
Number adults in family	2.08	0.91	1.00	12.00
Number of children 0-5	0.26	0.58	0.00	5.00
Number of children 6-18	0.60	0.93	0.00	11.00
Potential experience, years	22.81	11.97	0.00	76.50
Percent Married[b]	63.6%		0.00	1.00
Percent H.S. Degree or Equiv.[b]	32.5%		0.00	1.00
Percent College Degree or More[b]	29.1%		0.00	1.00
Part-Time Worker	12.3%		0.00	1.00
Public Sector - Federal	3.1%		0.00	1.00
Public Sector - Local	10.8%		0.00	1.00
Public Sector - State	5.5%		0.00	1.00

Source: CPS-ASEC 1994-2015, CPS-ORG 1994-2015. Notes: (a) set of demographic controls also includes indicators of gender (2), race-ethnicity (4), foreign-born, metropolitan size (7), industry (13), occupation (7). Sample includes 241,906 observations of 120,953 individuals employed over 2 consecutive years each without missing variables or imputed union status. (b) Controls for marital status include 6 groups and educational status include 4 groups. All means are weighted using sample weights and all dollar amounts are inflated to 2015 dollars.

Table 3.3: Summary statistics for variables and underlying components in full sample, union subsample and non-union subsample

Sample:	All	Union	Non-union
Net fiscal impact	8862.4 (14326.5)	11505 (13304.0)	8398.5 (14448.8)
Taxes paid	10289.7 (13030.1)	12635.2 (12358.8)	9877.9 (13101.2)
Federal income tax liability before credits	5540.3 (13378.7)	6446.4 (13357.5)	5381.2 (13376.1)
State income tax liability before credits	1605.1 (3918.9)	2051.3 (3831.6)	1526.8 (3928.8)
Annual property taxes	1003.6 (2138.0)	1251.9 (2208.8)	960.1 (2122.4)
Social security retirement payroll deduction	3155.7 (2353.1)	3451.6 (2290.4)	3103.7 (2360.1)
Federal retirement payroll deduction	141.6 (1036.5)	315.3 (1333.0)	111.1 (972.0)
Earned income tax credit (-)	235.6 (888.9)	115.1 (603.9)	256.8 (928.4)
Additional child tax credit (-)	53.66 (329.3)	29.44 (251.4)	57.91 (340.9)
Child tax credit (-)	154.3 (549.1)	183 (608.9)	149.2 (537.7)
Credit received from making work pay (-)	45.14 (170.6)	43.68 (170.0)	45.40 (170.7)
Federal stimulus payment (-)	38.73 (221.4)	39.37 (230.0)	38.62 (219.9)
Income from public benefits	1427.3 (5477.8)	1130.2 (4922.3)	1479.4 (5568.0)
Supplemental Security Income (SSI)	18.27 (421.1)	11.64 (369.6)	19.44 (429.5)
Welfare (public assistance)	14 (343.3)	5.99 (205.4)	15.40 (362.2)
Person market value of Medicare	1029.3 (9588.6)	895.8 (9169.3)	1052.8 (9660.1)
Person market value of Medicaid	1028.8	891.4	1053.0

	(9580.6)	(9166.9)	(9651.2)
Person value of food stamps	63.97 (464.0)	24.97 (264.4)	70.82 (490.4)
Person value of housing subsidy	2.39 (22.6)	1.627 (17.6)	2.524 (23.36)
Person value of school-lunch subsidy	57.08 (184.8)	41.22 (145.6)	59.87 (190.7)
Person value of energy subsidy	2.863 (40.8)	1.829 (31.5)	3.044 (42.21)
Educational assistance (beyond HS)	80.2 (1014.2)	69.27 (890.5)	82.12 (1034.4)
Social security	343.9 (2389.9)	166.6 (1745.2)	375.1 (2484.6)
Unemployment benefits	104 (1071.3)	135.7 (1183.5)	98.39 (1050.3)
Worker's compensation	35.47 (788.3)	92.28 (1345.0)	25.49 (642.0)
Veteran's benefits	63.25 (1258.6)	69.88 (1171.7)	62.08 (1273.2)
Disability benefits	23.1 (931.2)	35.31 (1265.4)	20.96 (859.2)
Survivor's benefits	84.73 (2393.2)	96.21 (2514.9)	82.71 (2371.2)
Private income	51821.2 (36378.4)	60338.7 (30772.9)	50325.8 (37074.4)
Alimony	18.99 (666.0)	12.81 (484.2)	20.07 (693.0)
Non-farm business income	144.8 (3716.4)	128.2 (3146.2)	147.7 (3807.8)
Child support	164.7 (1406.7)	154.8 (1356.1)	166.4 (1415.3)
Dividends	278.4 (2555.9)	235.9 (2011.9)	285.9 (2639.8)
Farm	20.98 (1296.3)	19.69 (1428.0)	21.20 (1271.8)
Interest	385.3 (2714.9)	400.9 (2400.2)	382.5 (2766.5)
Income from other source not specified	21.23 (706.7)	25.3 (614.5)	20.52 (721.7)

Rent	188.2 (3103.7)	223.1 (3499.0)	182.1 (3028.9)
Retirement	395.2 (3873.7)	287.8 (3179.9)	414.1 (3982.8)
Wage and salary income	47903.7 (47113.7)	54553.7 (35829.1)	46736.1 (48733.1)
Assistance from friends/relatives not in HH	23.69 (720.8)	20.96 (622.0)	24.17 (736.8)
Implied value of owner-occupied housing	4047.9 (5386.7)	5030.9 (5988.4)	3875.3 (5255.0)
Observations (individual-year)	241,906	36,098	205,808

Source: CPS-ASEC 1994-2015, CPS-ORG 1994-2015. Notes: Standard deviations presented in parentheses. All means are weighted using sample weights and all dollar amounts are inflated to 2015 dollars.

Table 3.4: Estimates of conditional association of union-membership on four outcomes using longitudinally-matched observations and various sets of conditioning variables

Specification:	1	2	3
	Outcome: net fiscal impact		
1(union member)	1289.8***	1264.3***	540.0**
	(91.5)	(138.1)	(254.4)
	Outcome: taxes paid		
1(union member)	1108.6***	1240.2***	216.3
	(85.6)	(129.7)	(208.5)
	Outcome: public benefits received		
1(union member)	-181.2***	-24.1	-323.7**
	(35.0)	(47.1)	(144.9)
	Outcome: private income earned		
1(union member)	4661.6***	4588.0***	1614.0***
	(205.3)	(302.3)	(575.1)
Demographics	Yes	Yes	Yes
State-year FE	Yes	Yes	Yes
Individual FE			Yes

Source: CPS-ASEC 1994-2015, CPS-ORG 1994-2015. Notes: Coefficient (within-individual, correlation-corrected SE). Significant at: *10 **5 ***1 percent level. 241,906 observations of 120,953 individuals over 2 consecutive years each. Coefficient estimates on 1(union member) are presented for each {outcome}x{specification} regression model. All regressions are weighted using sample weights and all dollar amounts are inflated to 2015 dollars.

Table 3.5: Estimated union-membership coefficients when controlling for various functions of private income

Specification:	1	2	3	1	2	3	1	2	3
	Net fiscal impact			Taxes paid			Benefits received		
				Base Model					
1(union member)	1289.8***	1264.3***	540.0**	1108.6***	1240.2***	216.3	-181.3***	-24.2	-323.7**
	(91.6)	(138.1)	(254.4)	(85.6)	(129.7)	(208.5)	(35)	(47.1)	(144.9)
				Base Model + Income					
1(union member)	71.5	91.9	143.7	-73.5	97.6	-163.7	-145.0***	5.7	-307.4**
	(69.8)	(113)	(209)	(61.5)	(101.5)	(155.6)	(35)	(47.2)	(144.8)
				Base Model + Income + Income^2					
1(union member)	219.1***	184.4*	171.7	166.6***	252.7**	-122.8	-52.5	68.3	-294.5**
	(70.3)	(111.7)	(208.8)	(62)	(99.9)	(154.8)	(34.9)	(47.4)	(144.6)
				Base Model + Income + Income^2 + Income^3					
1(union member)	237.8***	200.4*	178.7	204.7***	280.4***	-110.4	-33.1	80.0*	-289.0**
	(70.4)	(111.7)	(208.8)	(62)	(99.4)	(154.6)	(34.9)	(47.1)	(144.4)
				Base Model + Income + Income^2 + Income^3 + Income^4					
1(union member)	241.9***	201.4*	182.4	211.3***	260.9***	-105.6	-30.7	59.6	-288.1**
	(70.4)	(111.7)	(208.7)	(62)	(99.6)	(154.3)	(34.9)	(46.7)	(144.4)

Source: CPS-ASEC 1994-2015, CPS-ORG 1994-2015. Notes: These models use the base models and three specifications from Table 3.4 but add progressively adds controls for different functional forms of income. Coefficient (within-individual, correlation-corrected SE). Significant at: *10 **5 ***1 percent level. 241,906 observations of 120,953 individuals over 2 consecutive years each. Coefficient estimates on 1(union member) are presented for each {outcome}x{specification} regression model. All regressions are weighted using sample weights and all dollar amounts are inflated to 2015 dollars.

Table 3.6: Estimated coefficients by selected subsamples for Specification 1

Outcome:	Net fiscal impact		Taxes paid		Benefits received		Private income	
	Public	Private	Public	Private	Public	Private	Public	Private
1(union member)	510.7*** (157.9)	1717.1*** (114.1)	210.7 (148.9)	1619.6*** (104.2)	-300.0*** (56.4)	-97.5** (43.4)	1769.4*** (323.2)	6192.0*** (292.1)
Difference	-1206.5***		-1408.9***		-202.5***		-4422.5***	
N	49,403	192,503	49,403	192,503	49,403	192,503	49,403	192,503
	College	No college	College	No college	College	No college	College	No college
1(union member)	741.4*** (257.6)	1575.5*** (98.2)	713.8*** (244.2)	1358.7*** (90.8)	-27.7 (90.1)	-216.8*** (38.6)	2723.6*** (479.9)	5541.1*** (238)
Difference	-834.1***		-645.0**		189.1*		-2817.5***	
N	47,283	194,623	47,283	194,623	47,283	194,623	47,283	194,623

Source: CPS-ASEC 1994-2015, CPS-ORG 1994-2015. Notes: These models use the base models from Table 3.4 but run the regressions for the particular subsamples noted. Coefficient (within-individual, correlation-corrected SE). Coefficients and differences between the subsample coefficients 1(union member) significant at: *10 **5 ***1 percent level. 241,906 observations of 120,953 individuals over 2 consecutive years each. Coefficient estimates on 1(union member) are presented for each {outcome}x{specification} regression model. All regressions are weighted using sample weights and all dollar amounts are inflated to 2015 dollars.

Table 3.7: Transitions including unemployed, idle, and in school

	Status in t+1					
Status in t	Non-Union	Union	Unemp.	Idle	School	Total
Non-Union	122,191	4,203	2,898	4,493	727	134,512
%	90.84	3.12	2.15	3.34	0.54	100
Union	3,925	14,327	272	446	19	18,989
%	20.67	75.45	1.43	2.35	0.1	100
Unemp.	2,840	255	1,851	1,341	93	6,380
%	44.51	4	29.01	21.02	1.46	100
Idle	3,762	225	931	73,237	337	78,492
%	4.79	0.29	1.19	93.31	0.43	100
In School	1,334	40	212	619	2,904	5,109
%	26.11	0.78	4.15	12.12	56.84	100

Source: CPS-ASEC 1994-2015, CPS-ORG 1994-2015. Notes: Cell frequencies and row percentages reported. Sample includes the 243,482 individuals linked across 2 outgoing rotation groups and meeting sample inclusion criteria in both waves of being at least 18 years with non-imputed union status and no missing covariate or outcome data.

Table 3.8: Results including idle, unemployed and in school

Specification:	1	2	3
	Outcome: net fiscal impact		
1(union member)	2006.6***	1534.1***	976.4***
	(108.8)	(134.5)	(292.4)
1(unemployed)	-5533.2***	-5935.9***	-745.8**
	(189.8)	(245.5)	(333.6)
1(idle)	-6373.6***	-4338.5***	-1155.2***
	(192.8)	(237.1)	(397.8)
1(in high school or college)	-469.2**	-3831.6***	-928.2*
	(230.1)	(307.9)	(524.8)
	Outcome: taxes paid		
1(union member)	1513.9***	1326.1***	411.8**
	(83.5)	(108.3)	(187.8)
1(unemployed)	-3282.7***	-3586.4***	-413.6**
	(106.3)	(156.1)	(196.8)
1(idle)	-2473.3***	-1876.7***	-214.6
	(110.8)	(174.8)	(219.8)
1(in high school or college)	-1026.6***	-2409.7***	-998.0***
	(122.9)	(180.9)	(219.6)
	Outcome: public benefits received		
1(union member)	-492.7***	-208.0***	-564.6**
	(62.9)	(70.9)	(220.2)
1(unemployed)	2250.5***	2349.5***	332.2
	(131.8)	(169.7)	(261.4)
1(idle)	3900.3***	2461.9***	940.6***
	(145.6)	(164.0)	(335.0)
1(in high school or college)	-557.4***	1421.9***	-69.8
	(179.1)	(249.9)	(480.8)
	Outcome: private income earned		
1(union member)	6541.9***	5571.0***	2094.4***
	(198.1)	(244.6)	(534.8)
1(unemployed)	-17644.6***	-18258.9***	-2638.0***
	(403.3)	(540.5)	(625.2)
1(idle)	-14093.6***	-10858.0***	-2130.4***

	(338.8)	(474.8)	(711.6)
1(in high school or college)	-7514.4***	-14207.1***	-7112.6***
	(385.3)	(488.0)	(762.4)
Demographics	Yes	Yes	Yes
State-year FE	Yes	Yes	Yes
Individual FE			Yes

Source: CPS-ASEC 1994-2015, CPS-ORG 1994-2015. Notes: Coefficient (within-individual, correlation-corrected SE). Significant at: *10 **5 ***1 percent level. 486,964 observations of 243,482 individuals over 2 consecutive years each. Coefficient estimates on 1(union member) are presented for each {outcome}x{specification} regression model. All regressions are weighted using sample weights and all dollar amounts are inflated to 2015 dollars.

Table 3.9: Results using Displaced Worker Survey supplement of CPS

Outcome:	Net Fiscal Impact	Taxes Paid	Benefits Received	Private Income
Current Union Member	2939.4**	2646.5**	-293.0	8028.6**
	(1343.1)	(1208.2)	(575.2)	(3119.2)
Prior Union Member	701.5	448.4	-253.1	1744.1
	(1194.4)	(1056.7)	(541.7)	(2637.7)

Source: CPS-Displaced Worker Survey (1996,1998, 2000, 2002, 2004, 2006, 2008, 2010, 2012, 2014), CPS-ORG 1994-2015. Notes: Coefficient (within-individual, correlation-corrected SE). Significant at: *10 **5 ***1 percent level. 5,646 observations of 2,823 individuals over 2 consecutive years each. Coefficient estimates on current union member and prior union membership are presented for each outcome. Controls include family and individual characteristics as well as state-year fixed effects. All regressions are weighted using sample weights and all dollar amounts are inflated to 2015 dollars.

Table 3.10: Results excluding FICA contributions

Specification:	1	2	3
	Outcome: net fiscal impact		
1(union member)	930.9***	905.6***	364.8
	(84.6)	(129.4)	(238.0)
	Outcome: taxes paid		
1(union member)	749.7***	881.4***	41.0
	(78.0)	(120.1)	(189.0)
	Outcome: public benefits received		
1(union member)	-181.2***	-24.1	-323.7**
	(35.0)	(47.1)	(144.9)
	Outcome: private income earned		
1(union member)	4661.6***	4588.0***	1614.0***
	(205.3)	(302.3)	(575.1)
Demographics	Yes	Yes	Yes
State-year FE	Yes	Yes	Yes
Individual FE			Yes

Source: CPS-ASEC 1994-2015, CPS-ORG 1994-2015. Notes: Coefficient (within-individual, correlation-corrected SE). Significant at: *10 **5 ***1 percent level. 241,906 observations of 120,953 individuals over 2 consecutive years each. Coefficient estimates on 1(union member) are presented for each {outcome}x{specification} regression model. All regressions are weighted using sample weights and all dollar amounts are inflated to 2015 dollars.

Chapter 4

Assessing the effect of E-Verify mandates on employment

4.1 Introduction

Immigration reform is a timely and important policy issue in the United States. As of 2015, according to the best estimates available, there were slightly over 11 million unauthorized immigrants living in the U.S. which represents a significant increase relative to the estimated 3 million unauthorized immigrants in the early 1990s (Warren, 2017; Warren and Warren, 2013; Passel, Cohn and Gonzalez-Barrera, 2013).[1] To date, policymakers continue to push comprehensive immigration reform to curtail the number of unauthorized immigrants. The electronic employment verification system, known as E-Verify, is widely considered an important component of immigration reform. As immigration reform continues to be debated at a national level, any law that requires mandatory employment verification would only increase the use of E-Verify.

Lacking federal action, states have passed laws requiring the use of E-Verify to certain employers, some to all, and some states have even made it illegal to use E-Verify at all. This state-by-state variation in the implementation of E-Verify has been used to analyze the policy's effect on various outcomes. Indeed, the variation allows the direct test of its

[1] The unauthorized immigrant population has declined since 2008 when unauthorized immigrants totaled about 12 million.

effectiveness since the explicit goal of the E-verify program is to "reduce unauthorized employment without undue burden on employers or contributing to discrimination" (Westat, 2009). Therefore, the first-order effect of the policy is that of the employment of unauthorized immigrants. Do these mandates reduce the employment levels of unauthorized immigrants? An answer in the affirmative would be evidence of the policy's effectiveness. A second-order effect is how the employment levels of authorized immigrants and native-born workers change as a result of these mandates. An indirect effect studied in this paper is how the E-Verify mandates affect the population of unauthorized immigrants. Stricter immigration enforcement policies should decrease the size of the unauthorized population. Since the population of unauthorized immigrants has been falling since 2008 and E-Verify mandates have been increasing concurrently, it is a natural extension to analyze the role of these mandates on the changes in unauthorized immigrant populations.

The growing literature exploiting this variation highlights an important methodological issue. Economic research related to unauthorized immigrants relies on some representative definition for unauthorized immigrants. The most popular definition researchers have used is that of "Hispanic, non-citizen, young working-age (under 45) individuals with a high school education or less" citing the work of Passel and Cohn (2010), arguing that this a "group workers previously shown to be a very good representation of the most likely unauthorized" (Amuedo Dorantes and Bansak, 2014). A central question of this paper is: How representative is this group for unauthorized immigrants? That is, does this particular group of Hispanics represent a random sample of unauthorized immigrants in the U.S.? The definition fits well into the common media portrayal of who is unauthorized but, as the data will show, does not capture the full distribution of unauthorized immigrants' educational, country of birth, or geographical choice of living. The representativeness of this definition has two direct practical implications. On one hand, showing that the misclassification error associated with this definition is negligible will highlight the need to make available data on likely unauthorized immigrants to the broader research community. On the other hand, finding evidence supporting the use of this definition gives more substantial evidence for continuing to do so. As a benchmark, this paper uses data from the "logical edit" method of identifying likely unauthorized immigrants in microdata. While this method is also a proxy measure for unauthorized immigrants, it is widely considered the most reliable data on unauthorized immigrants in the U.S.

This paper adds to the literature in two ways. First, this paper directly assesses the representativeness of working-age Hispanics with low-levels of education relative to the "logical edit" method. Second, given that existing papers use data up to December 2012 and half of the enactment dates for universal E-Verify mandates are in 2012 (four of eight states), this paper uses data through 2014 to account for any potential lag in the effects of E-Verify on employment. This paper proceeds as follows. First, the paper gives an overview of the E-Verify program and the research related to its impact of various outcomes. Second, the paper provides a general overview of the different methods for identify likely unauthorized immigrants. Third, the paper presents descriptive evidence on the differences between the typical definitions of unauthorized immigrants and the full population of unauthorized immigrants. Fourth, the paper gives direct evidence on the representativeness of typical definitions for unauthorized immigrants by analyzing the effect of E-Verify on population size. Fifth, the paper focuses on the role of E-Verify on employment. The paper concludes with a discussion of the results and proposes various practical options for researchers studying unauthorized immigrants.

4.2 E-Verify

As the Department of Homeland Security explains "E-Verify is an Internet-based system that compares information from an employee's Form I-9, Employment Eligibility Verification, to data from U.S. Department of Homeland Security and Social Security Administration records to confirm employment eligibility." Employers can voluntary enroll in the E-Verify program but, unless the employer conducts their business in a state with an E-Verify mandate, there is no requisite that the employer do so. In 2017, E-Verify was used nationwide by more than 700,000 employers of all sizes and was joined by about 1,400 new participating companies every week. Since E-Verify was launched at a nationwide scale, the number of participating employers had grown nearly thirty-fold (from about 25,000 in 2007 to about 700,000 in 2017) (Citizenship and Services, 2014).

Form I-9 is used for verifying the identity and employment authorization of individuals hired for employment in the United States. All U.S. employers must ensure proper completion of Form I-9 for each individual they hire for employment in the United States.

However, an employer does not necessarily need to check whether the information on the I-9 is legitimate unless the employer is a federal contractors or they work in a state that mandates E-Verify use. The key innovation of the E-Verify program is that it allows employers to compare the I-9 information of a new worker via an online platform to information from the Social Security Administration and the Department of Homeland Security. Thus, in theory, an employer can positively identify unauthorized immigrants.

Using the E-Verify program has become simpler and faster for employers to use. Within three days of a newly hired worker starting their job, the employer must take the Form I-9 information and submit the information to the E-Verify webpage, thus creating a "case" for the new employee. Two outcomes are possible. First, which is the case for the vast majority of new employees, the result is "Employment Authorized" and is confirmed within seconds or, in some cases, up to 24 hours. For 98.82 percent of over 8 million cases in fiscal year quarter 1 of 2017 (October 2016 through December 2016), the new employee was automatically confirmed. In these cases, the employer and employee are free to continue business as usual. A second possible outcome is a "Tentative Nonconfirmation" which means that the Form I-9 information does not match the SSA and/or DHS records. In this case, the employee has the option to appeal the result within eight days and must contact the SSA or DHS to resolve the issue. In 2016, 1.18 percent of employees received these "Tentative Nonconfirmations" and only 0.16 percent (or 13.5 percent of all "Tentative Nonconfirmations") were able to reconcile the mismatch and receive their "Employment Authorized". The remaining 1.02 percent (or 86.5 percent of all "Tentative Nonconfirmations") were found to be not work authorized. In these cases, the employer must terminate employment though, in many cases, the worker walks away from the job on their own (Citizenship and Services, 2014).

Table 4.1 lists the 21 states who have enacted legal status verification mandates by their date of adoption, date of implementation and the sectors the mandate affects (see Figure 4.1 for a map of these states). Seven states have enacted universal mandates that require all employers to use E-Verify on all new employees while two states (Colorado and Tennessee) require that employee verify the legal status of new hires but not necessarily through the E-Verify program. Twelve have enacted mandates that affect employers in the public sector or that contract with the state. The implementation dates of more than half of the universal E-Verify mandates begin in 2012. If there is a lag in the effect, the impact of the mandate may not take effect until after 2012. In order to address this lag, when analyzing the effect of E-Verify on employment, data up to December 2014 is used.

4.3 Basic theory and empirical evidence

4.3.1 Theory

From a theoretical standpoint, E-Verify mandates have various labor market effects that vary by the immigration status of workers. In this section, the focus will be on the theoretical effects of E-Verify mandates on employment and wages in states with universal E-Verify mandates. Beginning with unauthorized workers, the most straightforward prediction is that their labor demand should fall. Labor demand decreases as the employers face higher employment costs. Most directly, demand for unauthorized labor declines as a result of better information on the authorization status of employees. Since it is illegal to hire unauthorized workers, it follows that knowing with certainty that a worker is unauthorized will decrease the demand for those workers as employers face an increased likelihood of sanctions/penalties for hiring such workers. Secondly, employment costs increase as a result of tentative nonconfirmations. Recall that employers have three days to start an E-Verify case for a new employee. Furthermore, after these days, for these tentative nonconfirmations, there are an additional eight days for workers to redress the nonconfirmation. As a result, employers who hire unauthorized workers incur these transaction costs in the form of the cost of training the unauthorized worker for the time they were employed and having to find another person to hire.

The labor supply of unauthorized workers should also decline as the E-Verify mandates make it more difficult for unauthorized workers to gain employment. Moreover, if unauthorized workers end up migrating to states with no E-Verify mandates, then the labor supply will decline further. As a result, the employment of unauthorized workers should theoretically fall because of a decline in both labor supply and demand. The effect on wages, however, will depend on the elasticity of supply and demand for unauthorized workers. Assuming that some employers continue to hire unauthorized workers, as a result of noncompliance or fraud, then, all else equal, the decreased labor supply alone of unauthorized workers would result in higher wages for these workers. However, because of the lower demand for these unauthorized workers, the overall wage change will depend on the relative elasticities of the labor demand and supply.

E-Verify mandates also affect the employment and wages of authorized workers. In general, since universal E-Verify mandates affect all employers, the costs associated with E-Verify use could drop the demand for all workers. This drop in demand would likely be most pronounced in sectors that employ higher shares of unauthorized immigrants since these are likely to be the employers that face higher numbers of "tentative nonconfirmations."[2] Other subgroups of workers will be affected differently. For those workers who are near substitutes for unauthorized workers, there should be an increase in demand and, since the labor supply should remain largely unaffected, their employment and wages should rise as well. However, an unintended consequence of the higher costs associated with "tentative nonconfirmations" is that employers may begin to discriminate against those they perceive as likely to be unauthorized. Thus, the increase in employment and wages for this substitute group of authorized workers may be tempered by this discrimination effect. Overall employment in states with E-Verify mandates should decrease as the overall labor supply and overall labor demand fall. Once again, the overall wage effect will depend on the elasticities of labor demand and supply. As many of the effects of E-Verify mandates, particularly the effect on wages, are ambiguous, the literature has focused on empirically measuring the impact of E-Verify mandates.

It should be noted that the explicit purpose of E-Verify mandates is to reduce the number of employed unauthorized immigrants in those states. Therefore, empirical evidence showing a decline in the employment of unauthorized immigrants is evidence of the program's and mandates' effectiveness. On the other hand, evidence that the mandates lead to adverse effects for authorized workers undermine the mandates' effectiveness.

[2] Stark and Jakubek (2012) construct a model demonstrating that once sanctions are set high enough for firms hiring unauthorized workers, the firm will shift resource towards verifying the legality of employees. This in turn reduces production efficiency and thus the returns to labor (i.e. wages) of native workers.

4.3.2 Empirical evidence

The literature on E-Verify mandates has grown substantially in the past years, covering not only the effects on employment and wages but also on outcomes as varied as population changes, foreign direct investment (see Amuedo-Dorantes, Bansak and Zebedee (2015)), migratory experiences of individual immigrants (see Amuedo-Dorantes, Puttitanun and Martinez-Donate (2013)), the inflow of native workers (see Good (2012)), and the use of the temporary agricultural program (H2A) (see Henry (2015)). To narrow the focus, this section will review the literature's findings on E-Verify mandates' effect on employment and wages as well the methodology these papers use for identifying unauthorized immigrants. On this latter point, it will be shown that the best way to address the representativeness of unauthorized immigrants is to focus on the effect of E-Verify mandates on the population of unauthorized immigrants.

The literature generally demonstrates mixed results on the employment and wages of unauthorized workers. Two papers have looked at these effects. Amuedo Dorantes and Bansak (2014) find that universal E-Verify mandates reduce the likelihood of employment of likely unauthorized workers while raising the wages of likely unauthorized women. They find no statistically significant effect on the wages of likely unauthorized men. In contrast, Orrenius and Zavodny (2015) find no statistically significant effect on the likelihood of employment of likely unauthorized workers but find that the wages of likely unauthorized men fall as a result of universal E-Verify mandates. Regarding authorized workers, the two papers also demonstrate mixed results. Amuedo Dorantes and Bansak (2014) find no effect on the group of substitute labor (naturalized Hispanic immimgrants) while Orrenius and Zavodny (2015) find that employment rises for this group. These mixed results are inherently the focus of the current paper.

Two potential methodological issues may underlie the reason for this discrepancy.[3] First, though the data for both papers come from the Basic Monthly Current Population Surveys, the timeframes span different years. Amuedo Dorantes and Bansak (2014) use data from January 2004 to December 2011 while Orrenius and Zavodny (2015) use the same data set but from January 2002 to December 2012. This different time horizon implies that

[3] The two papers also differ in the empirical methodology. Both use a difference-in-difference approach with slightly different controls.

Amuedo Dorantes and Bansak (2014) include fewer states that passed universal E-Verify laws (only four states have implementation dates before December 2011) and thus the studies do not study the same states. Importantly, though this point may not explain the differing results, it is important to note that both Amuedo Dorantes and Bansak (2014) and Orrenius and Zavodny (2015) have short post-enactment periods for most universal E-Verify mandating states. Amuedo Dorantes and Bansak (2014) do not include half of all universal E-Verify states while Orrenius and Zavodny (2015) include all these states but four of them (Alabama, Georgia, North Carolina and South Carolina) have implementation dates in 2012. The present study will test whether E-Verify mandates are sensitive to the length of post-enactment periods.

Second, the definition of likely unauthorized differs between the two papers.[4] Orrenius and Zavodny (2015) define likely unauthorized as immigrants "who have at most a high school diploma, are from Mexico, and are not naturalized citizens" (p. 952). Amuedo Dorantes and Bansak (2014) use a slightly different definition: Hispanic, non-citizens of working-age (under 45), with a high school education or less. Indeed, the definition of likely unauthorized in the E-Verify literature as a whole is not uniform. In a separate paper, Orrenius and Zavodny (2016) extend their definition to include immigrants from Central America but focus on immigrants age 20-54. Bohn, Lofstrom and Raphael (2015) and Bohn, Lofstrom and Raphael (2013) use the same definition as Amuedo Dorantes and Bansak (2014) (i.e. working age Hispanics with low levels of education).

This methodological issue is a key focus of the present study. Finding a group of immigrants that is most likely to be affected by E-Verify mandates can reveal the effect subgroup of all unauthorized immigrants but may not necessarily be reflective of the overall effect the policy has on all unauthorized immigrants. In other words, are a subgroup of immigrants that are extremely likely to be unauthorized reflective of the entire population of unauthorized immigrants? What effect is being measured given this definition of unauthorized immigrants? Future sections of this paper will go into more depth regarding the definition of likely unauthorized immigrants as well as provide direct evidence of how the definition of unauthorized immigrants matters for policy analysis. Orrenius and Zavodny (2016) look

[4] Pia Orrenius mentioned at the PAA Annual Meeting in San Diego that they updated their definition to include all Hispanics and that this change reconciled the difference between their paper and Amuedo Dorantes and Bansak (2014).

at how universal E-Verify mandates affect the population size of unauthorized immigrants (using the definition of working-age Mexican and Central Americans with low education) and find that the mandates reduce the number of likely unauthorized.[5] Importantly, they use the American Community Survey (ACS) from 2005 through 2014. As will be shown, the ACS provides the most reliable benchmark for comparing definitions of likely unauthorized immigrants.

4.4 Identifying likely unauthorized immigrants

The fundamental problem in identifying unauthorized immigrants in microdata is that there exists no survey that specifically asks the authorization status of an immigrant and that is also nationally-representative and continuously updated. Bachmeier, Van Hook and Bean (2014) highlight seven different publicly available surveys that include measure on legal status but only one of these is designed to be nationally representative. Specifically, the Survey of Income and Program Participation (SIPP) is a panel survey designed to be representative of U.S. household and focuses on household income and program participation. The 2004 panel includes 12 different waves and the second wave includes detailed questions on citizenship. In this set of questions, foreign-born respondents describe their status upon arrival as legal permanent resident, refugee/asylee, non-immigrant status (these are typically tourists or students), or "other." This final "other" category is considered to be the unauthorized immigrant group. The drawback of these data is that the SIPP is not designed for state-level analysis and, more importantly, sample of foreign-born respondents is relatively small (Bachmeier, Van Hook and Bean (2014) cite 9,178 foreign-born persons responding to the arrival status question) and thus makes sub-national analysis for this subgroup even less reliable.[6] Furthermore, while the SIPP includes potentially reliable data on unauthorized immigrants, the public-use files lump the unauthorized immigrant responses with those of refugees.. Thus, unless a researcher gains access to the restricted-use data, the most reliable SIPP responses on unauthorized immigrants are not readily available to researchers.

[5] This finding is in line with Bohn, Lofstrom and Raphael (2013) who found a significant reduction in the proportion of the Arizona population that is foreign-born and in particular, that is Hispanic noncitizen.

[6] The SIPP 2008 panel is representative for the 20 largest states.

The other six surveys that include question on immigration status are not designed to be nationally representative and are also temporally limited. Four of the surveys are limited to smaller geographies and cover only short periods of time (i.e. Immigration and Intergenerational Mobility in Metropolitan Los Angeles (IIMMLA), Immigrant Second Generation in Metropolitan New York (ISGMNY), Multi-City Study of Urban Inequality (MCSUI) and Los Angeles Family and Neighborhood Survey (LAFANS)) (Bachmeier, Van Hook and Bean, 2014). The other two surveys are national but focus on specific subgroups. The National Agricultural Workers Survey (NAWS) is nationally-representative of the agricultural workforce while the National Asian American Survey (NAAS) looks only at Asian Americans (Bachmeier, Van Hook and Bean, 2014). In sum, "surveys that have included measures [of legal status] are limited by the fact they are typically relatively small, regionally targeted, and/or focused on a particular subpopulation of immigrants" (Van Hook et al., 2014).

Furthermore, even with these existing datasets, this is an a priori concern about significant underreporting by unauthorized immigrants. That is, how reliable are self-responses from unauthorized immigrants? The literature on this topic is rather sparse but Bachmeier, Van Hook and Bean (2014) provide evidence supporting the reliability of legal status measures. Using the SIPP and LAFANS, Bachmeier, Van Hook and Bean (2014) analyze the non-response rates of legal status questions and show that these rates compare favorably to other immigration-related questions. Specifically, they show that the non-response to the legal status question is no higher than the non-response rates of year of immigration. A second concern with legal status questions is that their presence alone will reduce the response rates to surveys overall, a so-called "chilling effect". Bachmeier, Van Hook and Bean (2014) also show that the attrition rates from wave to another are no higher for unauthorized immigrants as they are for other subgroups. In all, even with some evidence supporting the reliability of legal status measures, the issue comes full circle in that the SIPP is not designed for sub-national analyses.

4.4.1 The "logical edit" method

Given that legal status questions are not available on the more popular datasets for economic research (i.e. the Current Population Survey, (CPS) the American Community Survey

(ACS), etc.), researchers have relied on different strategies for imputing legal status in these surveys. The most popular of these is the "logical edits" method which "have come to be trusted and widely cited outside of academia" (Van Hook et al., 2014) but the "specific details of the [logical edits] method are not publicly available" and thus are not easily replicated by researchers (Van Hook et al., 2014).[7] The development of the "logical edit" methodology is key to understanding the data used in this paper.

Identifying unauthorized immigrants in microdata first requires knowing the total number of unauthorized immigrants residing in the U.S. Indeed, both the methodology and researchers who developed the methodology are key to the development of the microdata on unauthorized immigrants. The aggregate total was developed by Robert Warren and Jeffrey Passel in 1987, documented in a paper titled "A Count of the Uncountable: Estimates of Undocumented Aliens Counted in the 1980 United States Census" (Warren and Passel, 1987). It has subsequently been used for years by the Department of Homeland Security (DHS) to produce official estimates on the unauthorized population. At its core, measuring the total unauthorized population is based on the difference between the total foreign-born population to the legally resident foreign-born population (Warren and Warren, 2013). The foreign-born population is estimated using larger government survey such as the Decennial Census, the ACS or the CPS.[8] The legal resident population relies on data from DHS on legal permanent residents and non-immigrant residents (such as students, tourists), and data on refugees from the Office of Refugee Resettlement. After other adjustments for mortality, undercount, emigration and removals, the total number of unauthorized immigrants is finalized. Because this methodology relies on the difference between total immigrants and total authorized immigrants, the methodology is known as

[7] These data, however, have never been evaluated for its reliability. Therefore, researchers have relied on statistical imputations that use these aforementioned smaller surveys to impute legal status in larger surveys. In an evaluation of these statistical imputation methods and a pseudo-"logical edits" method, (Van Hook et al., 2014) show that only the statistical imputation methods produced unbiased results when jointly observed with the given dependent variable. The "logical edits" method resulted in biased estimates in all scenarios. While this last result would seem to condemn the use of "logical edits", it is important to note that (Van Hook et al., 2014) were not able to replicate the actual "logical edits" method and thus it remains largely untested. In sum, though there is some indirect evidence against using the "logical edits" method for large datasets such as the CPS and the ACS, the method continues to be the most widely cited. Indeed, the direct evaluation of the "logical edits" method would be a worthwhile endeavor.

[8] There are also various adjustment made for the assumed undercount of the foreign born population. See Warren and Warren (2013) for a complete overview of the methodology.

the "residual method." Today, both Robert Warren and Jeffrey Passel produce separate estimates of the total unauthorized population using similar methodologies.

The development of the "logical edit" microdata stems from the "residual method" and both Passel and Warren have separately produced estimates using similar approaches (see Warren (2014) and Passel, Cohn and Gonzalez-Barrera (2013)). The starting point is the fact that survey data includes foreign-born respondents that are both authorized *and* unauthorized (Warren, 2014). Using either the CPS or the ACS, all foreign-born respondents are selected and then a series of "logical edits" remove those immigrants who are likely *authorized*. There are seven major groups that are considered to be authorized immigrants: 1) all immigrants who arrived before 1980 (as they were all granted amnesty after IRCA in 1986), 2) occupations that require legal status (such as lawyers, judges and police), 3) legal temporary migrants (such as students or tourists), 4) immediate relative of U.S. citizens, 5) those receiving public benefits, 6) immigrants age 60 or older at entry, and 7) being from a refugee country.

After these edits, the resulting number of unauthorized immigrants is higher in aggregate than the "residual method" estimate. To bring down this number, a random selection step is implemented that ties the number of counted likely unauthorized immigrants to a set of population controls that reflects the distribution of unauthorized immigrants by country of origin. A final step adjusts the population weights to account for the lower response rate of unauthorized immigrants to government surveys. The unauthorized immigrants identified here are more aptly called *likely unauthorized immigrants*. Though the totals match the total unauthorized population, it is important to recall that the population controls are tied to the totals from the residual method and, therefore, match by construction. In sum, the methodologies for estimating the total number of unauthorized immigrants, the "residual method", is closely tied to the "logical edit" method of identifying likely unauthorized immigrants in microdata.

To be clear, the "logical edit" method does not capture the true authorization status of an immigrant. Rather, it is a measure that captures those immigrants who a very likely to be unauthorized and, when summed up to the population total, closely reflect the size and composition of published estimates on unauthorized immigrants. Indeed, the "logical edit" is a measure that is, by construction, error-prone. In the remainder of this paper, the "logical edit" will serve as a benchmark for previous results.

4.5 Data

Three different datasets are used in this paper. To analyze the different definitions of unauthorized immigrants, the American Community Survey is used from 2005 through 2015 (Ruggles et al., 2015b). The ACS is the largest nationally-representative survey of U.S households, surveying approximately 1 percent of household or about 3.5 million people each year.[9] Topics in the ACS are rather comprehensive, covering demographic, housing, economic and social characteristics. Because of its breadth of topics and number of people surveyed, the ACS is currently the dataset of choice for the "logical edit" microdata.

The second and third datasets come from the Current Population Survey. Specifically, the paper uses the Basic Monthly Current Population Survey files from January 2002 through December 2014 as well as the 2002 through 2014 Annual Social and Economic Supplement (ASEC) (Ruggles et al., 2015a). The Basic Monthly CPS files are the source of the employment statistics of the U.S. and, as its name suggests, are fielded monthly to about 50 thousand households or 130 thousand people. The survey focuses primarily on the labor force characteristics of respondents and asks a limited number of questions on social and economic topics. In this paper, the Basic Monthly CPS files are used primarily to compare results to previous literature. The ASEC has been conducted since the 1960s and historically included detailed questions on the same sort of issues covered in the ACS. In fact, before 2010, the ASEC was the basis of the "logical edit" method. Both the ACS and ASEC are used to analyze the employment effects of E-Verify mandates and to test the reliability of the definitions for unauthorized immigrants.

This paper uses data from the "logical edit" method developed by Robert Warren.[10] Indeed, the ACS data used in this paper contains the likely unauthorized flag for 2010 through 2015. Since access to the code was granted, this paper extends the code to cover

[9] Beginning in 2005, the ACS replaced the decennial long-form, covering almost all of the topics covered in the long-form. Each Decennial Census looks to enumerate all people living in the U.S. For most people, the Census questionnaire contains a few questions (in 2010, the Census was 10 questions) focusing primarily on age, sex, and race. Between 1970 and 2000, about 15 percent of households received the long-form.

[10] The code for this flag was developed by the author as a consultant. The Center for Migration Studies of New York houses the data and granted use of the data for this project.

back to the ACS 2005. Moreover, the code was then adapted for use in the ASEC, covering back to 2002. Importantly, since a majority of the variables necessary for the imputation of likely unauthorized flags are not available in the Basic Monthly CPS files, the code was not adapted for the Basic Monthly files.[11]

4.6 Assessing the representativeness of unauthorized immigrants

The typical measure of unauthorized immigrants in the literature has been Hispanic, non-citizen, young prime-aged (20-54 years) individuals with a high school education or less. For simplicity, this measure will be referred to as "Hispanics with low-levels of education" or H.L.E. Moreover, the "logical edits" population with be referred to as L.E. and used interchangeably with likely unauthorized immigrants. The H.L.E. has been used because they are a good representation of a subpopulation that is very likely to be unauthorized. The first question this paper looks to answer is: How representative are H.L.E. of all unauthorized immigrants in the U.S.? From published figures, it is clear that H.L.E. misses about one-third or more of the entire unauthorized population. More importantly, as it relates to policies in different states, this distribution of unauthorized immigrants is not necessarily uniform across states. Consider the composition of likely unauthorized immigrants by state. For 2013, in California, the most represented countries of origin for likely unauthorized immigrants was Mexico and El Salvador followed by China and the Philippines. Meanwhile, in Illinois, the most represented country was Poland. In Florida, the most represented country was Haiti (Warren, 2014). The point here is that there are many likely unauthorized immigrants from non-Hispanic countries that are widely represented in each state but are omitted from the H.L.E. sample. To give evidence on this issue, this paper uses the likely unauthorized sample from the "logical edit" method as the representative sample of all unauthorized immigrants in the U.S.

[11] Most recently, Borjas (2017a) and Borjas (2017b) have used similar data. Borjas gained access to two years of likely unauthorized CPS-ASEC data (2012-2013) from the Pew Research Center but not the underlying code. Borjas then reverse engineered the code to extend the data to CPS-ASEC 1994 through 2014 and then the ACS 2005-2014. The present study uses data generated from the actual code that produces the logical edits data for ACS 2010-2015 and extends the underlying code to include earlier years of ACS and the CPS-ASEC. That is, no reverse engineering is needed.

4.6.1 Descriptives on unauthorized immigrants: L.E. v H.L.E.

As can be seen in Figure 4.2, the population size of prime-aged H.L.E.'s has remained rather steady around 7 million immigrants since 2005. The prime-aged "logical edits" population has always been greater, closer to 9 million, but has shown a decrease from 10 million in 2007 to about 9 million in 2015. While the H.L.E. population has always been below the logical edits population, the population size alone does not reveal the reveal the representativeness of H.L.E.'s for the entire unauthorized immigrant population. To be sure, a majority of H.L.E.'s are likely unauthorized immigrants under L.E. definitions but they are not a representative sample of all likely unauthorized immigrants. Figure 4.3 describes the characteristics of the H.L.E. relative to the prime-aged (20-54) likely unauthorized immigrants for 2015 using the ACS.[12] There were close to 9 million prime-aged likely unauthorized immigrants in the U.S. in 2015. Of this population, the H.L.E. sample represents about sixty percent. Moreover, it also includes an extra 1.9 million people who are not likely unauthorized. This extra group is, on average, older (40 years) than the overlap of H.L.E. and the "logical edits" by about 4 years and have spent about seven more years living in the U.S. This pattern points out a concern of misclassification error. A random draw of the HLE sample would lead to choosing an authorized worker one out of every four pulls (about 27 percent).

Moreover, as a matter of representativeness, the H.L.E. leaves out about 3.75 million unauthorized immigrants. While the H.L.E. covers about 80 percent of all Mexican and Central Americans (3.5 million of 4.5 million), there are still one million Mexican and Central Americans not included in the H.L.E.. By definition, the H.L.E. will exclude a large amount of non-Mexican/Central Americans and it is important to consider who these are. Of the 3.75 million immigrants missing from the H.L.E., a majority of these are more highly educated Asians and other Latin Americans. About 1.3 million Asians and nearly one million immigrants from other Latin American countries are excluded from the H.L.E. Indeed, about 2.7 million have high school degrees or more; half of these 2.7 million have college degrees or more. Lastly, the 3.75 million immigrants missing from the H.L.E. tend to have been in the U.S. for only 11 years. Table 4.2 presents statistics on the distribution of

[12] See Appendix Figure C.1 for the descriptives on the entire population including the non-prime aged. The patterns are similar though the inclusion of children adds in more kids with less than H.S. and brings down the average age.

the years lived in the U.S. across the two definitions. 20 percent of unauthorized immigrants have lived in the U.S. between 1 and 5 years whereas only 13 percent of H.L.E.'s have lived in the U.S. for that amount of time. The H.L.E. reflects a higher proportion having lived in the U.S. for at least five years (about 83 percent) while the proportion for unauthorized immigrants is closer to 74 percent. Moreover, because of the inclusion of more immigrants from outside Mexico and Central America, the racial composition clearly differs between H.L.E. and the "logical edits." Similarly, the H.L.E. show a higher proportion of married immigrants relative to the "logical edits" (58 percent v. 52 percent) which results from the H.L.E. sample excluding immigrants arriving in the past year. These new immigrants tend to be single at a higher proportion than those who have established themselves in the U.S. The employment rates across the two definitions are similar.[13]

The implication for policy analysis is important. H.L.E. will likely reflect the effects of policies on lower educated subgroups of unauthorized immigrants. However, is there good reason to believe that the patterns exhibited by these more highly educated immigrants will be in the same direction as H.L.E.? Or will there be a counterbalancing effect? Ultimately, these questions are best answered empirically.

4.6.2 Population change of unauthorized immigrants as a result of E-Verify: L.E. v H.L.E.

In order to demonstrate how well the H.L.E. captures the effect of E-Verify on all unauthorized immigrants, this section replicates the findings of Orrenius and Zavodny (2016). This choice is made because Orrenius and Zavodny (2016) use the ACS. Since the code used to develop the "logical edit" data used in the paper was designed for the ACS, comparisons to Orrenius and Zavodny (2016) are straightforward. The analysis looks at how universal E-Verify mandates affect the total population of unauthorized immigrants. Using a difference-in-difference estimation, Orrenius and Zavodny (2016) estimate the following:

$$\ln Pop_{st} = \alpha + \beta_1 E - Verify_{st} + \beta_2 EconConditions_{st-1} + State_s + Year_t + Trend_{st} + \varepsilon_{st}. \tag{4.1}$$

[13] Appendix Table C.1 shows the summary statistics for the full sample (ACS 2005-2015). The main difference in patterns is that including children brings down the average age, the percent of married immigrants, and the overall education level.

s denotes states and t denotes year. Since the ACS does not include month surveyed, *E-Verify* controls for the fraction of a year that a state had a universal E-Verify mandate in effect. *EconConditions* control for state-level business cycle conditions: real state GDP per capita, unemployment rates, local and state government spending per capita, and the number of housing permits and housing starts. *Trend* is linear state-year time trend to control for any underlying trends particular to a state. Standard errors are clustered at the state level and the data are weighted using the sum of the person weights for each cell in the regression.

The identification of E-Verify mandates comes from comparing the size of unauthorized immigrant population before and after universal E-Verify mandates take effect, controlling for these various state, year, and state-specific time trends. Table 4.3 presents the results of this model using the H.L.E. (which replicates the Orrenius and Zavodny (2016) findings) and also the results using the "logical edits." For completeness, the model is run on the 2005-2014 ACS sample which is the same sample used in Orrenius and Zavodny (2016) but then adds the 2015 ACS to update the results. The model is also run on subsample of unauthorized immigrants by years in the U.S. (specifically, greater than 5 years, between 1 and 5 years, and less than a year).

The 2005-2014 results show that the H.L.E. understates the effect of universal E-Verify mandates on unauthorized population by about 10 percent (-0.060(0.022) v -0.067(0.023)) though the coefficients are not statistically different from each other. That is, on average, the number of unauthorized immigrants falls by about 6.7 percent if the state had a universal E-Verify mandates the entire year. This finding supports the prediction laid out in Orrenius and Zavodny (2016) that their estimates "may reflect the lower bound of the effect of E-Verify laws" (p. 4). However, once the 2015 ACS is added to the sample, the coefficients increase to about -0.075 for both the H.L.E. and "logical edits." There are various implications of this result. First, the model specification may be incorrect and thus results may be unreliable. However, as Orrenius and Zavodny (2016) show, their results are robust to many validation checks and thus this implication seems unlikely. Second, the results potentially contradicts the idea that H.L.E. provide a lower bound for the true effect of E-Verify mandates. As will be shown, the results by year of entry give evidence to why this contradiction may be occurring.

Universal E-Verify mandates reduce the population of unauthorized immigrants under both H.L.E. and "logical edits." However, the driving forces behind this reduction is different for each definition. For both definitions, the largest effects are found in the recent arrival group (1-5 years) though this effect is largest for H.L.E. than for the "logical edits" (-0.226(0.062) v -0.160(0.095)). By adding the 2015 ACS, the coefficient increase by 0.05 to -0.276(0.083) for H.L.E. but only by 0.017 for LE. Therefore, under the H.L.E., the effect of universal E-Verify is larger than for the "logical edits" and this effect increases significantly for H.L.E. but less so for "logical edits." This difference highlights the how the H.L.E. is not representative of the overall unauthorized population. In particular, this pattern is noticed for the non-recent (greater than 5 years) and new immigrants (less than a year). Under the H.L.E., there is a large negative effect of universal E-Verify on new immigrants under H.L.E. but not the "logical edits." On the other hand, there is a small significant and negative effect for non-recent immigrants but no effect under H.L.E.

The results presented in Table 4.3 use only a prime-aged sample (20-54). Relaxing this restriction only for the "logical edits" sample shows that the effect of E-Verify is understated, for the "logical edit" sample, by excluding non-prime aged immigrants. Table 4.4 presents the results of the analysis that includes all immigrants. Specifically, the effect of universal E-Verify mandates increases from -0.074(0.024) (using the ACS 2005-2015) to -0.090(0.021). This result is likely due to the fact that, under the "logical edits", the effect of E-Verify is greater for non-recent and recent immigrants relative to the new immigrants. These groups tend to have a higher proportion of families relative to the recent immigrants who tend to be single at higher proportions. Thus, by including the sample of non-prime aged immigrants, universal E-Verify mandates are shown to have a larger effect on reducing unauthorized immigrant populations. The patterns by subgroup are still noticeable in this estimation in that the non-recent and recent immigrants drive the overall effect for "logical edits" while the H.L.E. are primarily driven by recent and new immigrants.

The implication for policy analysis is that the use of the H.L.E. may, on average, result in estimates that capture the overall direction of E-Verify mandates. However, subgroups of the H.L.E. do not seem to reflect this same pattern. The analysis here shows that, in the case of years in the U.S., the H.L.E. does not accurately reflect the effect of universal E-Verify mandates as compared to the unauthorized immigrants defined by the "logical edits." Overall, the H.L.E. surely captures a population that is likely to be affected by E-Verify mandates but this group is not necessarily reflective of the overall unauthorized immigrant population and thus subgroup analysis with the HLE may not be reflective of those subgroups in the entire unauthorized population.

4.7 E-Verify mandates and employment

Turning to the effect of E-Verify mandates on employment, this section looks at how universal E-Verify mandates affect the probability of employment of unauthorized immigrants as well as other subgroups. To best make comparisons with previous literature, the model used here follows that of Amuedo Dorantes and Bansak (2014). Importantly, Amuedo Dorantes and Bansak (2014) use data up to 2011 but this paper extends the timeframe to 2014. Thus, the following analysis gives insight into how a longer post-treatment period changes the estimated effects of E-Verify as well as how the use of the "logical edits" compares to the H.L.E. As the previous section showed, the H.L.E. may be reflect the overall result in regards to population but it is unclear how the H.L.E. fares with the employment.

The model estimated here is a linear probability model:

$$L_{ist} = \alpha + \beta_1 \textit{Universal E-Verify}_{st} + \beta_2 \textit{Public E-Verify}_{st} + X_{ist}\gamma + \beta_3 U_{st} + \delta_s + \theta_t + \delta_s t + \varepsilon_{ist}, \tag{4.2}$$

where L is a dummy for employment, *Universal E-Verify* is a dummy for a state with a universal E-Verify mandate, *Public E-Verify* is a dummy for a state with a public E-Verify mandate (see Table 4.1). X are individual-level controls including gender, race, age, marital status, number of children, educational attainment and industry fixed effects. U_{st} are monthly state unemployment rates, δ_s are state fixed-effects, θ_t are time fixed-effects and $\delta_s t$ are state-level time trends. The identification strategy in this model comes from the difference between the probability of employment for unauthorized immigrants living in states with universal E-Verify and those unauthorized immigrants living in states with no E-Verify mandates. After controlling for observable characteristics, β_1 captures the difference in the change in the likelihood of employment between observably similar unauthorized immigrants in universal E-Verify states relative to non-E-Verify states. Because the model includes a control for *Public E-Verify*, the non-E-Verify states do not include states with public-sector

E-Verify mandates. A major assumption is made in order to claim that E-Verify mandates cause the change in the likelihood of employment of unauthorized immigrants. Namely, the assumptions is that labor market changes in the states that adopt universal E-Verify mandates did not lead to those states adopting E-Verify mandates. For example, it would be plausible that states with large growth in the unauthorized worker population would be the states most likely to adopt E-Verify mandates. The present study does not delve further into the issue of causality as the focus is on how the choice of definitions for unauthorized immigrants affects the results of such analysis.

In order to use the logical edits, the code used to develop the ACS flags are adapted to the CPS-ASEC 2002-2014. Equation (2) is then estimated for four groups: 1) unauthorized immigrants ("logical edits"), 2) H.L.E.'s, 3) naturalized Hispanics, and 4) US-born non-Hispanics. The model is further broken down by sex. The first set of results use only the in labor force population for each subgroup such that β_1 captures the effect of E-Verify mandates relative only to those who are in the labor force. However, because it is possible for people to drop out of the labor force in response to E-Verify mandates, the model is also run on a sample that includes all people over 18 years of age.

Table 4.5 presents the results of this analysis. Focusing first on the difference between H.L.E. and "logical edits", the results are sensitive to the choice of sample (in labor force versus all people). The in labor force results show that, on average, workers in universal E-Verify states using the "logical edits" see an increase of 5 percentage points (0.050(0.004)) while using H.L.E. shows no statistically significant effect. However, including all people changes the results significantly. The effect of E-Verify mandates under the "logical edits" is now negative and statistically significant (-0.04) while the H.L.E. results are also negative and statistically significant. These results suggest that there could be considerable movement out of the labor force for people in E-Verify states relative to those in non-E-Verify states. Regarding the difference between "logical edits" and H.L.E., since the results depend on the choice of sample, it seems as though the H.L.E. are not necessarily representative of the effect of universal E-Verify mandates on all unauthorized immigrants.

This same inconsistency is noticed for other populations. For naturalized Hispanics, the overall effect of universal E-Verify mandates is positive for the in labor force sample but negative for the full sample. However, for U.S. born non-Hispanics, the effect of universal E-Verify mandates is negative and statistically significant under both samples. The implication of this finding is that universal E-Verify mandates decrease the likelihood of employment of US-born non-Hispanics. This finding is the opposite of that found by Amuedo Dorantes and Bansak (2014), suggesting that the additional years included in this study are important for identifying the effect of E-Verify mandates.[14]

Given that these results depend on whether the sample includes only the in labor force population or all adults, the conclusions of this paper are not cleanly summarized. Using the in labor force sample leads to the following conclusions: Universal E-Verify mandates result in an increase in the likelihood of being employed for unauthorized immigrants and naturalized Hispanics but a decrease in the likelihood of employment for US-born non-Hispanics. The H.L.E. fails to pick up an overall effect in this sample. On the other hand, including people not in the labor force leads to a different picture overall. Universal E-Verify mandates reduce the likelihood of all groups. This effect is largest for naturalized Hispanics who see a reduction of about 16 percentage points in the likelihood of employment as a result of E-Verify mandates. Unauthorized immigrants also experience a reduction in their likelihood of employment (-0.040(0.007)) but this effect is larger for the H.L.E. (-0.103(0.004)). Lastly, US-born non-Hispanics still see a reduction in their likelihood of employment (-0.012(0.001)) as a result of universal E-Verify.

In sum, the analysis presented here highlights that the sensitive of results to the definition of unauthorized immigrants and the sample used for analysis. H.L.E. do not necessarily reflect the effects of E-Verify mandates found with the "logical edits." The longer post-treatment period seems to also be an important factor for the effect of E-Verify mandates. By including more years of data, the results Amuedo Dorantes and Bansak (2014) no longer hold.

[14] Another plausible reason for the difference between this study and Amuedo Dorantes and Bansak (2014) is that different datasets are used. Amuedo Dorantes and Bansak (2014) use Basic Monthly CPS files but this dataset does not allow for the use of the logical edits. Appendix Table C.2 runs the analysis on the Basic Monthly CPS files. The results show that the effect of universal E-Verify mandates are also sensitive to the choice of sample (in labor force v. all people 18 and over).

4.8 Future research and conclusion

Overall, this paper highlights that assessing the effects of policies that target unauthorized immigrants, such as E-Verify, requires a good measure of unauthorized immigrants. While the H.L.E. is a subgroup of the unauthorized population that is very likely to be unauthorized, the H.L.E. is not a representative sample of the entire unauthorized population. In the setting of population changes and universal E-Verify, the H.L.E. and "logical edits" show a reduction in the unauthorized population but subgroup analysis is inconsistent across the definitions. Misclassification error is a likely explanation for these results. When applied to the setting of employment and E-Verify mandates, the results show that the measure of unauthorized immigrants matters. Results are sensitive to the sample chosen (in labor force v. all adults). Given these inconsistent findings, it is clear that the H.L.E. do not accurately the effect of E-Verify mandates on the likelihood of employment. Moreover, the results here suggest that additional years of data are important in capturing the effect of E-Verify mandates.

The main question left to answer is what a researcher should do if data on unauthorized immigrants is not available. One potential solution is to derive the misclassification probabilities of H.L.E.'s relative to unauthorized immigrants as measured through the "logical edits." These probability could then be used to correct estimates using H.L.E. Another issue arises for samples where the "logical edits" are not available. Specifically, in this paper, the Basic Monthly CPS files do not include enough information to replicate the "logical edits." In this scenario, it would be possible to apply as many "logical edits" as possible to the Basic Monthly files and then apply a selection criteria to mimic the distribution of immigrants by country of origin, state of residence, and year of immigration. Lastly, the most obvious option is to make public-use files of the "logical edits" microdata.

On the accurate estimation of E-Verify mandates, a few threats to validity warrant discussion. First, it is likely that states that pass E-Verify mandates also pass other anti-immigration laws that may affect employment rates. Indeed, Arizona's Legal Arizona Workers Act included an E-Verify mandate as one of many measures to curtail unauthorized immigrants' employment. Without explicitly controlling for these other measures, the estimated effects of E-Verify mandates may very well be due to these other factors. Thus, it is necessary to find a way to control for this "policy climate." Various authors have developed a measure that would account for this immigration policy climate. For example, Leerkes, Leach and Bachmeier (2012) conduct factor analysis to code states into three different levels of immigration control: high, moderate and low. Using data on employers participation in E-Verify, restrictive state laws, county and city involvement in the 287(g) program, the authors are able to construct a single measure ("internal control index") for each state by year that is then used to classify each state into the different levels of control. By using these more refined measures by state and year, it would be possible to capture the immigration policy climate of each state.

One last threat to validity is the level of enforcement within each state of the E-Verify mandate. While the E-Verify program can track the number of cases employers process, there is no guarantee that employers are processing all potential hires. The need here is to find a measure that quantifies the enforcement of anti-immigration laws. Fortunately, various authors have conducted these studies and some would be suitable for my study. In particular, Watson (2010) codes information on 287(g) on a year-by-year basis between 1993 and 2002. Using a dataset that "consists of counts of Immigration and Naturalization Services (INS) 'deportable aliens located' as the result of internal investigations, by INS internal district, country of origin, and fiscal year" (Watson, 2010). The correlation between 287(g) enforcement and E-Verify mandates is arguably strong enough for this measure to be a good proxy of enforcement. Thus, using a measure like the one presented by Watson (2010) (but extended through 2014), it would be possible to control for enforcement levels of E-Verify mandates that may be confounding the analysis conducted thus far in this paper.

Table 4.1: Overview of E-Verify mandates

State	Adoption date	Implementation date
Universal E-Verify mandate		
Alabama	June 2011	April 2012
Arizona	June 2007	January 2008
Georgia	May 2011	January 2012
Mississippi	March 2008	July 2008
North Carolina	June 2011	October 2012
South Carolina	June 2011	January 2012
Utah	March 2010	July 2010
Universal check mandate (alternate process[a])		
Colorado	January 2007	August 2008
Tennessee	June 2011	January 2012
Other mandate[b]		
Florida	January 2011	January 2011
Idaho	July 2009	July 2009
Indiana	April 2011	July 2011
Louisiana	July 2011	July 2011
Minnesota	July 2011	July 2011
Missouri	July 2008	January 2009
Nebraska	April 2009	October 2009
Oklahoma	May 2007	November 2007
Pennsylvania	July 2012	January 2013
Texas	December 2014	September 2015
Virginia	March 2011	December 2013
West Virginia	March 2012	June 2012

Source: http://www.troutmansanders.com/immigration/. Notes: (a) These alternate processes mandate that an employer verify the legal status of a newly hired employee but not necessarily through the E-Verify program. (b) Other mandates refer to mandates that are not universal but cover different sectors. Typically these sectors are public employers, public contractors/subcontractors or state agencies.

Table 4.2: Summary statistics for H.L.E. and Logical Edits, prime-aged (20-54) sample

	H.L.E.	Logical Edits		H.L.E.	Logical Edits
Years in U.S.	14.1	10.9	Age	35.4	34.4
	(8.7)	(6.9)		(8.8)	(8.7)
% >5 years	83.3	73.9	% Female	43.1	43.9
	(37.3)	(43.9)		(49.5)	(49.6)
% Between 1 and 5 years	13.3	20.2	% Employed	71.0	73.1
	(34.0)	(40.2)		(45.4)	(44.4)
% Last year	3.4	5.9	% Unemployed	6.3	5.9
	(18.0)	(23.5)		(24.2)	(23.5)
% Mexico and Central Am.	100.0	69.7	% N.I.L.F.	22.7	21.1
	(0.0)	(45.9)		(41.9)	(40.8)
% Other North Am.	0.0	0.4	% White	58.4	50.0
	(0.0)	(5.9)		(49.3)	(50.0)
% Other Latin Am.	0.0	10.6	% African American	0.6	5.4
	(0.0)	(30.8)		(7.4)	(22.6)
% Europe	0.0	3.1	% Asian	0.1	12.7
	(0.0)	(17.4)		(3.4)	(33.3)
% Asia	0.0	13.3	% Other	40.2	31.3
	(0.0)	(33.9)		(49.0)	(46.4)
% Africa	0.0	2.7	% Married	57.5	52.2
	(0.0)	(16.2)		(49.4)	(50.0)
% Oceania	0.0	0.2	% Single	33.3	37.9
	(0.0)	(4.2)		(47.1)	(48.5)

% Less than H.S.	69.1	45.8	% More than H.S.	0.0	13.7
	(46.2)	(49.8)		(0.0)	(34.3)
% H.S. or equiv.	30.9	26.4	% College or more	0.0	14.1
	(46.2)	(44.1)		(0.0)	(34.8)
Observations	545,209	658,576	Observations	545,209	658,576
Weighted N's	78,010,809	102,089,870	Weighted N's	78,010,809	102,089,870

Source: ACS 2005-2015. Notes: Averages presented with standard deviations in parentheses. H.L.E. refers to the "Hispanics with Low Education" proxy.

Table 4.3: Testing the role of proxies (H.L.E. v. Logical Edits): The effect of E-Verify laws on population sizes of unauthorized immigrants, prime-aged (20-54) sample

	All		Not recent (>5 years)		Recent (1-5 years)		New (<1 year)	
	H.L.E.	Logical Edits	H.L.E.	Logical Edits	H.L.E.	Logical Edits	H.L.E.	Logical Edits
ACS 2005 - 2014								
E-Verify in	-0.060***	-0.067***	-0.028	-0.050**	-0.226***	-0.160*	-0.394	-0.150
current year	(0.022)	(0.023)	(0.028)	(0.021)	(0.062)	(0.095)	(0.253)	(0.173)
Observations	510		510		510		510	
ACS 2005 - 2015								
E-Verify in	-0.075***	-0.074***	-0.038	-0.053**	-0.276***	-0.177**	-0.481*	-0.198
current year	(0.023)	(0.024)	(0.027)	(0.022)	(0.083)	(0.079)	(0.277)	(0.190)
Observations	561		561		561		561	

Source: ACS 2005-2015. * $p<0.1$, ** $p<0.05$, *** $p<0.01$. Note: Standard errors are robust, clustered by state (shown in parentheses). H.L.E. refers to the "Hispanics with Low Education" proxy and are meant to replicate the results of Orrenius and Zavodny (2016). Their results use ACS 2005-2014: All: $-0.061^{**}(0.023)$, Not recent: -0.026(0.026), Recent:$-0.258^{***}(0.071)$, New: $-0.464^{*}(0.259)$. The second set of results extends the analysis to include 2015.

Table 4.4: Testing the role of proxies (H.L.E. v. Logical Edits): The effect of E-Verify laws on population sizes of unauthorized immigrants, full sample

	All		Not recent (>5 years)		Recent (1-5 years)		New (<1 year)	
	H.L.E.	Logical Edits	H.L.E.	Logical Edits	H.L.E.	Logical Edits	H.L.E.	Logical Edits
ACS 2005 - 2014								
E-Verify in current year	-0.060***	-0.083***	-0.028	-0.063***	-0.226***	-0.177**	-0.394	-0.195
	(0.022)	(0.021)	(0.028)	(0.022)	(0.062)	(0.075)	(0.253)	(0.140)
Observations	510		510		510		510	
ACS 2005 - 2015								
E-Verify in current year	-0.075***	-0.090***	-0.038	-0.065***	-0.276***	-0.200***	-0.481*	-0.228
	(0.023)	(0.021)	(0.027)	(0.024)	(0.083)	(0.064)	(0.277)	(0.148)
Observations	561		561		561		561	

Source: ACS 2005-2015. * $p<0.1$, ** $p<0.05$, *** $p<0.01$. Note: Standard errors are robust, clustered by state (shown in parentheses). H.L.E. refers to the "Hispanics with Low Education" proxy and are meant to replicate the results of Orrenius and Zavodny (2016). Their results use ACS 2005-2014: All: $-0.061^{**}(0.023)$, Not recent: -0.026(0.026), Recent:$-0.258^{***}(0.071)$, New: $-0.464^{*}(0.259)$. The second set of results extends the analysis to include 2015.

Table 4.5: Estimates of the impact of E-Verify mandates on probability of employment - 2002 - 2014 - Annual Social and Economic Supplement

	In Labor Force Population			Full CPS		
	All	Male	Female	All	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)
Logical Edits						
Universal	0.050***	0.042***	0.024***	-0.040***	-0.050***	0.103***
	(0.004)	(0.005)	(0.006)	(0.007)	(0.006)	(0.010)
Observations	60,266	39,338	20,928	83,335	44,552	38,783
HLE						
Universal	-0.005	-0.029***	0.041***	-0.103***	-0.155***	0.072***
	(0.005)	(0.005)	(0.012)	(0.004)	(0.010)	(0.009)
Observations	47,993	32,596	15,397	67,650	36,817	30,833
Naturalized Hispanic						
Universal	0.047***	0.048***	0.107***	-0.161***	-0.407***	-0.012
	(0.004)	(0.008)	(0.017)	(0.011)	(0.013)	(0.020)
Observations	31,193	16,712	14,481	45,748	21,184	24,564
US-Born non-Hispanic						
Universal	-0.029***	-0.035***	-0.021***	-0.012***	-0.026***	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	1,007,447	515,796	491,651	1,536,977	728,680	808,297

Source: CPS-ASEC 2002-2014. * p<0.1, ** p<0.05, *** p<0.01. Notes: Controls include gender (when applicable), race, age, marital status, number of children in household, educational attainment, industry fixed effects, state fixed effects, time (year, month) fixed effects, state specific time trends, unemployment rates. Standard errors clustered at the state level. All regressions use survey weights (wtsupp).

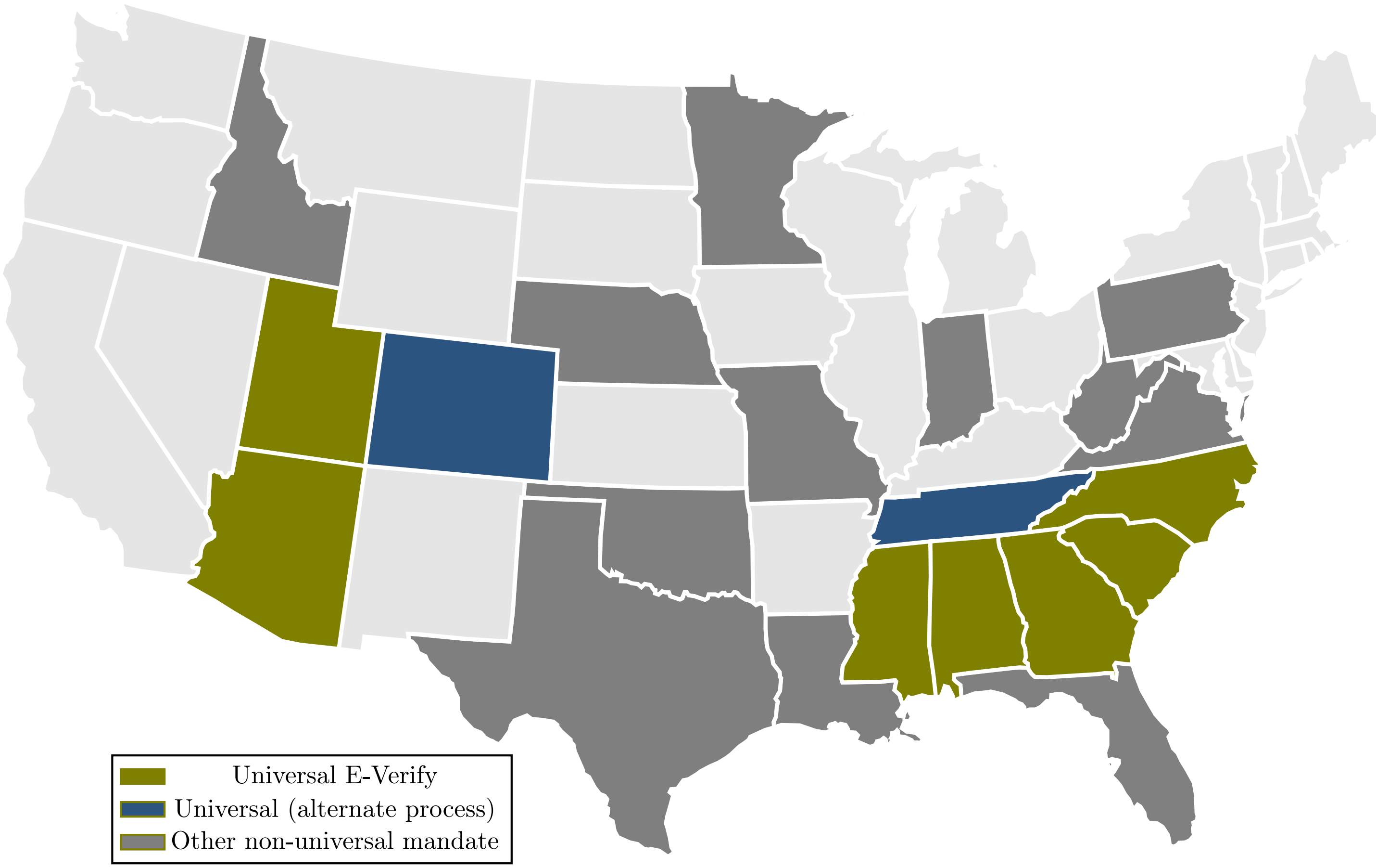
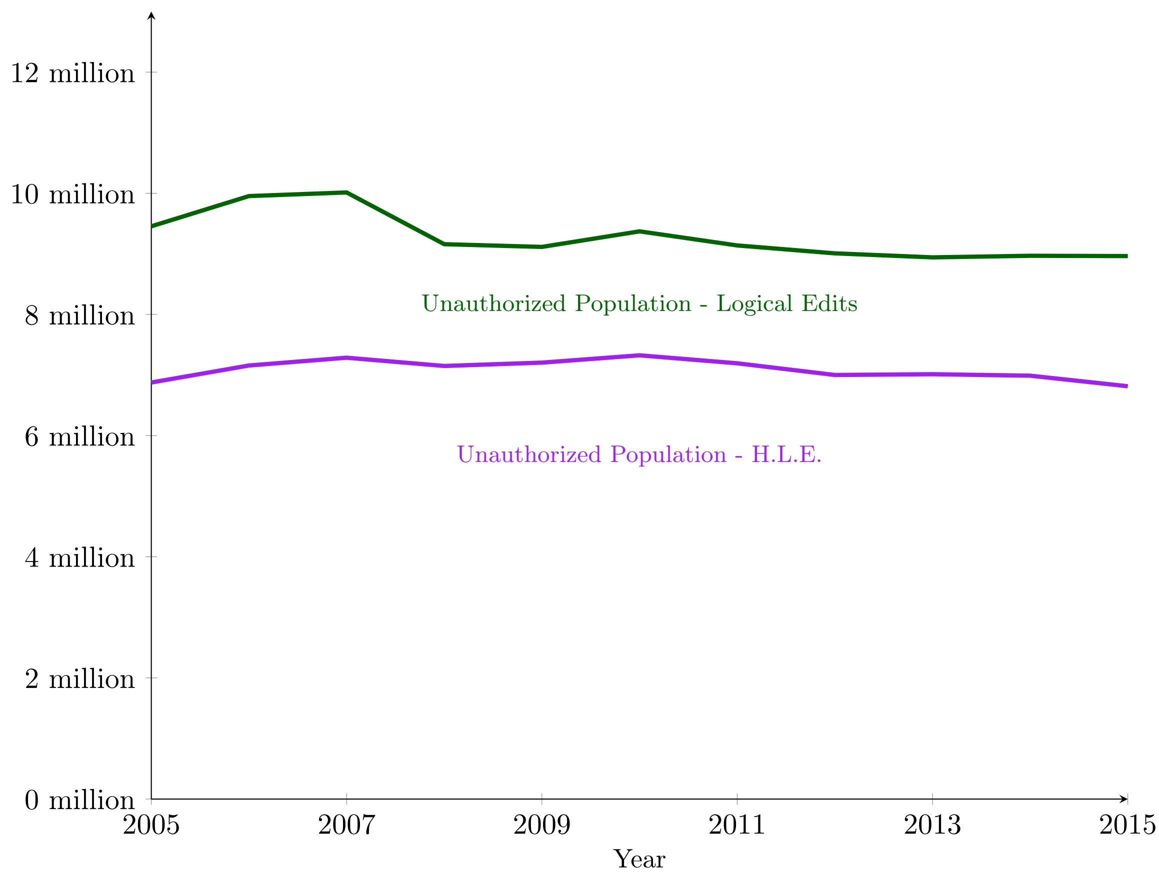


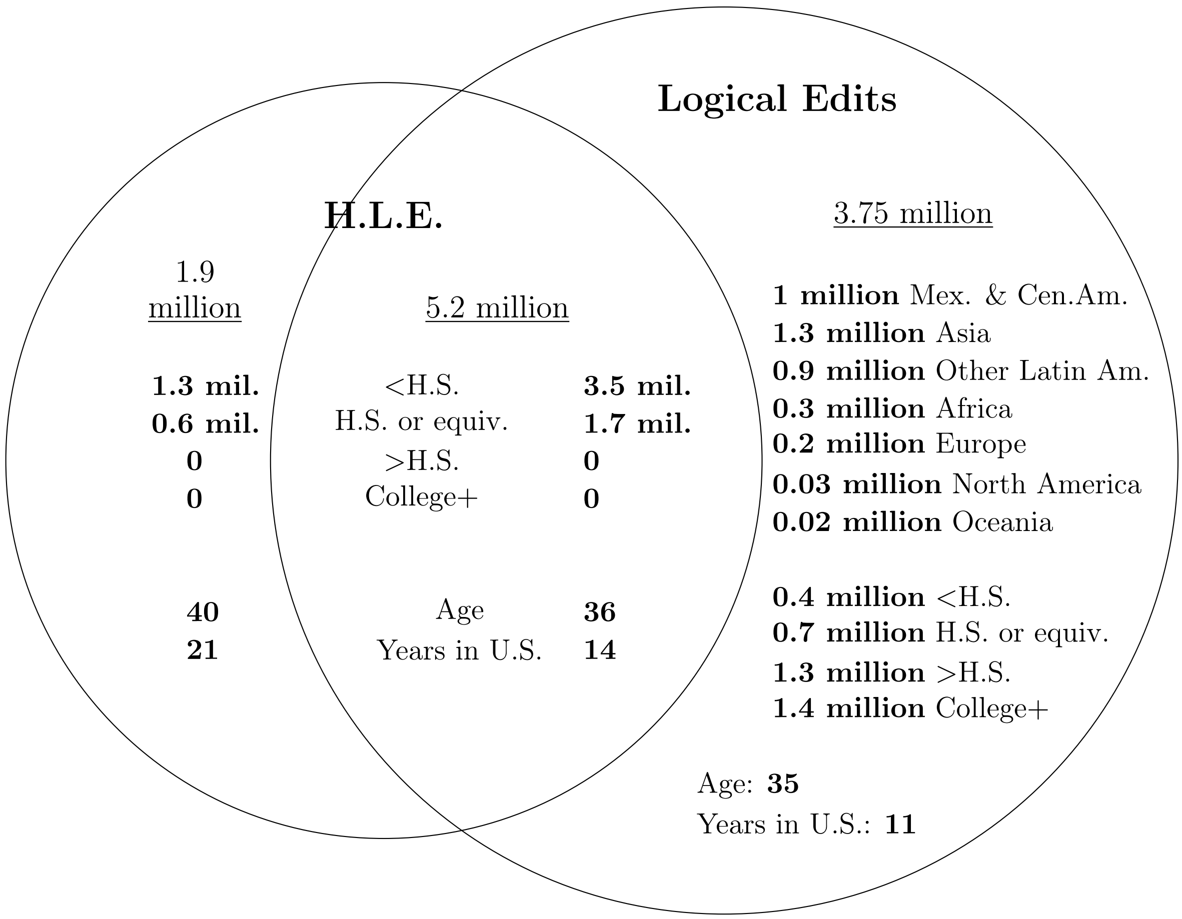
Figure 4.1: States with E-Verify Mandates

Figure 4.2: Population for prime-aged unauthorized immigrants in US - H.L.E. v Logical Edits, ACS 2005-2015



Notes: H.L.E. refers to the "Hispanics with Low Education" proxy. Estimates are weighted using ACS person weights.

Figure 4.3: Comparing H.L.E. and Logical Edits - Characteristics of likely unauthorized by different proxies - Prime-aged workers - 2015



Notes: H.L.E. refers to the "Hispanics with Low Education" proxy. Estimates are weighted using ACS person weights.

References

Amuedo Dorantes, Catalina, and Cynthia Bansak. 2014. "Employment verification mandates and the labor market outcomes of likely unauthorized and native workers." Contemporary Economic Policy, 32(3): 671–680.

Amuedo-Dorantes, Catalina, Cynthia Bansak, and Allan A Zebedee. 2015. "The impact of mandated employment verification systems on state-level employment by foreign affiliates." Southern Economic Journal, 81(4): 928–946.

Amuedo-Dorantes, Catalina, Thitima Puttitanun, and Ana P Martinez-Donate. 2013. "How do tougher immigration measures affect unauthorized immigrants?" Demography, 50(3): 1067–1091.

Auerbach, Alan J, and Philip Oreopoulos. 1999. "Analyzing the fiscal impact of US immigration." The American Economic Review, 89(2): 176–180.

Bachmeier, James D, Jennifer Van Hook, and Frank D Bean. 2014. "Can we measure immigrants' legal status? Lessons from two US Surveys." International Migration Review, 48(2): 538–566.

Bane, Mary Jo, and David T Ellwood. 1983. "Slipping into and out of poverty: The dynamics of spells."

Bitler, Marianne, and Hilary Hoynes. 2016. "The more things change, the more they stay the same? The safety net and poverty in the Great Recession." Journal of Labor Economics, 34(S1): S403–S444.

Blank, Rebecca M, and Harry J Holzer. 1997. It takes a nation: A new agenda for fighting poverty. JSTOR.

Blau, Francine D. 1984. "The use of transfer payments by immigrants." ILR Review, 37(2): 222–239.

Bohn, Sarah, Magnus Lofstrom, and Steven Raphael. 2013. "Did the 2007 Legal Arizona Workers Act reduce the state's unauthorized immigrant population?" Review of Economics and Statistics.

Bohn, Sarah, Magnus Lofstrom, and Steven Raphael. 2015. "Do E-verify mandates improve labor market outcomes of low-skilled native and legal immigrant workers?" Southern Economic Journal, 81(4): 960–979.

Bollinger, Christopher, Barry T Hirsch, Charles M Hokayem, and James P Ziliak. 2014. "Trouble in the Tails? Earnings Nonresponse and Response Bias across the Distribution Using Matched Household and Administrative Data."

Bollinger, Christopher R, and Barry T Hirsch. 2006. "Match bias from earnings imputation in the Current Population Survey: The case of imperfect matching." Journal of Labor Economics, 24(3): 483–519.

Bollinger, Christopher R, and Barry T Hirsch. 2013. "Is earnings nonresponse ignorable?" Review of Economics and Statistics, 95(2): 407–416.

Borjas, George J. 2017a. "The earnings of undocumented immigrants." National Bureau of Economic Research.

Borjas, George J. 2017b. "The labor supply of undocumented immigrants." Labour Economics, 46: 1–13.

Brady, David, Regina S Baker, and Ryan Finnigan. 2013. "When unionization disappears: State-level unionization and working poverty in the United States." American Sociological Review, 78(5): 872–896.

Buchmueller, Thomas C, John DiNardo, and Robert G Valletta. 2002. "Union effects on health insurance provision and coverage in the United States." ILR Review, 55(4): 610–627.

Budd, John W. 2007. "The effect of unions on employee benefits and non-wage compensation: Monopoly power, collective voice, and facilitation." In What Do Unions Do? A Twenty Year Perspective. Transaction Publishers.

Carnes, Nicholas. 2012. "Does the numerical underrepresentation of the working class in Congress matter?" Legislative Studies Quarterly, 37(1): 5–34.

Carnes, Nicholas. 2013. White-collar government: The hidden role of class in economic policy making. University of Chicago Press.

Cellini, Stephanie Riegg, Signe-Mary McKernan, and Caroline Ratcliffe. 2008. "The dynamics of poverty in the United States: A review of data, methods, and findings." Journal of Policy Analysis and Management, 27(3): 577–605.

Citizenship, US, and Immigration Services. 2014. "History and milestones."

Council of Economic Advisers. 2015. "The economics of early childhood investments." White House, Washington D.C.

DiNardo, John, and David S Lee. 2004. "Economic impacts of new unionization on private sector employers: 1984–2001." The Quarterly Journal of Economics, 119(4): 1383–1441.

Eidelson, J. 2013. "Video: McDonald's tells workers to get food stamps." http://www.salon.com/2013/10/23/videomcdonaldstellsworkerstogetfoodstamps/.

Elango, Sneha, Jorge Luis García, James J Heckman, and Andrés Hojman. 2015. "Early childhood education." National Bureau of Economic Research.

Eller, TJ. 1996. Dynamics of economic well-being: Who stays poor? Who doesn't? US Department of Commerce, Economics and Statistics Administration, Bureau of the Census.

Feng, Shuaizhang. 2013. "Identification and statistical inference using matched March CPS data, with an application to US poverty dynamics." Journal of Economic and Social Measurement, 38(2): 159–170.

Flood, Sarah, and José Pacas. 2016. "Using the Annual Social and Economic Supplement with Current Population Survey panels." Minnesota Population Center Working Paper No. 2016-4.

Flood, Sarah, Miriam King, Steven Ruggles, and J Robert Warren. 2015. "Integrated public use microdata series, Current Population Survey: Version 4.0. [Machine-readable database]." Minneapolis: University of Minnesota.

Fox, Liana, Christopher Wimer, Irwin Garfinkel, Neeraj Kaushal, and Jane Waldfogel. 2015. "Waging war on poverty: Poverty trends using a historical supplemental poverty measure." Journal of Policy Analysis and Management, 34(3): 567–592.

Frandsen, Brigham R. 2012. "Why unions still matter: The effects of unionization on the distribution of employee earnings." Manuscript. Cambridge, MA: MIT.

Freeman, Richard B. 1981. "The effect of unionism on fringe benefits." ILR Review, 34(4): 489–509.

Freeman, Richard B. 1984. "Longitudinal analyses of the effects of trade unions." Journal of Labor Economics, 2(1): 1–26.

Freeman, Richard B., and JL Medoff. 1984. "What do unions do."

Good, Michael. 2012. "Do immigrant outflows lead to native inflows? An empirical analysis of the migratory responses to US." Economics Research Working Paper Series. 10.

Henry, Raymond. 2015. "Effect of state E-Verify laws on H2A program utilization." Master's diss. Georgia State University.

Hirsch, Barry T, and David A Macpherson. 2003. "Union membership and coverage database from the current population survey: Note." ILR Review, 56(2): 349–354.

Hirsch, Barry T, and Edward J Schumacher. 2004. "Match bias in wage gap estimates due to earnings imputation." Journal of Labor Economics, 22(3): 689–722.

Hokayem, Charles, and Misty L Heggeness. 2014. "Factors influencing transitions into and out of near poverty: 2004-2012." SEHSD Working Paper 2014-05. Washington, DC: US Census Bureau.

Hokayem, Charles, Christopher R Bollinger, and James P Ziliak. 2014. "The role of CPS non-response on the level and trend in poverty." University of Kentucky Center for Poverty Research Discussion Paper Series, DP2014-05.

Lee, David S, and Alexandre Mas. 2012. "Long-run impacts of unions on firms: New evidence from financial markets, 1961–1999." The Quarterly Journal of Economics, 127(1): 333–378.

Leerkes, Arjen, Mark Leach, and James Bachmeier. 2012. "Borders behind the border: An exploration of state-level differences in migration control and their effects on US migration patterns." Journal of Ethnic and Migration Studies, 38(1): 111–129.

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, Part 1): 489–506.

McKernan, Signe-Mary, and Caroline Ratcliffe. 2005. "Events that trigger poverty entries and exits." Social Science Quarterly, 86(s1): 1146–1169.

Naifeh, Mary. 1998. Dynamics of economic well-being, poverty, 1993-94: trap door? revolving door? or both? Census Bureau.

National Academies of Sciences, Engineering, and Medicine. 2016. The Economic and Fiscal Consequences of Immigration: A Report. National Academies Press.

O'Hara, Amy. 2006. "Tax variable imputation in the Current Population Survey." Department of the Treasury Internal Revenue Service, 169.

Orrenius, Pia M, and Madeline Zavodny. 2015. "The impact of E-Verify mandates on labor market outcomes." Southern Economic Journal, 81(4): 947–959.

Orrenius, Pia M, and Madeline Zavodny. 2016. "Do state work eligibility verification laws reduce unauthorized immigration?" IZA Journal of Migration, 5(1): 5.

Passel, Jeffrey S, and D'Vera Cohn. 2010. "US unauthorized immigration flows are down sharply since mid-decade." Washington, DC: Pew Hispanic Center.

Passel, Jeffrey S, D Cohn, and A Gonzalez-Barrera. 2013. "Population decline of unauthorized immigrants stalls, may have reversed." Population.

Pew Research Center. 2016. America's shrinking middle class: A close look at changes within metropolitan areas.

Preston, Ian. 2014. "The effect of immigration on public finances." The Economic Journal, 124(580).

Proctor, Bernadette D, Jessica L Semega, and Melissa A Kollar. 2016. "US Census Bureau, Current Population Reports, P60-252, Income and Poverty in the United States. 2015. US Government Printing Office: Washington, DC, 2016." https://www.census.gov/content/dam/Census/library/publications/2016/demo/p60-256.pdf.

Rank, Mark R, and Thomas A Hirschl. 2001. "Poverty across the life cycle: Evidence from the PSID." Journal of Policy Analysis and Management, 20(4): 737–755.

Renwick, Trudi, and Liana Fox. 2016. "The Supplemental Poverty Measure: 2015." Current Population Reports, Series P60, 258.

Ribar, David, and Karen S Hamrick. 2003. "Dynamics of poverty and food sufficiency."

Rivera Drew, Julia A, Sarah Flood, and John Robert Warren. 2014. "Making full use of the longitudinal design of the Current Population Survey: Methods for linking records across 16 months." Journal of Economic and Social Measurement, 39(3): 121–144.

Romig, Kathleen, and Arloc Sherman. 2016. "Social Security keeps 22 million Americans out of poverty: A state-by-state analysis." Center for Budget and Policy Priorities.

Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. 2015a. "Integrated Public Use Microdata Series, Current Population Survey: Version 4.0 [dataset]." Minneapolis: University of Minnesota.

Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. 2015b. "Integrated Public Use Microdata Series: Version 6.0 [dataset]." Minneapolis: University of Minnesota.

Sanders, K. 2012. "Alan Grayson says more Walmart employees on Medicaid, food stamps than other companies." http://www.politifact.com/truth-o-meter/statements/2012/dec/06/alan-grayson/alan-grayson-says-more-walmart-employees-medicaid-/.

Schreur, Elliot, and Kristen Arnold. 2015. "What is the Social Security retirement age?" https://www.nasi.org/learn/socialsecurity/retirement-age.

Sherman, Arloc, Robert Greenstein, and Kathy Ruffing. 2012. "Contrary to 'Entitlement Society' rhetoric, over nine-tenths of entitlement benefits go to elderly, disabled, or working households." Center on Budget and Policy Priorities, 10.

Sojourner, Aaron J. 2013. "Do unions promote members' electoral office holding? Evidence from correlates of state legislatures' occupational shares." ILR Review, 66(2): 467–486.

Sojourner, Aaron J, Brigham R Frandsen, Robert J Town, David C Grabowski, and Min M Chen. 2015. "Impacts of unionization on quality and productivity: Regression discontinuity evidence from nursing homes." ILR Review, 68(4): 771–806.

Stark, Oded, and Marcin Jakubek. 2012. "Employer sanctions and the welfare of native workers." Economics Letters, 117(3): 533–536.

Stewart, Jay. 2002. Recent trends in job stability and job security: Evidence from the March CPS. US Department of Labor, Bureau of Labor Statistics, Office of Employment and Unemployment Statistics.

Storesletten, Kjetil. 2000. "Sustaining fiscal policy through immigration." Journal of Political Economy, 108(2): 300–323.

Turek, Joan, Fritz Scheuren, Brian Sinclair-James, Bula Ghose, and Sameer Desale. 2009. "Effects of imputation on CPS Income and Poverty Series: 1981–2007." JSM Proceedings: Section on Survey Research Methods, 1892–1900.

U.S. Congressional Budget Office. 2014. "Taxing capital income: Effective marginal tax rates under 2014 law and selected policy options." Congressional Budget Office, Washington D.C.

U.S. Internal Revenue Service. 2012. "Tax gap for tax year 2006." Internal Revenue Service, Washington D.C.

Van Hook, Jennifer, James D Bachmeier, Donna L Coffman, and Ofer Harel. 2014. "Can we spin straw into gold? An evaluation of immigrant legal status imputation approaches." Demography, 1–26.

Warren, Robert. 2014. "Democratizing data about unauthorized residents in the United States: Estimates and public-use data, 2010 to 2013." Journal on Migration and Human Security, 2(4): 305–328.

Warren, Robert. 2017. "Zero undocumented population growth is here to stay and immigration reform would preserve and extend these gains." Journal on Migration and Human Security, 5(2).

Warren, Robert, and Jeffrey S Passel. 1987. "A count of the uncountable: estimates of undocumented aliens counted in the 1980 United States Census." Demography, 24(3): 375–393.

Warren, Robert, and John Robert Warren. 2013. "Unauthorized immigration to the United States: Annual estimates and components of change, by state, 1990 to 2010." International Migration Review, 47(2): 296–329.

Watson, Tara. 2010. "Inside the refrigerator: Immigration enforcement and chilling effects in Medicaid participation." National Bureau of Economic Research Working Paper 16278.

Westat. 2009. "Findings of the E-Verify program evaluation."

Wheaton, Laura, and Kathryn Stevens. 2016. "The effect of different tax calculators on the Supplemental Poverty Measure." Urban Institute.

Wooldridge, Jeffrey M. 2005. "Violating ignorability of treatment by controlling for too many factors." Econometric Theory, 21(5): 1026–1028.

Appendix A

Chapter 1

Table A.1: Descriptive statistics of families by OPM poverty transitions

	Full Sample	Never Poor	NonPoor to Poor	Poor to NonPoor	Always Poor
Number of People - T1	2.14	2.18	1.89	1.92	2.02
	(0.004)	(0.004)	(0.016)	(0.017)	(0.017)
Number of People - T2	2.13	2.17	1.86	1.96	2.03
	(0.004)	(0.004)	(0.016)	(0.017)	(0.017)
Selected Demographics of Householder					
Age of Householder	53.65	54.26	53.60	51.85	47.45
	(0.047)	(0.051)	(0.211)	(0.223)	(0.187)
% White - Non. Hisp.	76.93	79.97	67.91	66.70	55.60
	(0.118)	(0.124)	(0.543)	(0.552)	(0.540)
% Afr. Amer. - Non. Hisp.	8.23	6.69	13.21	13.32	18.73
	(0.077)	(0.077)	(0.394)	(0.398)	(0.424)
% Amer. Ind. - Non. Hisp.	0.47	0.37	0.70	0.90	1.17
	(0.019)	(0.019)	(0.097)	(0.111)	(0.117)
% Asian - Non. Hisp.	3.34	3.39	3.09	3.21	3.08
	(0.050)	(0.056)	(0.201)	(0.207)	(0.188)
% Other - Non. Hisp.	1.13	1.05	1.29	1.63	1.60
	(0.030)	(0.032)	(0.131)	(0.148)	(0.136)
% Hispanic	9.89	8.53	13.80	14.24	19.83
	(0.084)	(0.087)	(0.401)	(0.410)	(0.433)
% Non-Citizen	6.35	5.35	8.78	10.71	12.97
	(0.068)	(0.070)	(0.329)	(0.362)	(0.365)
% Naturalized Citizen	7.00	7.04	7.06	6.63	6.70
	(0.071)	(0.079)	(0.298)	(0.292)	(0.272)
Observations	127,404	104,257	7,387	7,284	8,476

Source: CPS-ASEC 2010-2016. Note: Standard errors in parentheses.

Table A.2: Family composition changes by OPM poverty transitions

	Full Sample	Never Poor	NonPoor to Poor	Poor to NonPoor	Always Poor
Summary of Changes[a]					
% Departure of Any Family Member	5.90	5.90	8.78	4.53	4.50
	(0.066)	(0.073)	(0.329)	(0.244)	(0.225)
% Arrival of Any Family Member	4.54	4.40	4.87	5.79	4.91
	(0.058)	(0.063)	(0.250)	(0.274)	(0.235)
% Both Departure and Arrival	0.61	0.59	0.79	0.71	0.60
	(0.022)	(0.024)	(0.103)	(0.098)	(0.084)
% No Change in Family Composition	88.95	89.11	85.56	88.97	89.98
	(0.088)	(0.096)	(0.409)	(0.367)	(0.326)
Arrivals[a]					
% Newborn	1.90	1.82	2.28	1.49	2.90
	(0.038)	(0.041)	(0.174)	(0.142)	(0.182)
% Arrival Child (2-6)	0.35	0.32	0.70	0.39	0.45
	(0.017)	(0.017)	(0.097)	(0.073)	(0.072)
% Arrival Child (7-17)	0.59	0.52	1.05	0.72	0.98
	(0.021)	(0.022)	(0.119)	(0.099)	(0.107)
% Arrival Adult (18-64)	2.77	2.73	2.61	4.24	2.01
	(0.046)	(0.051)	(0.186)	(0.236)	(0.152)
% Arrival Elder (65+)	0.30	0.29	0.23	0.56	0.16
	(0.015)	(0.017)	(0.056)	(0.087)	(0.044)
% Marriage	0.73	0.72	0.37	1.57	0.47
	(0.024)	(0.026)	(0.071)	(0.146)	(0.074)
Departures[a]					
% Departure Child (0-6)	0.48	0.44	0.41	0.85	0.81
	(0.019)	(0.020)	(0.074)	(0.108)	(0.097)
% Departure Child (7-17)	0.76	0.69	0.54	1.38	1.17
	(0.024)	(0.026)	(0.086)	(0.137)	(0.117)

% Departure Adult (18-64)	5.07	5.13	6.75	3.86	3.86
	(0.061)	(0.068)	(0.292)	(0.226)	(0.209)
% Departure Elder (65+)	1.03	0.99	2.66	0.67	0.34
	(0.028)	(0.031)	(0.187)	(0.096)	(0.063)
% Departure Spouse	0.89	0.84	2.26	0.76	0.49
	(0.026)	(0.028)	(0.173)	(0.102)	(0.076)
% Departure Head	1.12	0.99	3.12	0.92	1.02
	(0.029)	(0.031)	(0.202)	(0.112)	(0.109)
% Divorce/Separation	0.57	0.50	1.79	0.50	0.44
	(0.021)	(0.022)	(0.154)	(0.083)	(0.072)
% Widowed	0.65	0.65	1.46	0.31	0.13
	(0.022)	(0.025)	(0.140)	(0.065)	(0.040)
Observations	127,404	104,257	7,387	7,284	8,476

Source: CPS-ASEC 2010-2016. Notes: Standard errors in parentheses. (a) The summary of changes are mutually exclusive but the breakdown of arrivals and departures are not mutually exclusive. Changes in family composition can be concurrent.

Table A.3: Average recipiency amount of SPM components by poverty transition type and time period

	All		NonPoor to Poor		Poor to NonPoor	
	t_1	t_2	t_1	t_2	t_1	t_2
SPM Total Resources	54,206	53,697	38,543	9,252	9,604	37,380
	(55,976)	(56,546)	(36,098)	(15,034)	(14,940)	(36,547)
Total Cash Income	74,505	74,583	48,523	15,017	15,427	47,434
	(84,684)	(86,669)	(52,888)	(15,001)	(13,863)	(56,558)
Wages/Salaries	53,641	53,353	25,673	7,229	7,666	25,216
	(80,362)	(82,233)	(45,793)	(13,993)	(13,222)	(52,598)
Other Cash Income	20,864	21,230	22,850	7,788	7,761	22,218
	(36,919)	(37,466)	(35,858)	(9,104)	(9,270)	(29,026)
Taxes Paid	14,159	14,646	6,264	381	281	6,254
	(28,462)	(29,825)	(16,457)	(3,441)	(2,973)	(19,915)
Necessary Expenses	2,075	2,049	1,329	1,023	1,092	1,353
	(2,770)	(2,725)	(1,875)	(2,760)	(2,788)	(1,862)
Government Subsidies	454	442	858	1,001	1,038	875
	(1,762)	(1,709)	(2,465)	(2,493)	(2,589)	(2,422)
Medical Expenditures	4,538	4,632	3,269	5,362	5,505	3,323
	(6,969)	(6,948)	(4,293)	(16,265)	(15,061)	(4,307)
Observations	127,146		10,308		10,170	

Source: CPS-ASEC 2010-2016. Notes: All income values are adjusted to reflect 2015 dollars. Adjustment made using annual average Consumer Price Index Research Series (CPI-U-RS). Standard deviations in parentheses.

Table A.4: Job loss and poverty transitions

Panel A: Work schedule changes and poverty transitions

Changes in Householder Weeks Worked		Total Families	NonPoor to Poor[a]	NonPoor to NonPoor[a]
Weeks Worked in t_1	Weeks Worked in t_2	N	N (% Row)	N (% Row)
52-50 weeks	52-50 weeks	48,557	1,295 (2.7%)	45,471 (93.6%)
52-50 weeks	49-40 weeks	2,043	133 (6.5%)	1,791 (87.7%)
52-50 weeks	39-27 weeks	1,157	93 (8.0%)	991 (85.7%)
52-50 weeks	26-14 weeks	512	76 (14.8%)	408 (79.7%)
52-50 weeks	13-1 week	330	59 (17.9%)	253 (76.7%)
52-50 weeks	Unemployed	2,371	806 (34.0%)	1,241 (52.3%)
52-50 weeks	Retired	1,376	256 (18.6%)	1,048 (76.2%)

Panel B: The influence of the change in wages on poverty transitions

Changes in Householder Weeks Worked			Percent of NP-P where change in wages is sufficient[b]
Weeks Worked in t_1	Weeks Worked in t_2	N	N (%)
52-50 weeks	52-50 weeks	1,295	631 (48.7%)
52-50 weeks	49-40 weeks	133	74 (55.6%)
52-50 weeks	39-27 weeks	93	70 (75.0%)

52-50 weeks	26-14 weeks	76	50 (66.2%)
52-50 weeks	13-1 week	59	48 (81.4%)
52-50 weeks	Unemployed	806	539 (66.9%)
52-50 weeks	Retired	256	144 (56.3%)

Source: CPS-ASEC 2010-2016. Notes: (a) Two transitions are omitted (Poor to NonPoor, Poor to Poor) and therefore rows do not add up to 100%. (b) Sufficient refers to the framework established in section 2.6.2.

Table A.5: Avoiding poverty after job loss

Panel A: The role of government subsidies in keeping families out of poverty after job loss

Changes in Householder Weeks Worked			Percent of NP-NP where gov't subs and EITC keep families out of poverty[a]
Weeks Worked in t_1	Weeks Worked in t_2	N	N (%)
52-50 weeks	52-50 weeks	45,471	555 (1.2%)
52-50 weeks	49-40 weeks	1,791	33 (1.8%)
52-50 weeks	39-27 weeks	991	44 (4.4%)
52-50 weeks	26-14 weeks	408	21 (5.1%)
52-50 weeks	13-1 week	253	15 (5.9%)
52-50 weeks	Unemployed	1,241	95 (7.7%)
52-50 weeks	Retired	1,048	9 (0.9%)

Panel B: The role of unemployment compensation in keeping families out of poverty after job loss

Changes in Householder Weeks Worked			Percent of NP-NP where unemployment compensation keeps families out of poverty
Weeks Worked in t_1	Weeks Worked in t_2	N	N (%)

52-50 weeks	52-50 weeks	45,471	44 (0.01%)
52-50 weeks	49-40 weeks	1,791	11 (0.6%)
52-50 weeks	39-27 weeks	991	35 (3.5%)
52-50 weeks	26-14 weeks	408	31 (7.6%)
52-50 weeks	13-1 week	253	33 (13.0%)
52-50 weeks	Unemployed	1,241	150 (12.1%)
52-50 weeks	Retired	1,048	2 (0.2%)

Source: CPS-ASEC 2010-2016. Notes: (a) Government subsidies include those outlined in Table 2.2 (SNAP, lunch subsidies, WIC, housing subsidies and energy subsidies) *as well as* EITC.

Table A.6: Family member departures and poverty transitions

Panel A: Departures of male and female adults and poverty transitions

Type of Departure	Total Families	NonPoor to NonPoor	NonPoor to Poor	Poor to NonPoor	Poor to Poor
	N	N (% Row)	N (% Row)	N (% Row)	N (% Row)
Male Departs	1,657	1,080 (65.2%)	351 (21.2%)	112 (6.8%)	114 (6.9%)
Divorce	550	370 (67.2%)	114 (20.7%)	34 (6.2%)	32 (5.8%)
Death	528	349 (66.1%)	119 (22.5%)	32 (6.1%)	28 (5.3%)
Female Departs	1,347	972 (72.2%)	173 (12.8%)	125 (9.3%)	77 (5.7%)
Divorce	526	402 (76.4%)	67 (12.7%)	36 (6.8%)	21 (4.0%)
Death	281	224 (79.7%)	19 (6.8%)	28 (10.0%)	10 (3.6%)

Panel B: The influence of the loss in wages from departure on the transition into poverty

Type of Departure	Total NonPoor to Poor Families	Percent of NP-P where change in wages is sufficient
	N	N (%)
Male Departs	351	142 (40.5%)
Divorce	114	62 (54.9%)
Death	119	15 (12.6%)

Female Departs	173	73 (42.2%)
Divorce	67	38 (56.1%)
Death	19	2 (10.5%)

Source: CPS-ASEC 2010-2016.

Appendix B

Chapter 2

B.1 Additional robustness analysis

Table B.4 presents estimates that do not use weights. Results are largely stable.

Recall that our main analysis focuses on the set of cases with full-observed data and excluded cases with any variable (outcome, treatment, or control) missing or imputed. Table B.5 describes how the sample size is affected by this set of exclusions. Table B.5 displays estimated coefficients on each of 4 outcomes under each of the three specifications after including only one kind of imputation at a time, then, in the final row, including all of them. The cross-sectional results of specifications 1 and 2 are, by-and-large quite stable. The fixed effect estimates are qualitatively stable except for inclusion of full-line imputes and inclusion of all imputes. These findings are in line with Hirsch and Schumacher (2004) and Bollinger and Hirsch (2006). Imputations attenuate the union coefficient towards zero.

B.2 Data note

Understanding the sampling methodology of the CPS is key to understanding our sample. The CPS is a monthly survey designed to collect data primarily on employment; the Basic Monthly CPS's are the source of the official unemployment statistics. The Basic Monthly CPS consists of about 60,000 dwellings. Each dwelling is selected to be in the CPS for 4 consecutive months, then out of the CPS rotation for 8 months, and then back again for 4 more months. Each of these months is referred to as a Month-In-Sample (MIS) for a total of 8 MIS's for any given dwelling. The ORG questions refer to a survey that is given only to dwellings from MIS groups 4 or 8 (i.e. these are the months after which these dwellings will either be out of sample for 8 months or out of the CPS entirely). The questions encompassed in the ORGs focus on more specific labor questions, most important to our study is the union membership question. The union membership question is thus asked only of one-fourth of any given Basic Monthly CPS.

Every March, the CPS administers the Annual Social and Economic Supplement (ASEC) to all dwellings in the March Basic CPS. Among the questions asked here are detailed breakdowns of annual income sources and social program benefit recipiency. In order to conduct our analysis, it is necessary for us to link the March Basic CPS to the ASEC which is a more tedious ordeal in practice. We use the newly created identifiers of the Minnesota Population Center (MARBASECID) for this purpose; the exact algorithm and more detailed explanation of the CPS sampling methodology is included in Flood and Pacas (2016).

As part of our sensitivity analysis, we use our linked-CPS sample with the biennial Displaced Worker Survey (1998-2014). Because this survey is fielded in January and February, it is only possible to link a smaller subsample of March respondents. Specifically, from the 36,531 (unweighted) displaced workers, 2,823 workers are linked to our sample, about 7.8 percent of the total Displaced Worker Survey sample.

Table B.1: Sample description

Sample:	1	2	3	4	5	6	7
Year:	N	N	N	N	N	N	N
1994	150,943	133,669	121,386	86,477	11,658	5,927	2,245
1995	149,642	100,490	94,223	67,658	11,658	5,927	2,245
1996	130,476	114,667	104,622	74,532	13,190	6,668	1,967
1997	131,854	115,963	105,774	75,723	26,652	13,485	4,033
1998	131,617	115,369	105,640	75,736	25,672	13,059	3,934
1999	132,324	115,066	105,165	75,627	23,483	11,897	3,589
2000	133,710	115,800	105,968	76,632	22,062	11,189	3,420
2001	218,269	111,062	101,365	73,573	20,726	10,636	3,238
2002	217,219	133,615	120,342	87,153	21,778	11,153	3,347
2003	216,424	135,524	122,174	88,847	23,912	12,124	3,578
2004	213,241	131,818	119,186	86,624	22,802	11,562	3,480
2005	210,648	129,816	118,784	86,697	22,741	11,401	3,436
2006	208,562	128,322	117,205	85,777	24,913	12,424	3,634
2007	206,639	127,990	116,991	85,685	26,498	13,479	3,992
2008	206,404	127,219	116,347	86,102	27,546	13,776	4,153
2009	207,921	128,976	117,473	86,767	28,148	13,625	4,054
2010	209,802	129,156	117,896	87,378	27,916	13,434	3,893
2011	204,983	126,241	115,228	86,110	27,292	13,296	3,793
2012	201,398	125,256	114,506	85,819	27,182	13,164	3,757
2013	202,634	124,254	113,802	85,412	24,583	11,498	3,292
2014	199,556	123,438	112,740	75,403	18,763	8,597	2,446
2015	199,024	122,467	111,277	83,695	7,789	3,585	1,026
Total	4,083,290	2,716,178	2,478,094	1,803,427	486,964	241,906	72,552

Source: CPS-ASEC 1994-2015, CPS-ORG 1994-2015. Sample 1 - Full ASEC. Sample 2 - Keep if month-in-sample (MIS) equals 4 or 8 and then only the people that link to their ORG. This should include all March respondents and respondents who link to April, May, and June. Sample 3 - Drop respondents who do not match on age, sex, race and seem to be at the same job. Sample 4 - Drops respondents aged less than 18. Sample 5 (this is the sample used for Tables 7 and 8) - Keep NILF and unemployed . Drop if respondent not in both years, or if union is not in universe (civilians 15+ wage/salary workers, excludes self-employed), or do not match in sex, hispanic, black, asian, or foreign born. Sample 6 (main sample) - Drop if respondent not in both years, or if union is not in universe (civilians 15+ wage/salary workers, excludes self-employed), or do not match in sex, hispanic, black, asian, or foreign born. Sample 7 - Keep only March Basic observations from Sample 6.

Table B.2: Details of variables

Variable	Var Name (IPUMS)	Variable (Census)	Type	Record type	Special construction
		Taxes			
Federal income tax liability before credits	FEDTAX	FEDTAX_BC	Imputed	Person	None
State income tax liability before credits	STATETAX	STATETAX_BC	Imputed	Person	None
Annual property taxes	PROPTAX	PROP_TAX	Imputed	Household	Divided by total adults (age>18) in household
Social security retirement payroll deduction	FICA	FICA	Imputed	Person	None
Federal retirement payroll deduction	FEDRETIR	FED_RET	Imputed	Person	None
Earned income tax credit	EITCRED	EIT_CRED	Imputed	Person	None
Additional child tax credit	ACTCCRD	ACTC_CRD	Imputed	Person	None
Child tax credit	CTCCRD	CTC_CRD	Imputed	Person	None
Credit received from Making Work Pay	MWPVAL	MWP_VAL	Imputed	Person	CPS ASEC 2010-2011
Federal stimulus payment	STIMULUS	STIMULUS	Imputed	Person	CPS ASEC 2009
		Income from Public Benefits			
Supplemental Security Income	INCSSI	SSI_VAL	Collected	Person	None
Welfare (Public assistance)	INCWELFR	PAW_VAL	Collected	Person	None
Person market value of Medicare	PMVCARE	P_MVCARE	Receipt	Person	None
Person market value of Medicaid	PMVCAID	P_MVCAID	Receipt	Person	None
Person value of food stamps	FMFDSTAMP (Not in IPUMS)	F_MV_FS	Collected	Family	Divided by total adults (age>18) in family
Person value of housing subsidy	FMVHOUSSUB (Not in IPUMS)	FHOUSSUB	Receipt	Family	Divided by total adults (age>18) in family

Person value of school-lunch subsidy	FMVSCHLUNCH (Not in IPUMS)	F_MV_SL	Receipt	Family	Divided by total adults (age>18) in family
Person value of energy subsidy	HEATVAL	HENGVAL	Collected	Household	Divided by total adults (age>18) in household
Educational assistance (beyond HS)	INCEDUC	ED_VAL	Collected	Person	None
Social security	INCSS	SS_VAL	Collected	Person	None
Unemployment benefits	INCUNEMP	UC_VAL	Collected	Person	None
Worker's compensation	INCWKCOM	WC_VAL	Collected	Person	None
Veteran's benefits	INCVET	VET_VAL	Collected	Person	None
Disability benefits	INCDISAB	DSAB_VAL	Collected	Person	None
Survivor's benefits	INCSURV	SRVS_VAL	Collected	Person	None
		<u>Private Income</u>			
Wage and salary income	INCWAGE	PEARNVAL	Collected	Person	None
Alimony	INCALIM	ALM_VAL	Collected	Person	None
Non-farm business income	INCBUS	SEMP_VAL	Collected	Person	None
Child support	INCCHILD	CSP_VAL	Collected	Person	None
Dividends	INCDIVID	DIV_VAL	Collected	Person	None
Farm	INCFARM	FRM_VAL	Collected	Person	None
Interest	INCINT	INT_VAL	Collected	Person	None
Income from other source not specified	INCOTHER	OI_VAL	Collected	Person	None
Rent	INCRENT	RNT_VAL	Collected	Person	None
Retirement	INCRETIR	RTM_VAL	Collected	Person	None
Assistance from friends not in HH	INCASIST	FIN_VAL	Collected	Person	None
Implied value of owner-occupied housing	HOUSRET	HOUSRET	Imputed	Household	Divided by total adults (age>18) in household

Table B.3: Summary statistics for variables and underlying components in full sample, union subsample and non-union subsample

Sample:	All	Union	Non-union	Unemployed	Idle	In School
Net fiscal impact	1120.7	10833.7	7670.5	-686.8	-11856.7	-3497.4
	(18087.5)	(14324.0)	(15857.3)	(13264.0)	(15746.9)	(10150.3)
Taxes paid	7184.0	12368.6	9721.4	4134.3	2370.7	308.4
	(11877.1)	(12196.8)	(13295.8)	(8856.6)	(6981.0)	(2570.7)
Federal income tax liability before credits	4002.8	6366.3	5427.1	2260.8	1431.3	203.7
	(12426.1)	(14002.0)	(14436.7)	(8945.3)	(7484.5)	(4197.4)
State income tax liability before credits	1142.9	2018.2	1534.3	713.4	381.2	47.29
	(3741.8)	(3880.8)	(4425.9)	(2531.7)	(2179.4)	(566.5)
Annual property taxes	921.7	1237.7	966.8	753.2	843.9	53.35
	(2104.0)	(2176.2)	(2099.2)	(1909.4)	(2157.3)	(532.9)
Social security retirement payroll deduction	2059.2	3397.6	3145.6	1331.5	92.32	106.5
	(2665.3)	(2305.5)	(2809.6)	(1905.1)	(602.7)	(425.5)
Federal retirement payroll deduction	72.94	304.4	87.28	11.58	3.024	1.128
	(743.1)	(1310.3)	(859.3)	(279.8)	(124.0)	(70.97)
Earned income tax credit (-)	193.8	118.6	255.1	374.7	103.8	32.18
	(814.2)	(613.1)	(922.9)	(1090.8)	(620.7)	(324.4)
Additional child tax credit (-)	45.66	29.98	57.56	76.18	29.31	7.141
	(307.7)	(251.8)	(341.0)	(377.2)	(257.8)	(116.7)
Child tax credit (-)	106.0	180.9	140.1	97.07	37.72	6.981
	(462.4)	(605.6)	(523.1)	(426.2)	(289.3)	(107.1)
Credit received from making work pay (-)	31.86	43.52	42.29	52.81	11.45	4.450
	(145.6)	(169.3)	(164.7)	(181.3)	(94.69)	(41.30)
Federal stimulus payment (-)	30.40	39.84	37.57	51.79	16.00	3.634
	(194.2)	(230.3)	(215.5)	(246.6)	(139.1)	(46.63)

Income from Public Benefits	6063.3 (11977.3)	1534.8 (7618.6)	2050.9 (8360.3)	4821.1 (9711.3)	14227.4 (14028.9)	3805.9 (9938.5)
Supplemental Security Income (SSI)	233.2 (1498.3)	12.13 (386.4)	21.27 (461.4)	76.94 (881.8)	665.1 (2473.2)	153.0 (1113.4)
Welfare (public assistance)	45.09 (604.1)	7.252 (224.3)	17.85 (376.9)	145.1 (1036.3)	91.74 (869.4)	46.34 (570.0)
Person market value of Medicare	1939.6 (10195.7)	925.1 (9304.1)	1102.4 (9759.6)	1016.5 (9120.9)	3733.0 (10902.6)	1239.3 (10926.3)
Person market value of Medicaid	1749.7 (10238.5)	921.6 (9300.3)	1090.1 (9743.9)	1349.2 (9230.7)	3101.3 (11123.8)	1774.0 (11065.0)
Person value of food stamps	128.3 (637.3)	28.75 (292.4)	76.10 (509.8)	357.3 (1093.1)	219.2 (801.6)	155.2 (653.3)
Person value of housing subsidy	7.504 (44.70)	1.771 (18.59)	2.671 (24.77)	14.38 (64.54)	16.58 (66.85)	6.242 (36.87)
Person value of school-lunch subsidy	56.91 (190.4)	41.75 (147.7)	58.45 (187.9)	108.9 (283.0)	53.23 (194.8)	60.63 (162.7)
Person value of energy subsidy	8.045 (65.82)	1.946 (32.22)	3.363 (44.28)	17.03 (91.97)	16.91 (93.80)	4.900 (49.03)
Educational assistance (beyond HS)	138.7 (1442.8)	87.01 (992.2)	140.1 (1383.5)	108.8 (1089.5)	49.04 (949.3)	1607.7 (5068.8)
Social security	2433.6 (5763.4)	177.5 (1791.6)	469.1 (2807.4)	491.9 (2742.4)	6635.8 (7856.4)	172.1 (1353.1)
Unemployment benefits	158.7 (1465.2)	174.9 (1412.4)	122.4 (1189.9)	2008.8 (5292.7)	65.16 (999.4)	43.77 (750.0)
Worker's compensation	62.58 (1186.4)	96.76 (1371.9)	26.83 (721.2)	51.71 (974.3)	120.0 (1711.0)	3.699 (176.7)
Veteran's benefits	158.7 (2141.9)	73.04 (1226.3)	60.48 (1253.9)	74.35 (1200.8)	361.8 (3303.1)	35.94 (980.2)
Disability benefits	115.6 (1888.4)	37.08 (1280.7)	22.23 (863.3)	40.72 (1000.6)	305.1 (3027.3)	32.28 (986.5)
Survivor's benefits	171.5	93.05	85.28	79.01	353.4	45.76

	(2775.0)	(2455.6)	(2390.8)	(2129.6)	(3472.6)	(1574.2)
Private Income	35185.8	59204.0	48040.5	22264.9	11010.4	1996.9
	(36770.9)	(30709.1)	(37984.6)	(27150.8)	(18588.8)	(7061.0)
Alimony	19.07	12.25	19.46	7.355	21.97	6.196
	(758.9)	(471.5)	(741.9)	(427.4)	(875.8)	(413.5)
Non-farm business income	1727.6	135.0	3056.6	580.9	73.73	52.06
	(14536.1)	(3240.0)	(19351.2)	(7167.9)	(2083.0)	(2796.9)
Child support	123.1	152.4	155.9	150.3	64.51	32.67
	(1248.6)	(1361.0)	(1382.9)	(1276.5)	(984.2)	(489.5)
Dividends	365.4	229.6	326.7	85.30	511.1	31.18
	(3206.1)	(1967.1)	(2965.8)	(1214.7)	(3965.3)	(872.9)
Farm	131.2	19.23	232.2	16.56	6.020	0.435
	(4176.1)	(1400.1)	(5563.9)	(1175.8)	(900.2)	(66.06)
Interest	547.8	391.1	436.4	153.7	844.7	30.49
	(3698.1)	(2350.1)	(3182.2)	(1619.9)	(4838.4)	(474.8)
Income from other source not specified	42.91	29.05	26.43	45.32	76.65	5.187
	(1191.5)	(754.7)	(910.9)	(1004.9)	(1660.1)	(196.7)
Rent	282.1	220.2	280.6	117.2	331.1	34.19
	(3800.4)	(3459.2)	(3828.0)	(2586.4)	(4031.6)	(1177.6)
Retirement	1382.0	296.6	471.7	479.1	3361.8	8.916
	(7197.0)	(3251.5)	(4399.5)	(4560.3)	(10827.6)	(496.1)
Wage and salary income	27982.0	53706.1	41697.3	18012.0	1216.5	1447.8
	(45139.8)	(38910.9)	(51473.8)	(33011.8)	(9732.1)	(5643.9)
Assistance from friends/relatives not in HH	42.72	22.17	31.52	63.84	54.28	194.4
	(999.2)	(636.6)	(801.7)	(852.2)	(1185.3)	(2485.5)
Implied value of owner-occupied housing	4157.3	5003.6	4030.7	3292.0	4520.9	161.8
	(5457.7)	(5989.0)	(5377.8)	(5194.5)	(5520.2)	(1338.0)
Observations (individual-year)	486,964	38,039	268,564	12,544	158,628	9,189

Source: CPS-ASEC 1994-2015, CPS-ORG 1994-2015. Notes: Standard deviations presented in parentheses. All means are weighted using sample weights and all dollar amounts are inflated to 2015 dollars.

Table B.4: Main estimates without weights

Specification:	1	2	3
	Outcome: net fiscal impact		
1(union member)	1200.4***	1208.2***	571.8**
	(78.4)	(120.5)	(225.4)
	Outcome: taxes paid		
1(union member)	1013.8***	1148.2***	281.0
	(72.9)	(111.9)	(187.1)
	Outcome: public benefits received		
1(union member)	-186.6***	-60.1	-290.8**
	(29.9)	(41.0)	(125.3)
	Outcome: private income earned		
1(union member)	4268.7***	4442.2***	1468.4***
	(185.7)	(262.0)	(516.2)
Demographics	Yes	Yes	Yes
State-year FE	Yes	Yes	Yes
Individual FE			Yes

Source: CPS-ASEC 1994-2015, CPS-ORG 1994-2015. Notes: Coefficient (within-individual, correlation-corrected SE). Significant at: *10 **5 ***1 percent level. 241,906 observations of 120,953 individuals over 2 consecutive years each. Coefficient estimates on 1(union member) are presented for each {outcome}x{specification} regression model. All dollar amounts are inflated to 2015 dollars.

Table B.5: Effect of imputations on coefficients

Panel A - Specification:	1	2	3	1	2	3
	Net fiscal impact			Taxes paid		
	Main Sample (Drop all imputations)					
1(union member)	1289.8***	1264.3***	540.0**	1108.6***	1240.2***	216.3
	(91.6)	(138.1)	(254.4)	(85.6)	(129.7)	(208.5)
N	241,906					
Dropped N (%)	317,200 (56.73%)					
	Drop Imputed Union Status Only					
1(union member)	949.0***	1017.8***	547.5	722.6***	928.2***	280.2
	(105.0)	(147.8)	(506.1)	(70.5)	(89.8)	(186.8)
N	526,954					
Dropped N (%)	32,152 (5.75%)					
	Drop Imputed Income Only					
1(union member)	1145.3***	1155.9***	369.2	991.4***	1135.3***	319.8
	(82.8)	(117.9)	(252.5)	(75.7)	(107.5)	(205.8)
N	305,750					
Dropped N (%)	253,356 (45.31%)					
	Drop Full-Line Impute Only					
1(union member)	990.8***	1058.1***	99.4	771.6***	950.3***	18.9
	(90.5)	(117.7)	(340.9)	(72.9)	(92.8)	(176.5)
N	478,544					
Dropped N (%)	80,652 (14.43%)					
	Include All Imputations					
1(union member)	919.6***	996.5***	43.0	717.2***	898.6***	155.3
	(99.0)	(136.5)	(447.1)	(67.4)	(83.1)	(174.1)
N	559,106					
Dropped N (%)	0 (0%)					

Panel B - Specification:	1	2	3	1	2	3
	Benefits received			Private income		
	Main Sample (Drop all imputations)					
1(union member)	-181.3***	-24.2	-323.7**	4661.6***	4588.0***	1614.0***
	(35.0)	(47.1)	(144.9)	(205.3)	(302.4)	(575.1)
N	241,906					
Dropped N (%)	317,200 (56.73%)					
	Drop Imputed Union Status Only					
1(union member)	-226.4***	-89.6	-267.4	2834.6***	3517.1***	1119.2**
	(79.5)	(119.2)	(476.4)	(180.2)	(198.3)	(528.9)
N	526,954					
Dropped N (%)	32,152 (5.75%)					
	Drop Imputed Income Only					
1(union member)	-153.9***	-20.6	-49.4	3992.1***	4111.0***	1752.4***
	(34.9)	(52.1)	(148.0)	(188.1)	(262.1)	(580.3)
N	305,750					
Dropped N (%)	253,356 (45.31%)					
	Drop Full-Line Impute Only					
1(union member)	-219.3***	-107.8	-80.5	3195.4***	3808.9***	433.3
	(53.7)	(73.2)	(297.0)	(180.2)	(208.3)	(500.3)
N	478,544					
Dropped N (%)	80,652 (14.43%)					
	Include All Imputations					
1(union member)	-202.4***	-98.0	112.3	2786.8***	3455.3***	805.6
	(74.7)	(108.8)	(417.5)	(170.6)	(186.4)	(499.5)
N	559,106					
Dropped N (%)	0 (0%)					

Source: CPS-ASEC 1994-2015, CPS-ORG 1994-2015. Notes: Coefficient (within-individual, correlation-corrected SE). Significant at: *10 **5 ***1 percent level. Coefficient estimates on 1(union member) are presented for each {outcome}x{specification} regression model. All regressions are weighted using sample weights and all dollar amounts are inflated to 2015 dollars.

Appendix C

Chapter 3

Table C.1: Summary statistics for H.L.E. and Logical Edits, full sample

	H.L.E.	Logical Edits		H.L.E.	Logical Edits
Years in U.S.	14.1	10.3	Age	35.4	32.3
	(8.7)	(6.9)		(8.8)	(12.8)
% >5 years	83.3	70.4	% Female	43.1	44.7
	(37.3)	(45.6)		(49.5)	(49.7)
% Between 1 and 5 years	13.3	22.3	% Employed	71.0	69.7
	(34.0)	(41.6)		(45.4)	(45.9)
% Last year	3.4	7.3	% Unemployed	6.3	6.0
	(18.0)	(26.0)		(24.2)	(23.8)
% Mexico and Central Am.	100.0	68.4	% N.I.L.F.	22.7	24.3
	(0.0)	(46.5)		(41.9)	(42.9)
% Other North Am.	0.0	0.4	% White	58.4	49.9
	(0.0)	(6.5)		(49.3)	(50.0)
% Other Latin Am.	0.0	11.1	% African American	0.6	5.6
	(0.0)	(31.5)		(7.4)	(23.0)
% Europe	0.0	3.3	% Asian	0.1	13.2
	(0.0)	(17.9)		(3.4)	(33.9)
% Asia	0.0	13.8	% Other	40.2	30.7
	(0.0)	(34.5)		(49.0)	(46.1)
% Africa	0.0	2.8	% Married	57.5	45.0
	(0.0)	(16.4)		(49.4)	(49.8)
% Oceania	0.0	0.2	% Single	33.3	45.7
	(0.0)	(4.3)		(47.1)	(49.8)

% Less than H.S.	69.1	52.1	% More than H.S.	0.0	12.0
	(46.2)	(50.0)		(0.0)	(32.5)
% H.S. or equiv.	30.9	23.3	% College or more	0.0	12.0
	(46.2)	(42.3)		(0.0)	(32.6)
Observations	545,209	842,786	Observations	545,209	842,786
Weighted N's	78,010,809	127,813,114	Weighted N's	78,010,809	127,813,114

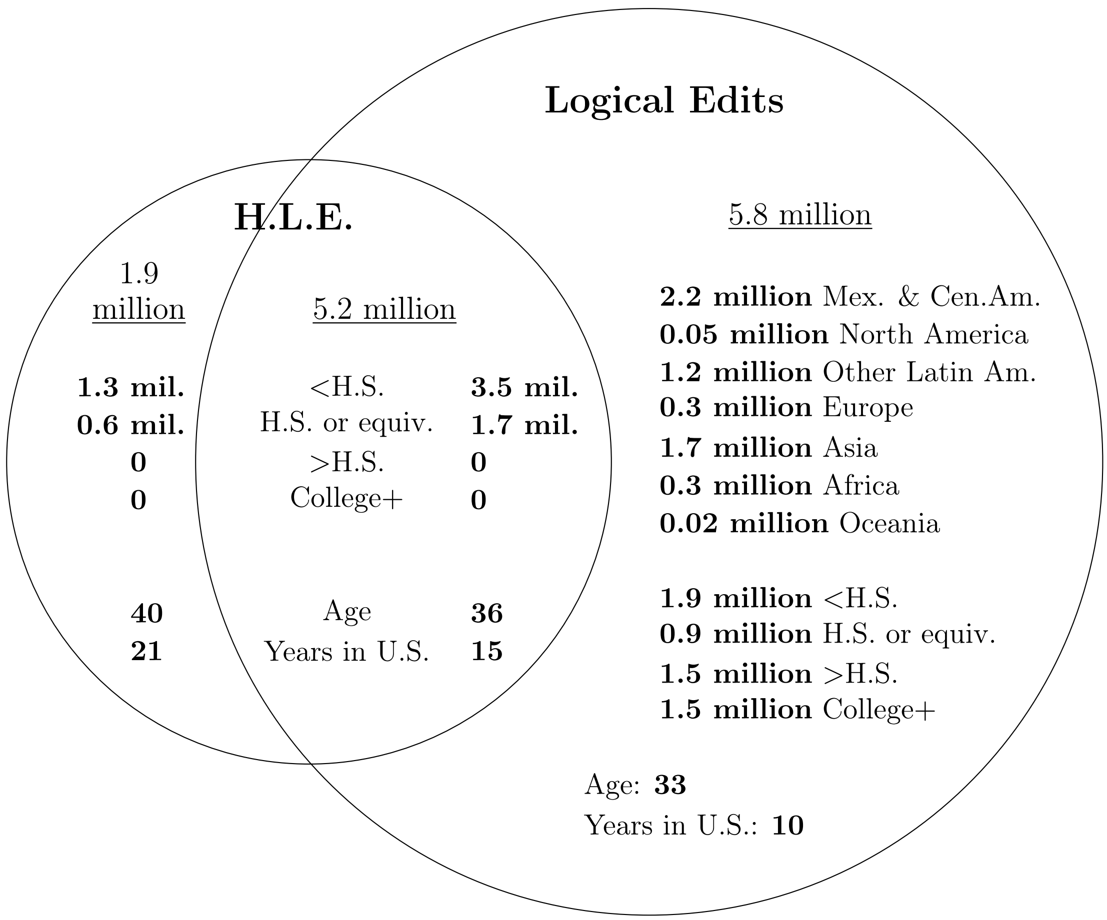
Source: ACS 2005-2015. Note: Averages presented with standard deviations in parentheses. H.L.E. refers to the "Hispanics with Low Education" proxy.

Table C.2: Estimates of the impact of E-Verify mandates on probability of employment, Basic Monthly CPS files

	In Labor Force Population			Full CPS		
	All	Male	Female	All	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)
H.L.E.						
Universal	0.028**	0.033***	0.046*	0.035	0.007	0.062**
	(0.013)	(0.007)	(0.025)	(0.021)	(0.049)	(0.029)
Observations	277,369	189,596	87,773	386,406	212,088	174,318
Naturalized Hispanic						
Universal	0.039	0.119*	-0.041	0.009	0.006	-0.024
	(0.046)	(0.063)	(0.037)	(0.046)	(0.084)	(0.020)
Observations	184,575	98,638	85,937	269,637	124,911	144,726
US-Born non-Hispanic						
Universal	-0.010***	-0.017***	-0.003	-0.003	-0.004	-0.005
	(0.001)	(0.001)	(0.003)	(0.006)	(0.006)	(0.005)
Observations	8,708,946	4,492,400	4,216,546	13,562,270	6,461,583	7,100,687

Source: Basic Monthly CPS 2002-2014. * p<0.1, ** p<0.05, *** p<0.01. Notes: Controls include gender (when applicable), race, age, marital status, number of children in household, educational attainment, industry fixed effects, state fixed effects, time (year, month) fixed effects, state specific time trends, unemployment rates. Standard errors clustered at the state level. All regressions use survey weights (wtsupp).

Figure C.1: Comparing H.L.E. and Logical Edits - Characteristics of likely unauthorized by different proxies - Full population



Notes: H.L.E. refers to the "Hispanics with Low Education" proxy. Estimates are weighted using ACS person weights.