From the Farm to the City:
How the Changing Supply of Immigrant Workers Impacts the United States

By

ZACHARIAH JUDSON RUTLEDGE
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Agricultural and Resource Economics

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Pierre Mérel, Chair

J. Edward Taylor

Giovanni Peri

Committee in Charge

2020

Dedicated to our nation's essential workers, both immigrant and native-born.

Without your dedication and hard work, life as we know it would cease to exist.

Contents

List of Figures

List of Tables

Abstract

From the Farm to the City: How the Changing Supply of Immigrant Workers Impacts the United States

Over the past few decades, the United States (U.S.) has experienced a massive influx of low-skilled immigrants. Between 1990 and 2018, the total immigrant population increased from 20 million (8% of the population) to 45 million (14% of the population).[1] However, not all low-skilled sectors of the economy have consistently experienced positive immigrant labor supply shocks. For example, recent evidence suggests that the agricultural sector has experienced a decline in the supply of immigrant workers.

Mexican and Central American immigrants are by far the largest group of low-skilled immigrants in the U.S., comprising half of the immigrant population with a high school education or less. About 80% of Mexican and Central American immigrants have at most a high school diploma, and they are typically employed in low-skilled sectors of the economy, such as agriculture and construction. Estimates suggest that 11 million unauthorized immigrants reside in the U.S., 8 million of whom are Mexican and Central American. The issue of unauthorized immigration has led to a contentious debate, driving a wedge between Americans. Opponents of immigration argue that these immigrants take American jobs, depress the wages of native-born workers (natives), and drain resources from the social welfare system. Proponents argue that these immigrants take low-paying, physically demanding jobs that Americans don't want, which reduces the cost of goods and services, and that immigrants often contribute to the tax base even if they are unauthorized to work. Economists have failed to come to a consensus on the debate, partly because it is difficult to find empirical settings that lend themselves to producing exogenous variation in the supply of immigrants. It is plausible that elements on both sides of the debate are

[1]The term "immigrant" used here is defined as a person residing in the U.S. who was not a U.S. citizen at birth. These statistics were obtained from the Migration Policy Institute at: `https://www.migrationpolicy.org/programs/data-hub/us-immigration-trends#history`.

valid to some extent, depending on the outcome of interest and the economic sector under consideration.

In this dissertation, I examine how changes in the supply of low-skilled immigrants affect various outcomes in the U.S. and the extent to which these immigrants have been able to achieve economic success. The first chapter examines how a decline in the supply of immigrant farm workers impacts labor-intensive crop production in the state of California. The second chapter investigates how increased immigration impacts native workers in non-farm sectors of the economy. The third chapter documents the extent to which Mexicans and Central Americans have been able to close the earnings and employment rate gap (relative to native workers) over time. As a whole, this dissertation sheds light on how low-skilled immigration creates winners and losers and documents the extent to which immigrants have been successful in assimilating into the U.S. labor market.

Chapter 1 extends the existing farm labor literature, which has found evidence of a declining farm labor supply, by quantifying the impacts such changes have on labor-intensive crop production. Specifically, I provide reduced-form estimates of the effects of shifts in the farm labor supply on the production of hand-harvested fruits and vegetables. Using crop production and employment data from California between 1990 and 2017, I estimate fixed-effects panel regressions linking farm employment (measured at the county-year level) to crop production outcomes (measured at the crop-county-year level). Because I use variation in equilibrium employment, as opposed to exogenous variation in the labor supply, I use an equilibrium displacement model to identify plausible sources of bias that may affect my empirical estimates. This exercise reveals that my point estimates should be interpreted as upper bounds for the effects of interest. Empirically, these bounds indicate that a one percent decrease in the farm labor supply (in terms of the number of workers) causes at most a 0.37% reduction in production in the top 10 producing counties, which together produce 86% of the total value of hand-harvested crops in the state. Production effects are channeled primarily through a reduction in harvested acreage, although there are some effects on yield. I also find that a 1% decrease in the labor supply causes at most a

0.46% decrease in the total value of hand-harvested crop production in the top 5 producing counties (or $600 million). The results from this chapter indicate that a declining farm labor supply could generate economically meaningful consequences for farmers, but that it will likely not devastate the aggregate production of fruits and vegetables in the near future.

Chapter 2 analyzes the impact of immigration on the labor market outcomes of native workers in the U.S.[2] The analysis focuses on workers in U.S. metropolitan statistical areas using U.S. Census and American Community Survey data between 1990 and 2011. We use a set of imperfect instruments to derive new bounds on the short-run impacts of immigration on the earnings, employment rate, and full-time employment rate of natives. We focus on nine sectors with higher immigrant penetration and instrument for the sectoral immigrant share using the immigrant share in all other sectors. We find negative effects of immigration on native earnings in sectors where we would most expect to find them: low-skilled sectors that produce non-traded goods where immigrant penetration has been high in recent decades. We uncover negative effects on native earnings in the construction, food service, and personal service sectors, with upper bounds ranging from -2.9% to -6.6% for each 10 percentage point increase in the immigrant share. Earnings effects in other sectors are not statistically significant. In the six low-skilled sectors we consider, immigration reduces the native employment rate, with effects ranging from -0.6 to -2.0 percentage points for each 10 percentage point increase in the immigrant share. Our findings indicate that increases in the low-skilled immigrant labor supply lead to worse labor market outcomes for some low-skilled native workers in the short run.

Chapter 3 investigates whether recently arrived low-skilled immigrants have been more successful than older cohorts at assimilating into the U.S. labor market.[3] Specifically, we use U.S. Census and American Community Survey data between 1970 and 2017 to examine how different Mexican and Central American cohorts of arrival compare to

[2]This chapter is co-authored with Dr. Pierre Mérel, Professor of Agricultural and Resource Economics at the University of California, Davis.

[3]This chapter is co-authored with Dr. Giovanni Peri, Professor of Economics at the University of California, Davis.

similarly aged and educated natives in terms of their earnings and employment rate over time. We find that, on average, they started with an earnings gap of 40-45 percent and eliminated half of it within 20 years of arrival. Recent cohorts that arrived after 1995 did better than earlier cohorts in terms of the initial gap and the convergence rate. All cohorts achieved employment rates that surpassed that of natives within 20 years of arrival. The most recent cohorts arrived in the U.S. with no employment rate gap. We conclude that Mexican and Central American immigrants enter the U.S. with a strong proclivity toward being employed. However, despite their successful integration into the labor market in terms of achieving gainful employment, they have not reached parity with their native counterparts in terms of earnings.

Acknowledgments

There is no way I can express in words the full amount of gratitude I have for all the individuals who helped me succeed in this Ph.D. program. Many people have touched my life in different ways, but I would like acknowledge some of the most influential people I have had the privilege of interacting with.

First and foremost, I would like to acknowledge and thank my dissertation committee chair, Dr. Pierre Mérel, who spent countless hours and many entire days in one-on-one meetings with me to help push my research projects forward and educate me on the fundamentals of academic research. I will forever be grateful for his commitment to my professional development, his willingness to help me pursue my personal research interests, and for pushing me to think like an economist. Witnessing Dr. Mérel's ability to analyze problems and derive meaningful conclusions from limited information has helped me realize that there is usually a logical answer to every research question. I feel privileged to have had exposure to his ingenious way of dissecting difficult problems and sifting through different layers of data, as well his ability to think through, and overcome, research challenges that, to me, often seemed too difficult to resolve. Dr. Mérel always pushed me to consider the fundamental concepts of economics and econometrics when trying to answer research questions or make sense of unexpected results, and he taught me that there is usually some theoretical or mathematical rationale that can explain what is going on. While Dr. Mérel encouraged me to think critically and solve problems on my own, whenever I ran into a challenge I absolutely needed help with, he was always there to help me resolve the issue. Dr. Mérel has been a constant influence in my life since I entered the program in 2015. Although it took over four years, Dr. Mérel helped me develop, step by step, my oversimplified undergraduate honors thesis project into a solid academic publication. Dr. Mérel also dedicated a tremendous amount of time helping me develop the research material contained in my job market paper, which I am ever so grateful for. Dr. Mérel was always there to double check my math, correct my proofs when they were wrong, consult me about new empirical results, and help me figure out what to do next.

Dr. Mérel usually left his office door slightly cracked when he was working and almost always allowed me to come in unannounced to ask him questions. Even if he was busy at the moment, he always freed up time to meet with me even if that meant coming back a little while later. Dr. Mérel's influence on my life over the past five-plus years cannot be overstated. I am forever indebted to Dr. Mérel for his investment of time and energy in me, which has helped me accumulate a great wealth of human capital during my time in the A.R.E. department at U.C. Davis.

I would also like to acknowledge my committee member Dr. J. Edward Taylor, who not only inspired me to apply to the Ph.D. program in A.R.E. when I was an undergraduate student at U.C. Davis but also provided invaluable support to me as a graduate student. As an undergraduate student, Dr. Taylor helped me develop an interest in academic research by taking me on as an undergraduate advisee and helping me learn to apply basic econometrics to simple research questions. As a Ph.D. student, Dr. Taylor again took me under his wing and helped me develop my research agenda by nudging me to get involved in farm labor research. He also provided me with numerous opportunities to develop relevant research skills, such as survey design when we worked with the C.F.B.F. or how to efficiently review academic articles when selecting participants for a conference. Dr. Taylor helped me gain confidence in my ability to write by allowing me to take an active role in a handful of research projects that were intended for both academic and non-academic audiences. Dr. Taylor also taught me the importance of being able to think more holistically about academic research by taking a step back and considering how all the little pieces serve to answer the larger research question. Dr. Taylor has had a significant positive influence on my life since I took his undergraduate course on development economics. Dr. Taylor's calm demeanor, patience, and ability to empathize with others always made it a pleasure to work with him, and his example is one that I strive to emulate.

I would also like to acknowledge my committee member Dr. Giovanni Peri, who has also served as a mentor to me since I was an undergraduate student. Without Dr.

Peri's guidance as my advisor in the Undergraduate Honors Program at U.C. Davis, it is highly unlikely I would have considered applying to graduate school. Dr. Peri helped me develop a core interest in immigration economics as an undergraduate student and continued to influence my research agenda in graduate school. Dr. Peri provided an avenue for me to learn about immigration research by inviting me to become a student affiliate of the Global Migration Center (formerly the Migration Research Cluster), where I was able to gain exposure to immigration research from a variety of different academic fields. In addition to providing me with many opportunities to expand the scope of my understanding about immigration, Dr. Peri provided an outlet for me to present various stages of my work both at the migration seminars and at brown bag meetings he organized, where he always provided critical input that helped strengthen my research. I also feel very privileged to have had the opportunity to engage in academic research with Dr. Peri, which has allowed me to expand the breadth of my knowledge and experience.

There are a number of other individuals who played a key role in my success, too. Among them are my wife, Rebekah, who provided me with emotional support along the way and allowed me to spend so much time in my office working on research instead of spending quality time with her. She has been with me through thick and thin, has helped me stay focused on the end goal, and has inspired me to keep moving forward during difficult times. I would also like to thank my parents Ron and Joyce and brother Zeke who cheered me on along the way and allowed me to vent when times were tough. I would also like to thank my friend and walking partner Jordan Vaughn, who helped me get out of the office and clear my mind on the weekends when we went on our neighborhood walks. I would also like to thank my former office mates, Heidi Schweizer, Jarrett Hart, and Louis Sears, whose shenanigans made coming to the office fun on many occasions and made the graduate student experience more bearable. I would also like to thank my two colleagues Edward Whitney and Curtis Morrill who provided much needed camaraderie as we took the longer journey through the program as split-core students. I would like to thank our Graduate Program Coordinator, Christy Hansen, who always dropped what

she was doing to answer my questions or resolve some technical issue I was having. I would also like to thank the our department's IT staff, Laurie Warren, Jeff Goettsch, and Arnon Erba, who were always there to help me resolve computer issues. I would like to thank Sara Neagu-Reed and Bryan Little of the California Farm Bureau Federation for partnering with me to conduct a survey to learn more about farm labor scarcity in California. I would like to thank Luc Christiaensen for involving me in the Future of Work in Agriculture project. I would also like to thank Phil Martin for helping me learn more about farm labor issues and for allowing me to get involved in research outside the scope of my dissertation. I would like to thank the Agricultural & Applied Economics Association for providing an outlet for me to share my research and network with other like-minded folks. I would like to thank the Gifford Center for Population Studies for providing financial support to travel to the A.A.E.A. annual meetings. And last but not least, I would like to express a heartfelt sense of gratitude to the Giannini Foundation of Agricultural Economics, who provided a significant amount of financial support, which enabled me to dedicate a substantial amount of time to the research contained in Chapter 1 of this dissertation.

Chapter 1: No Farm Workers, No Food? Evidence from Specialty Crop Production

1.1 Introduction

Farm labor is an essential input in the production of many fruit and vegetable crops because they typically need to be harvested manually. Recent studies document what many view as a worrying decline in the farm labor supply. For instance, Richards (2018) finds that there has been an "insufficient supply [of harvest workers] to meet the demand from firms, even in the steady state equilibrium" and that this issue is "chronic and not merely a feature of our current policy environment." Farmer surveys further suggest that this situation has been exacerbated by recent changes in immigration policy, including tighter border security and stronger internal enforcement (CFBF and UC Davis 2019; Rutledge and Taylor 2019a). A declining farm labor supply has the potential to reduce the nation's access to safe and healthy produce, increase food prices, and cause farmers to suffer significant economic losses, yet few existing studies have examined its implications. In this paper, I help fill the gap by estimating the effects of shifts in the farm labor supply on hand-harvested fruit and vegetable crop production in California, the leading specialty crop producer in the United States (U.S.).

Recent research has attributed the declining farm labor supply in the U.S. to demographic and structural changes in Mexico, more stringent security measures along the U.S. border, and a reduction in the number of workers willing to engage in follow-the-crop

migration (Zahniser et al. 2011; Passel et al. 2012; Hertz and Zahniser 2012; Taylor et al. 2012; Fan et al. 2015; Charlton and Taylor 2016; Zahniser et al. 2018, 2019). Enhanced border security has led to higher coyote (smuggler) fees, which has increased the financial cost of entering the U.S. (Massey et al. 2002; Orrenius 2004; Riosmena 2004; Dickerson and Medina 2017). And those who have attempted to cross the border have been driven further into the desert to avoid apprehension, leading to an increase in fatalities (Jones 2020). In some regions of the U.S., local immigration enforcement policies have caused farm workers to leave local labor markets due to the threat of deportation (Kostandini et al. 2013; Ifft and Jodlowski 2016). Evidence from the National Agricultural Workers Survey reveals an upward trend in the proportion of the farm workforce engaged in non-farm work, suggesting that competition from non-farm employers has added pressure to the farm labor supply (Rutledge and Taylor 2019b). Once farm workers are employed in other sectors of the economy, they are less likely to return to farm work because inter-sectoral movement requires irreversible investments, either in human capital or location, which makes the return to farm work more costly (Richards and Patterson 1998).

At the current wage rate, domestic workers seem reluctant to engage in farm labor and, as a result, are unlikely to replace immigrant workers (Taylor et al. 2012). During the great recession, when unemployment rates were close to 10%, the United Farm Workers (UFW) launched the nationwide "Take Our Jobs" campaign, which offered farm employment to any domestic worker who wanted a job.[4] Although the UFW received thousands of responses, the union's president, Arturo Rodriguez, said that only a few dozen domestic workers followed through on the offer after realizing that the work involved "back-breaking jobs in triple-digit temperatures that pay minimum wage, usually without benefits" (Smith 2010). Economic theory suggests that a shrinking farm labor supply should put upward pressure on wages. According to the USDA's Farm Labor Survey (NASS 2020), real farm wages have increased by 10% over the last decade, and estimates indicate that they will need to increase by an another 10% over the next decade

[4]The UFW is a farm worker union formerly led by Cesar Chavez.

just to keep the labor supply constant (Charlton et al. 2019b). Taken together, this body of evidence points to a farm labor supply that is likely shifting inward.

But how much do changes in the farm labor supply really affect the production and total value of labor-intensive crops? To answer these empirical questions, I estimate elasticities of hand-harvested specialty crop production and value of production with respect to the farm labor supply using data from California spanning the period from 1990 to 2017. My empirical strategy deploys fixed-effects panel regression models at the crop-county-year level of aggregation, where the main explanatory variable of interest is a measure of county-level farm employment during the peak harvest season. The identifying variation is generated from differences across counties in the evolution of employment about the county average net of smooth county-level trends. The fact that county-level crop employment, an equilibrium value, is used as the main explanatory variable in place of a labor supply variable causes several identification challenges.

First and foremost, to the extent that the labor supply curve is upward-sloping (as opposed to being perfectly inelastic) and the labor demand curve is downward sloping (as opposed to being perfectly elastic), a change in employment understates the underlying shock in labor supply (see section 1.3.2.1 for a detailed explanation and graphical illustration). Thus, the output-employment elasticity overstates the elasticity of interest, which is with respect to the underlying labor supply shift. Second, because farm employment is an equilibrium value, it may be influenced by other factors, such as productivity shocks caused by weather or technical change, that affect the outcome independently of labor supply, raising concerns about omitted variables bias. My strategy to cope with the challenge of omitted variables bias is to include a host of control variables in the panel regression, namely crop-by-county fixed effects, year fixed effects, county-level trends, and weather controls.

To gain insight into the remaining bias, I leverage an equilibrium displacement model and investigate several market scenarios depending on the nature of omitted factors. This exercise reveals that my empirical estimates should, in many respects, be interpreted as

upper bounds. Specifically, the equilibrium displacement model allows me to examine how labor supply shocks and other market shocks, such as productivity shocks, jointly affect the equilibrium quantities in the specialty crop and farm labor markets and therefore the relative changes in output and equilibrium employment observed in the data. The model provides a set of reduced-form equations that account for structural relationships between the labor and output markets. These equations can be used to derive formulas for the estimation bias. The bias formulas are functions of structural parameters with unknown magnitudes but known signs, notably elasticities, allowing the sign of the bias to be discussed. Thus, I am able to relate structural parameters descriptive of the specialty crop market to each potential source of bias in a transparent fashion. For example, the model allows me to decipher how changes in productivity, such as those that may result from weather events, or shifts in the supply of non-labor inputs, such as land or water, affect the empirically estimated output-employment elasticity. Therefore, although my regression analysis is reduced-form, it is directly linked to a structural model of fruit and vegetable supply.

My empirical results indicate that a one percent decrease in the farm labor supply (in terms of the number of workers) causes at most a 0.37% decrease in hand-harvested fruit and vegetable production in the top 10 labor-intensive crop producing counties, which together produce 86% of the value of all labor-intensive crops in the state.[5] My findings further suggest that reduced production is primarily channeled through a decrease in the number of acres harvested, but there are also small yield effects. In the top 10 counties, a one percent decrease in the farm labor supply causes at most a 0.23% reduction in harvested acreage and at most a 0.14% reduction in the average yield. An analysis of impacts on the total value of production reveals effects that are concentrated in the top 5 counties. In those counties, a 1% decrease in the farm labor supply causes at most a 0.46% decrease in the total value of production. These results suggest that moderate decreases in the farm labor supply could have meaningful economic impacts, but they would likely not

[5]Throughout this study, a "labor-intensive" crop is defined as a fruit or vegetable crop that did not have a viable automated harvest technology available at any point during the period of study.

devastate California's aggregate fruit and vegetable production. Falsification tests run on mechanically-harvested field and nut crops deliver estimates that are much smaller than those for labor-intensive crops, and are sometimes negative, consistent with the hypothesis that labor supply shocks have a smaller impact on crops that do not rely heavily on manual labor.

To the best of my knowledge, only one recent study examines how negative labor supply shocks affect the aggregate production of fruits and vegetables in the U.S. Zahniser et al. (2011) use a computable general equilibrium model to simulate how a change in the unauthorized immigrant labor force (including both farm and non-farm workers) would affect farm employment and agricultural production, among other things.[6] They find that a policy aimed at increasing immigration enforcement would lead to a 3.4% reduction in farm employment and a 2.0% (respectively 2.9%) reduction in fruit (respectively vegetable) production, implying an upper bound for the elasticity of production with respect to the farm labor supply of 0.58 (respectively 0.85).[7] These upper bounds are larger than those produced by my analysis, although their bound for fruit production is relatively close to the estimate of 0.47 that I find in the top 5 producing counties. Two other recent studies, Brady et al. (2016) and Cassey et al. (2018), use an equilibrium displacement model to examine how simultaneous shocks to output demand and labor supply affect the production of tree fruits rather than the aggregate production of all labor-intensive crops. Although both studies examine the joint effects of two contemporaneous shocks, the elasticities of production with respect to the labor supply can be deduced in each case. The elasticity of aggregate tree fruit production deduced from Brady et al. (2016) ranges from 0.21 to 0.54, which is consistent with the upper bounds produced by my analysis.[8] Cassey et al. (2018) estimate an elasticity of 0.42 for apples, which is consistent

[6]The first policy they consider simulates increased use of the H-2A visa program, and the second policy simulates increased immigration enforcement across the entire U.S. economy, which would arbitrarily lead to a 2.1 million person decrease in the unauthorized workforce.

[7]See section 1.3.2 for details about why these numbers are considered upper bounds.

[8]The aggregate tree fruit production analysis for the state of Washington in Brady et al. (2016) produces results that are very similar to Gunter et al. (1992) because the two studies use the same reduced-form equations and structural parameter values (except for a minor difference in the cost share of the labor input).

with my analysis when I focus on the top 5 producing counties, but their estimate of 0.93 for peaches lies outside the range of potential values I find even when I focus on the more dominant production areas.

This paper contributes to the literature in three ways. First, I extend the existing U.S. farm labor literature, which has found evidence of a declining farm labor supply, by quantifying the impacts such changes may have on hand-harvested fruit and vegetable production in California, the leading agricultural producer in the U.S. (Taylor et al. 2012; Fan et al. 2015; Charlton and Taylor 2016). Recent empirical studies indicate that farm labor shortages may be prevalent throughout the U.S., but farmers are notorious for finding new solutions to their production problems, so farm labor supply shocks are only relevant insofar as they affect the ability to produce crops (Glaister 2006; Plummer 2013; Good 2017; Oatman 2018; della Cava and Lopez 2019; CFBF and UC Davis 2019; Rutledge and Taylor 2019a; Richards 2018; Hertz and Zahniser 2012). By estimating an upper bound for the elasticity of labor-intensive crop production with respect to the farm labor supply, I am able to quantify the potential range of effects and demonstrate that moderate labor supply shocks could have economically meaningful impacts, but that they will likely not devastate the aggregate production of hand-harvested crops.

Second, I directly provide an empirical estimate of the elasticity of labor-intensive crop production with respect to the farm labor supply using a new approach that does not rely on a combination of structural parameter values. Existing studies that examine the impacts of farm labor supply shocks (i.e., Gunter et al. 1992; Brady et al. 2016; Cassey et al. 2018) use the traditional equilibrium displacement approach, which relies on parameter estimates taken from the literature, often from multiple sources, and plugs them into a system of equations (Gardner 1975; Duffy and Wohlgenant 1991; Wohlgenant 1993; Piggott et al. 1995; Gunter et al. 1992). The approach assumes that parameter values are known with certainty and, even when a range of parameter values are used in the analysis, may fail to capture the central tendency of the potential range of effects (Davis and Espinoza

1998).[9] Although I also leverage an equilibrium displacement model, I use it to determine the sign of the estimation bias and proceed to estimate an upper bound for the effect of interest using output, weather, and employment data while imposing little to no structure at the estimation stage.

Third, I demonstrate a novel utilization of the equilibrium displacement framework, which can be adapted to a variety of empirical settings both inside and outside the field of labor economics. The approach I develop delivers insights about the estimation bias that may result from the use of an equilibrium employment variable in place of a labor supply variable, a common problem in labor economics. For example, in the immigration literature, researchers have relied heavily on the use of instrumental variables to produce exogenous variation in the immigrant labor supply (e.g., Borjas 2014; Basso and Peri 2015a; Jaeger et al. 2018a; Mérel and Rutledge 2020), but a commonly-used class of instruments has been found to introduce more bias than the regressor of interest in certain settings (Jaeger et al. 2018a; Goldsmith-Pinkham et al. 2020). The approach I develop offers a new alternative to the use of instrumental variables, allowing researchers to determine if their empirical estimates can be interpreted as bounds for the parameters of interest.

The rest of the paper is organized as follows: section 1.2 provides some background on California crop production and labor, section 1.3 provides a theoretical framework to gain insight into what types of bias may arise and the methodology used in the empirical analysis, section 1.4 describes the data, section 2.5 discusses the results, and section 1.6 provides some concluding remarks.

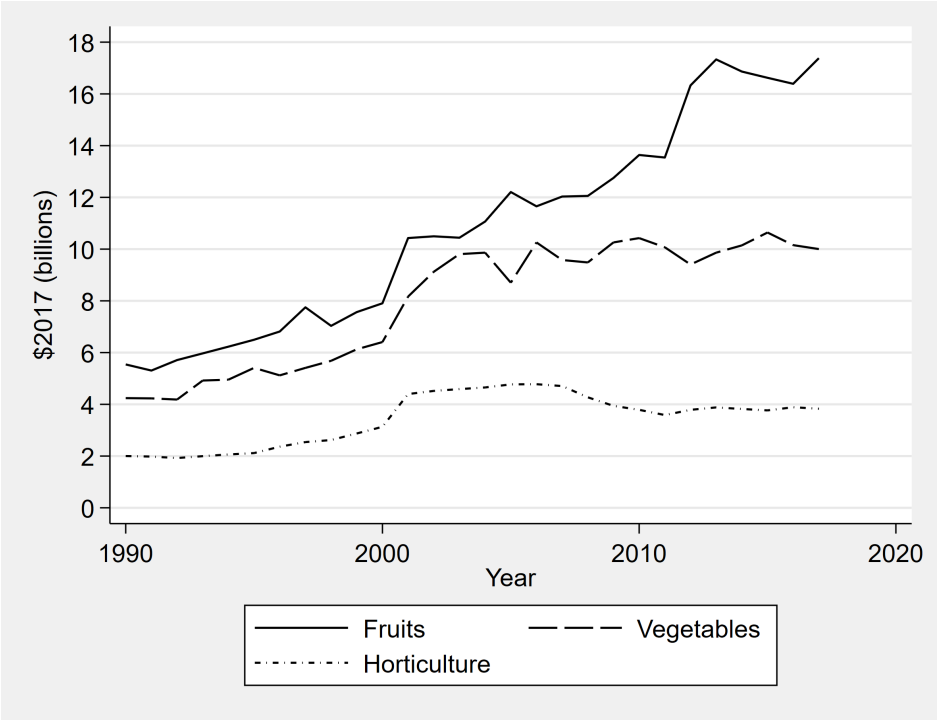
1.2 Background

In terms of the value of production, California is the leading agricultural state in the U.S., generating one third of all domestic vegetables and two thirds of the fruits and nuts (CDFA

[9] Davis and Espinoza (1998) develop an alternative to this approach, which includes assigning random distributions to the structural parameter values based on known priors and identifying the mean, median, and mode of the range of potential effects to identify the centrality of the distribution.

2018). In 2017, California's farms and ranches produced more than 400 commodities worth nearly $60 billion (NASS 2018).[10] Fruit, vegetable, and horticulture (FVH) crop production accounted for 52% ($31.2 billion) of the value of all agricultural production in the state and 68% of non-animal production. Of the total FVH crop value, 56% ($17.4 billion) was generated by fruits, 32% ($10.0 billion) by vegetables, and 12% ($3.8 billion) by horticulture (see Figure 1.1).

Fig. 1.1 Value of FVH Crop Production in California, 1990-2017



Note: Values have been adjusted to real $2017 using the current CPI for the U.S. city average for all items found at: `https://www.bls.gov/cpi/data.htm`.

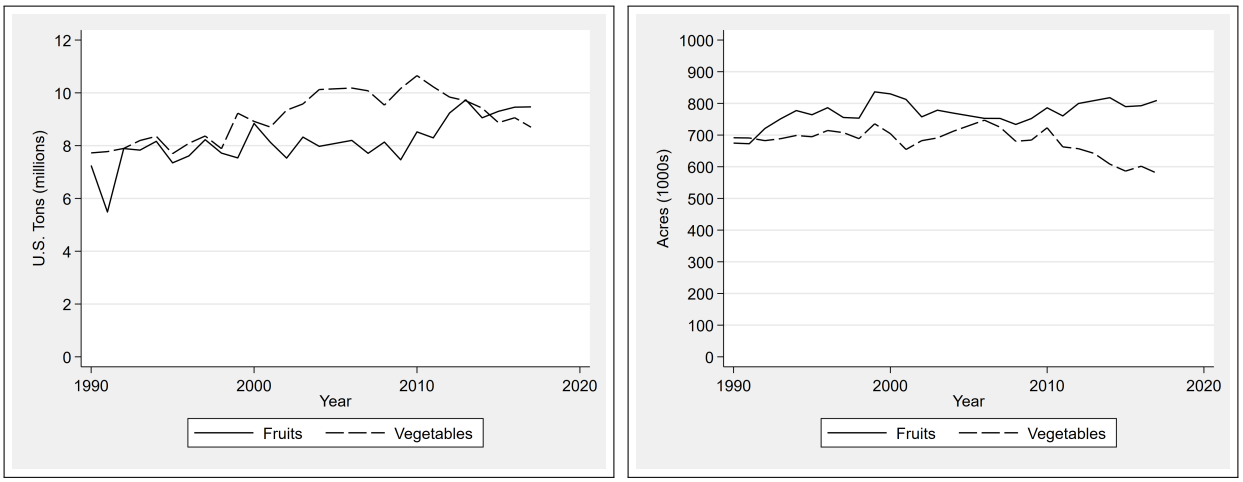
California is also the largest agricultural employer with labor expenses accounting for nearly one third of the nation's total (NASS 2019). California state law requires every county to appoint an agricultural commissioner, who compiles an annual report of the gross production, harvested acreage, and value of each commodity produced (California Agricultural Commissioners and Sealers Association 2020). Each year, the California De-

[10]This figure differs from the $50 billion in cash receipts due to the fact that some production, such as cattle feed, is used on the farm where it is produced. However, the value of fruit and vegetable crops should reflect cash receipts because these crops are not directly consumed by the farms that produce them.

partment of Food and Agriculture collaborates with the U.S. Department of Agriculture to consolidate the county-level statistics, providing a rich set of quantitative crop production data that can be utilized by researchers. These factors make California an ideal setting to study how changes in the farm labor supply affect labor-intensive crop production.

Fig. 1.2 Labor-Intensive Fruit and Vegetable Production
and Harvested Acreage in California, 1990-2017

A: Production B: Harvested Acres



The statewide production of hand-harvested vegetables increased steadily from about 8 million to 10 million tons between 1990 and 2010, and then started to decline (see Figure 1.2.A). The upward trend in vegetable production prior to 2010 was driven by a higher average yield rather than an expansion in acreage (see Figure 1.2.B).[11] Other than a significant drop in production in 1991, which was driven by a winter freeze that devastated orange crops (Brooks 1991), hand-harvested fruit production remained relatively stable at about 8 million tons until 2010, after which a noticeable expansion occurred.[12]

Over the past century, two major policy events have motivated economists to investigate the effects of farm labor supply shocks on U.S. agriculture. The first event was

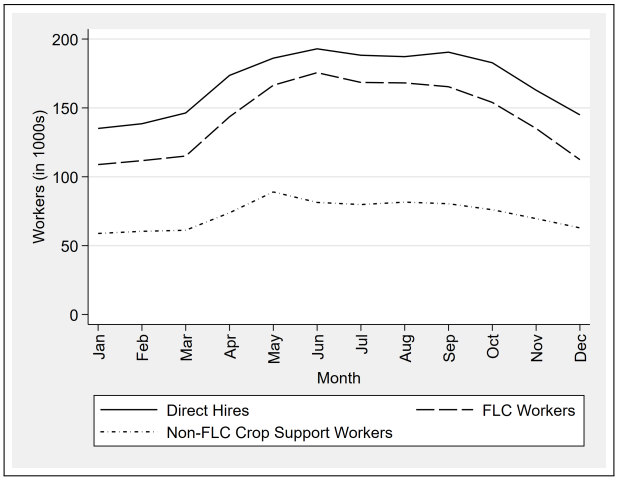
[11] I use a within crop-county fixed-effects estimator in my empirical analysis, so my results are not confounded by changes in the crop mix within, or across, counties.

[12] These numbers refer to crops that are only hand harvested. Crops that are both hand picked and mechanically harvested (such as wine and raisin grapes), as well as crops that are only mechanically harvested (such as potatoes and processing tomatoes) are not included in these calculations. Production data for horticulture crops are not expressed in terms of weight in the CDFA Agricultural Commissioners' Reports and are not included in these calculations.

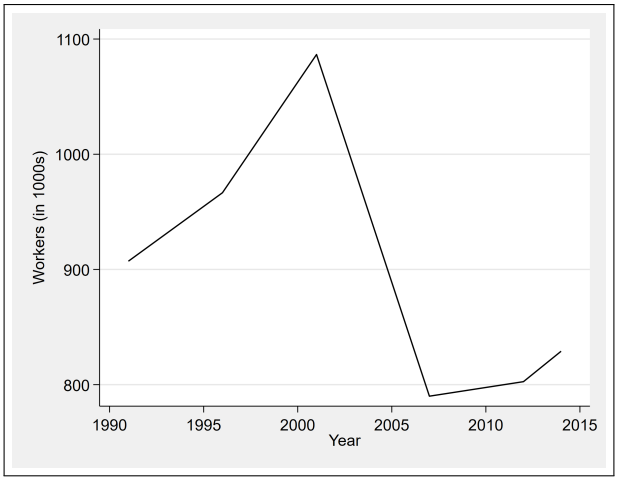
the termination of the Bracero Program in 1964, which ended an era of legal temporary migration from Mexico. The second event was the passage of the Immigration Reform and Control Act (IRCA) in 1986, which, among other things, legalized the unauthorized farm workforce. In each case, major shocks to the farm labor supply failed to materialize, and aggregate crop production was not seriously affected (Shultz 1965; Martin 1966; Duffield 1990; Gunter et al. 1992). Labor supply shocks were short lived in the post-Bracero years due to a sustained inflow of unauthorized Mexican workers who could earn wages as much as eight times higher in the U.S. (Martin 2006). During the post-IRCA years, the number of legalized farm workers who left for other sectors of the economy, the main concern from a policy standpoint, was met by an equal (or greater) inflow of new farm workers as a result of family reunification policies that granted visas to the spouses and dependents of those legalized under the law (Boucher et al. 2007).

Fig. 1.3 California Crop Employment Measures

A: Average Monthly Employment by Worker Type, 2017



B: Reported Farm Worker Social Security Numbers in California, 1991-2014



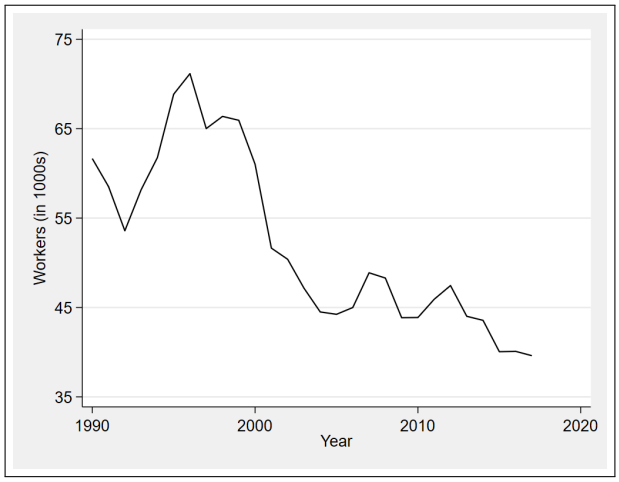
California's farm workers can be classified into three broad categories: (i) those who are recruited and hired directly by farmers (direct hires), (ii) those who are hired by farm labor contractors (FLCs) and are brought to farms to perform certain tasks (e.g., pruning, weeding, or harvesting), and (iii) non-FLC crop-support workers who are contracted to perform certain tasks, such as tilling the soil or providing mechanical harvesting services.

The non-FLC crop-support workers generally do not hand-harvest fruit and vegetable crops and, as a result, are not considered in the analysis. The direct hires and FLC workers *do* perform hand harvest labor and are the focus of this study.

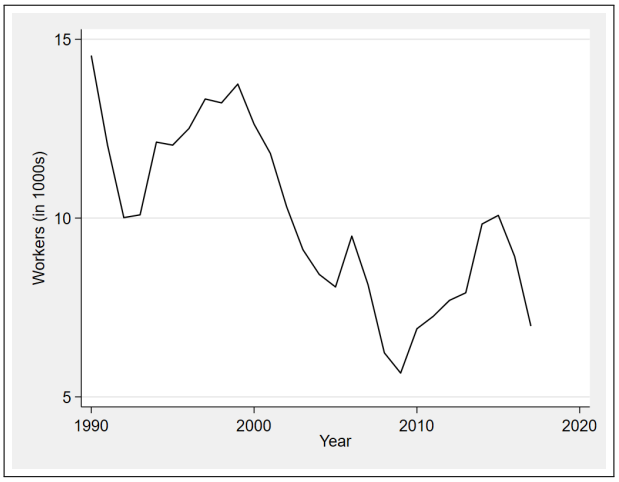
In 2017, California crop farmers employed an average of 386,000 workers each month (BLS 2018). However, due to the seasonal nature of agriculture, the number of workers employed at any given time fluctuates throughout the year. Figure 1.3.A shows the average crop employment for each month during 2017, broken down by type of worker, revealing that statewide employment peaks during the summer months when the bulk of the harvest activities take place. Figure 1.3.B shows the number of social security numbers reported by California farm workers for various years between 1991 and 2014 (Khan et al. 2004; Martin et al. 2016). If the number of social security numbers provides an accurate reflection of the number of workers employed on farms, then these statistics suggest that the number of farm workers employed in the state has declined since the turn of the millennium. However, it is unclear if a count of the social security numbers used by farm workers provides an accurate depiction of the number of workers because some undocumented workers may utilize more than one social security number and some social security numbers may be used by more than one worker. In addition, statewide statistics mask significant heterogeneity among local labor markets, as can be seen by the administrative employment data provided by the Quarterly Census of Employment and Wages in Figure 1.4.

Fig. 1.4 Average Direct Hire and FLC Employment
During the Peak Employment Quarter for Select Counties, 1990-2017

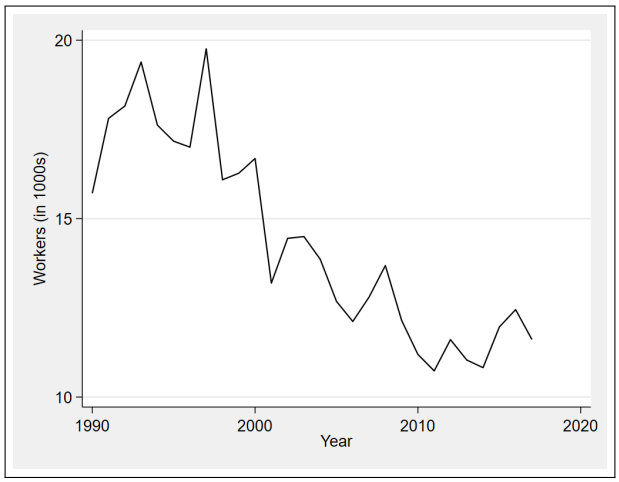
A: Fresno County



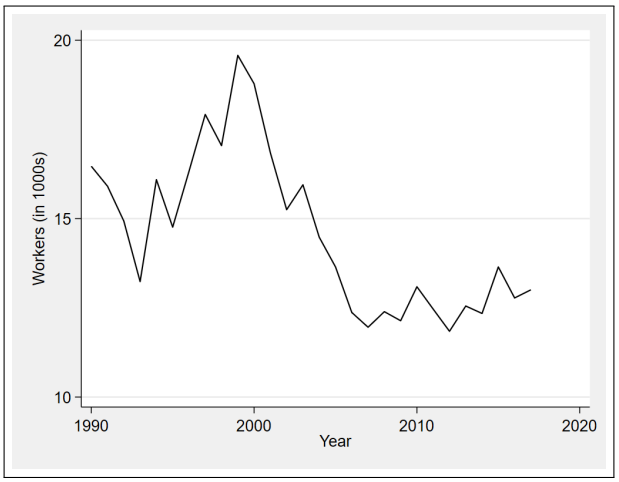
B: Imperial County



C: Riverside County



D: San Joaquin County



1.3 Theory and Methodology

1.3.1 The Model

To provide a theoretical framework that can be used to evaluate the effects of market and technology shocks on the equilibrium outcomes, I develop an equilibrium displacement model and use it to analyze the expected direction of bias caused by the use of an equi-

librium employment variable in place of labor supply variable and omitted variables.[13] I use the model to derive formulas for the expected bias of my empirical estimates under three scenarios that allow for different sets of shocks to the equilibrium. These derivations reveal that the empirical estimates can be interpreted as upper bounds for the parameters of interest in each case.

The model assumes that farmers are price takers in both the input and output markets and that they produce a single homogeneous labor-intensive fruit and vegetable good using two inputs: labor and a composite non-labor input. I define Q as a county's fruit and vegetable output, which is generated by a production function that is homogeneous of degree one, where A and B are the labor and non-labor inputs, respectively.

The model describes a county-level industry equilibrium characterized by six equations in six endogenous variables (Q, p, A, B, p_A, p_B) that define (1.1) the industry demand function, (1.2) the industry production function, two equations (1.3) and (1.4) that equate the marginal value product of each input to its respective price, and two equations (1.5) and (1.6) that define the input supply functions as follows:

$$Q = f(p) \tag{1.1}$$

$$Q = Q(A, B) \tag{1.2}$$

$$p_A = pQ_A \tag{1.3}$$

$$p_B = pQ_B \tag{1.4}$$

$$A = g(p_A) \tag{1.5}$$

$$B = h(p_B), \tag{1.6}$$

where p denotes the output price, p_A (respectively p_B) denotes the price of input A (respectively B), and Q_A and Q_B denote the partial derivatives of the production function with respect to inputs A and B, respectively.

[13]The model I develop builds off of the framework of Muth (1964).

I consider three types of shocks to the equilibrium. The first type of shock, defined as β, represents a shift in the labor supply (in percentage terms) in the direction of the quantity axis at a given price. The second type of shock, defined as γ, represents a shift in the non-labor input supply (in percentage terms) in the direction of the quantity axis at a given price. Positive values of β and γ represent increases in the supply of labor and non-labor inputs, respectively.[14] Examples of non-labor input supply shocks include decreases in the amount of agricultural land due to urban expansion and restrictions on surface water deliveries during periods of drought. The third shock, defined as δ, is a productivity shock.[15] δ can be expressed in logarithmic differential form as follows:

$$\delta = dQ_A^*\Big|_{dA^*=dB^*=0} = dQ_B^*\Big|_{dA^*=dB^*=0} = d\ln(\delta'),$$

where $d\ln(\delta')$ represents the percentage change in the marginal product of both inputs at a given level of input use. These types of productivity shocks can be caused by weather events, such as spring freezes, extreme heat, and rain during the pollination period, or through the diffusion of productivity-enhancing technologies like smart irrigation, all of which can shift the production function and impact the productivity of labor and non-labor inputs.

It is convenient to express changes in the endogenous variables in terms of percentage changes so that the coefficients in the reduced-form equations can be interpreted as elasticities. By taking the total derivative of (1.1), (1.5), and (1.6) and dividing each equation by its respective left-hand-side variable, one can derive equations (1.1′), (1.5′), and (1.6′). Using the homogeneity assumption defined above, additional manipulation of (1.2) - (1.4) leads to the derivation of equations (1.2′) - (1.4′).[16]

[14]Note that the β and γ used here differ from those used in Muth (1964). Here I define input supply shocks as shifts in the direction of the quantity axis at a given price, whereas Muth (1964) defines them as shifts in the direction of the price axis at a given quantity such that an increase in the supply of an input corresponds to a decrease in its price.

[15]Muth (1964) refers to δ as a "factor-neutral" productivity shock.

[16]These derivations are outlined in Muth (1964), although I have modified the model to consider fewer shocks to the equilibrium.

$$dQ^* - \eta dp^* = 0 \tag{1.1'}$$

$$dQ^* - k_A dA^* - k_B dB^* = \delta \tag{1.2'}$$

$$-dp^* + \frac{k_B}{\sigma} dA^* - \frac{k_B}{\sigma} dB^* + dp_A^* = \delta \tag{1.3'}$$

$$-dp^* - \frac{k_A}{\sigma} dA^* + \frac{k_A}{\sigma} dB^* + dp_B^* = \delta \tag{1.4'}$$

$$dA^* - e_A dp_A^* = \beta \tag{1.5'}$$

$$dB^* - e_B dp_B^* = \gamma \tag{1.6'}$$

Together (1.1′) - (1.6′) define a new system of equations, where $dX^* = d\ln(X)$ represents the percentage change in some variable X, and the following structural parameters are taken to be exogenous: η is the elasticity of demand for the final output, σ is the elasticity of substitution between inputs A and B, k_A (respectively k_B) is the production cost share of input A (respectively B), and e_A and e_B are the supply elasticities of inputs A and B, respectively. This system of equations can be solved for each of the six endogenous variables, generating six reduced-form equations.

The demand for fruits and vegetables at the county level is likely highly elastic. For example, I show in Appendix 1.A that even if a county produces 5% of the total output, the elasticity of demand facing that county will be at least 20 times larger than the aggregate demand elasticity, and can potentially be much larger if the supply of the other regions is not perfectly inelastic. Recent studies suggest that the aggregate demand elasticity for labor-intensive crops lies somewhere in the range of -0.59 to -2.44 (Brady et al. 2016; Cassey et al. 2018). Because the average county in my sample produces less than 5% of the national output, the relevant demand elasticity for this study is likely at least -11.8 (20 × -0.59) and may potentially be much larger. In order to simplify the model and keep the mathematics tractable, I assume the demand for output at the county level is perfectly elastic (i.e., $\eta \to -\infty$) in the derivations that follow. If I relax the perfectly elastic

output demand assumption and allow $\eta > -\infty$, the conclusions about the direction of bias remain unchanged as long as $\eta \leq -1$ and $\eta \leq -\sigma$, which seems plausible given the likely large magnitude of η and the fact that the substitutability between hand-harvest labor and non-labor inputs is limited under the current state of technology. Under the perfectly elastic output demand assumption, the reduced-form equations for dQ^* and dA^* can be expressed as

$$dQ^* = \overbrace{\left[\frac{k_A(\sigma + e_B)}{D}\right]}^{\xi_1}\beta + \overbrace{\left[\frac{k_B(\sigma + e_A)}{D}\right]}^{\xi_2}\gamma + \\ + \underbrace{\left[\frac{\sigma(1 + k_A e_A + k_B e_B) + k_B e_A + k_A e_B + e_A e_B}{D}\right]}_{\xi_3}\delta \tag{1.7}$$

and

$$dA^* = \underbrace{\left[\frac{\sigma + k_A e_B}{D}\right]}_{\rho_1}\beta + \underbrace{\left[\frac{k_B e_A}{D}\right]}_{\rho_2}\gamma + \underbrace{\left[\frac{(\sigma + e_B)e_A}{D}\right]}_{\rho_3}\delta, \tag{1.8}$$

where

$$D = \sigma + k_B e_A + k_A e_B \geq 0. \tag{1.9}$$

To simplify the notation, the coefficients on β, γ, and δ in equation (1.7) (respectively equation (1.8)) are denoted by ξ_1, ξ_2, and ξ_3 (respectively ρ_1, ρ_2, and ρ_3). Because β (respectively γ) is a percentage change, it can be expressed in logarithmic differential form as $d\ln(\beta')$ (respectively $d\ln(\gamma')$), where $\ln(\beta')$ and $\ln(\gamma')$ represent the unobserved (log) farm labor supply and non-labor input supply variables, respectively.

The relevant concern from a policy standpoint is with regard to how decreases in the supply of Mexican farm workers, which may result from increased border enforcement or an expanding Mexican economy, affect U.S. crop production. As a result, I am not particularly interested in how changes in equilibrium farm employment affect production because (i) variation in employment may be driven by fluctuations in labor demand

and (ii) even if the demand for labor is stable, under commonly accepted assumptions, changes in employment will not accurately reflect changes in labor supply. In light of these considerations, the parameter of interest is the elasticity of labor-intensive crop production with respect to the farm labor supply:

$$\frac{\partial Q^*}{\partial \beta} = \frac{\partial \ln(Q)}{\partial \ln(\beta')} = \frac{k_A(\sigma + e_B)}{\sigma + k_B e_A + k_A e_B} = \xi_1.$$

If the elasticity of substitution between labor and the composite non-labor input (σ), the input supply elasticities (e_A and e_B), and the cost shares (k_A and k_B) were known, a simple exercise could determine the value of ξ_1, and there would be no need to conduct an empirical analysis. However, the parameter values are not known with certainty, and estimates tend to vary widely. For example, the farm labor supply elasticity estimates used in the most recent equilibrium displacement studies range from 0.71 to 3.37 (Brady et al. 2016; Cassey et al. 2018). The matter is further complicated by the fact that the non-labor input is a composite good, making it difficult (or impossible) to determine the relevant supply and substitution elasticities. An examination of equation (1.7) reveals that if $e_A \geq e_B$, then ξ_1 is bounded by the interval $[0, 1]$. However, in the case where $e_A \leq e_B$, ξ_1 is bounded by the interval $[k_A, 1]$. Martin et al. (2016) estimate that $k_A \in [0.2, 0.4]$, which implies that when the supply of labor is less elastic than the supply of the non-labor input, $\xi_1 \in [0.2, 1]$. Even in this special case, the range of potential values is wide enough that an empirical estimate would provide useful information and, as a result, there is value in performing an empirical exercise.

If β' were observed, direct estimates of ξ_1 could be obtained by using ordinary least squares (OLS) regression with a (log) production variable as the outcome variable and $\ln(\beta')$ as the regressor of interest possibly controlling for factors one may think could be correlated with labor supply shocks. To the extent that the labor supply shocks are uncorrelated with the other shocks (δ and γ), conditional on the included covariates, the resulting estimate would be unbiased. However, when ξ_1 is estimated using an

employment variable (A) instead of a labor supply variable (β'), the regression coefficients are typically biased, so understanding the direction of bias is critical for inference. If the bias is positive, the empirical estimates can be interpreted as upper bounds. Fortunately, the equilibrium displacement model provides a framework from which one can derive formulas for the estimation bias under different scenarios. In the sections that follow, I use the model to derive bias formulas under the assumption that there are no productivity shocks or non-labor input supply shocks (i.e., $\delta = \gamma = 0$) and then consider more general cases where there are (i) unobserved productivity shocks (i.e., $\delta \neq 0$) and (ii) unobserved productivity shocks and non-labor input supply shocks (i.e., $\delta \neq 0$ and $\gamma \neq 0$). By starting with a simple case and examining cases that are progressively more complex, I am able to demonstrate that the use of an equilibrium employment variable in place of a labor supply variable may generate estimates that are biased upwards even in the absence of technology and non-labor input supply shocks, and that each additional shock adds another source of potential upward bias. These findings indicate that my empirical estimates should be interpreted as upper bounds for the parameter of interest, ξ_1.

1.3.2 Estimation Bias

1.3.2.1 Case I: Labor Supply Shock Only

First, I consider the case where there are no omitted variables. In this case $\beta \neq 0$, but $\gamma = \delta = 0$. By using equation (1.8) to solve for β and substituting the formula for β into (1.7), one can express the production-employment relationship as

$$dQ^* = \frac{\xi_1}{\rho_1} dA^*. \tag{1.10}$$

By integrating (1.10), the production-employment relationship can be expressed as

$$\ln(Q) = c + \frac{\xi_1}{\rho_1} \ln(A),$$

where c is the constant of integration. This relationship can be estimated empirically using the following OLS regression model:

$$\ln(Q) = c + \Gamma \ln(A) + \nu, \tag{1.11}$$

where ν is the error term satisfying the condition $E[\nu|\ln(A)] = 0$. The OLS coefficient on the (log) equilibrium employment variable has a probability limit equal to

$$\Gamma_{OLS} = \frac{\xi_1}{\rho_1} = \left[\frac{k_A(\sigma + e_B)}{\sigma + k_A e_B}\right] = \xi_1 + \underbrace{\left[\frac{k_A k_B e_A(\sigma + e_B)}{(\sigma + k_A e_B)(\sigma + k_B e_A + k_A e_B)}\right]}_{\theta_1},$$

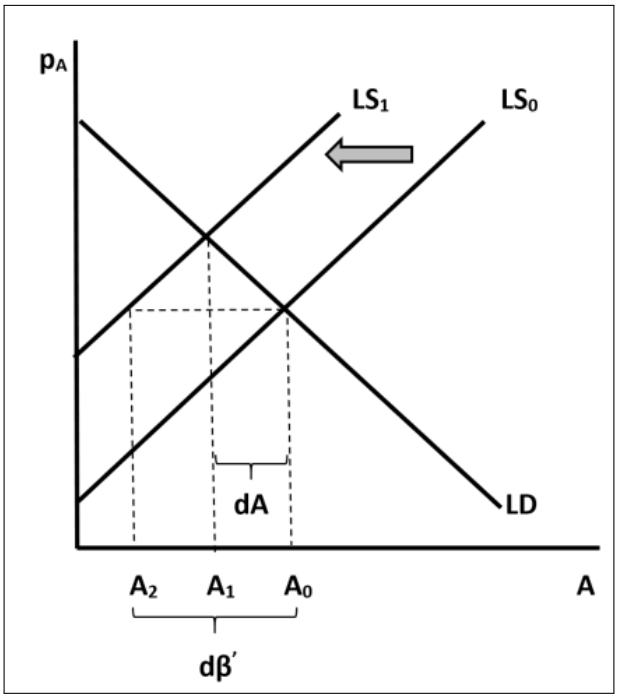
where ξ_1 is the parameter of interest, and $\theta_1 \geq 0$ represents the bias from using an equilibrium employment variable to estimate the effect of a change in the labor supply.[17] The fact that $\theta_1 \geq 0$ results from the fact that $\xi_1 \geq 0$ and $\rho_1 \in [0, 1]$ (see equations (1.7) and (1.8)). Therefore, if there are no omitted variables, Γ_{OLS} can be interpreted as an upper bound for ξ_1. It is important to note that Γ_{OLS} is only biased when the labor supply curve is not perfectly inelastic (i.e., $e_A > 0$) and the labor demand elasticity, defined as η_A below, is finite (i.e., $\eta_A > -\infty$). Figure 1.5 provides a graphical depiction of the local farm labor market to help demonstrate why this result emerges.

As can be seen in Figure 1.5.A, when the demand for labor is downward sloping and the supply of labor is upward sloping, a horizontal shift in the labor supply from LS_0 to LS_1 corresponds to a shift that is larger in magnitude than the change in equilibrium employment (i.e., $|d\beta'| = |A_2 - A_0| > |A_1 - A_0| = |dA|$). As a result, the effect of the labor supply shock the size of $d\beta'$ will be attributed to a smaller change in A, which will cause the empirical estimates to overstate the effect of the labor supply shock. When the labor supply is perfectly inelastic (i.e., $e_A = 0$), as shown in Figure 1.5.B, the shift from LS_0 to LS_1 causes an equivalent change in employment (i.e., $|dA| = |d\beta'| = |A_1 - A_0|$), so there is no bias. This same result emerges when labor demand is perfectly elastic (i.e., $\eta_A \to -\infty$)

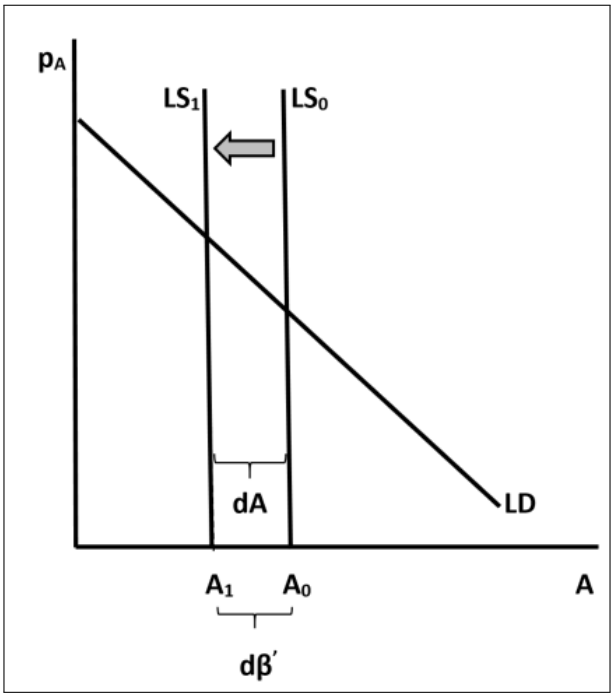
[17] I refer to this source of bias as the "employment-labor supply mismatch bias" in following sections.

Fig. 1.5 The Farm Labor Market Under Different Supply and Demand Elasticities

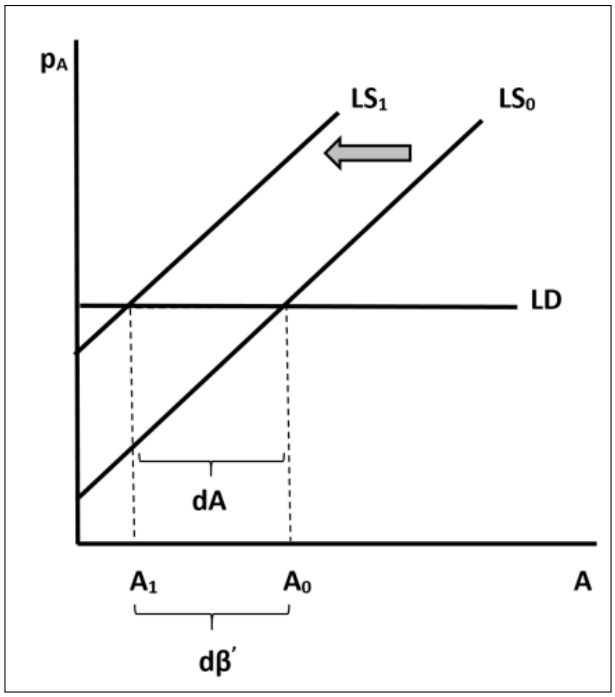
A: Upward-Sloping Labor Supply and
Downward Sloping Labor Demand



B: Perfectly Inelastic Labor Supply and
Downward Sloping Labor Demand



C: Upward Sloping Labor Supply and
Perfectly Elastic Labor Demand



as shown in Figure 1.5.C. The formula for the labor demand elasticity can be expressed as follows:

$$\eta_A \equiv \frac{\frac{\partial A^*}{\partial \beta}}{\frac{\partial p_A^*}{\partial \beta}} = -\frac{\sigma + k_A e_B}{k_B},$$

where the reduced form equation for dp_A^* is

$$dp_A^* = -\left[\frac{k_B}{D}\right]\beta + \left[\frac{k_B}{D}\right]\gamma + \left[\frac{\sigma + e_B}{D}\right]\delta,$$

and D is defined in equation (1.9). The demand for labor will be highly elastic if the cost share of input B (k_B) is small, if labor and non-labor inputs are highly substitutable (i.e., σ is large), or if the supply of the non-labor input is highly elastic (i.e., $e_B \gg 1$). The demand for labor is perfectly elastic if labor and the composite non-labor input are perfect substitutes (i.e., $\sigma \to \infty$), if the non-labor input supply is perfectly elastic (i.e., $e_B \to \infty$), or if labor is the only input (i.e., $k_B = 0$).

The formula for the relative bias can be expressed as the ratio of the labor supply elasticity to the negative of the labor demand elasticity:

$$\frac{\theta_1}{\xi_1} = \frac{k_B e_A}{\sigma + k_A e_B} = -\frac{e_A}{\eta_A} \geq 0. \tag{1.12}$$

Mathematically, the absence of bias when the labor supply is perfectly inelastic or when the labor demand is perfectly elastic can be seen by examining equation (1.12) and comparing the cases where $e_A = 0$ or $\eta_A \to -\infty$ to the case where $e_A > 0$ and $\eta_A > -\infty$. All else equal, (1.12) will tend to be smaller if the supply of labor is highly inelastic (i.e., $e_A \ll 1$) or the demand for labor is highly elastic (i.e., $\eta_A \ll -1$).

1.3.2.2 Case II: Unobserved Productivity Shocks

Now I consider the case where there are unobserved productivity shocks (δ). In this case, $\beta \neq 0$ and $\delta \neq 0$, but $\gamma = 0$. Using the same procedure outlined in section 1.3.2.1, one can

derive the following production-employment equation under this scenario:

$$\ln(Q) = c + \frac{\xi_1}{\rho_1}\ln(A) + \Lambda \ln(\delta'),$$

where c is the constant of integration, and $\Lambda \geq 0$ (see Appendix 1.B for proof) is defined as follows:

$$\Lambda \equiv \left[\frac{\xi_3\rho_1 - \xi_1\rho_3}{\rho_1}\right] \geq 0.$$

The elasticity of labor-intensive crop production with respect to the labor supply can be estimated empirically using (1.11), but in this case, the error term (ν) contains the term $\Lambda \ln(\delta')$, and $E[\nu|\ln(A)] \neq 0$ because equilibrium employment is affected by productivity shocks via equation (1.8). Under this scenario, the coefficient on the (log) equilibrium employment variable has a probability limit equal to

$$\Gamma_{OLS} = \xi_1 + \theta_1 + \underbrace{\left[\frac{\text{cov}(\ln(A), \Lambda \ln(\delta'))}{\text{var}(\ln(A))}\right]}_{\theta_2},$$

where $\text{cov}(X, Y)$ represents the covariance between two variables X and Y, $\text{var}(X)$ represents the variance of some variable X, and θ_2 represents the bias from the omitted productivity shock variable. In this case, Γ_{OLS} will be subject to two sources of bias: one from the use of an equilibrium employment variable in place of a labor supply variable (θ_1) and one from the unobserved productivity shocks (θ_2). By integrating equation (1.8), one can obtain formulas for $\ln(A)$ and $\text{var}(\ln(A))$ and substitute them into θ_2. Even under the assumption that labor supply shocks (β') are uncorrelated with productivity shocks (δ'), there will be omitted variables bias because the model is estimated with an equilibrium employment variable (A) instead of a labor supply variable (β'). Under this assumption, $\text{cov}(\ln(\beta'), \ln(\delta')) = 0$, and the relative bias from the omitted productivity shock variable can be expressed as

$$\frac{\theta_2}{\xi_1} = \frac{\Lambda}{\rho_3 \xi_1} \left[\frac{\rho_3^2 \text{var}(\ln(\delta'))}{\rho_1^2 \text{var}(\ln(\beta')) + \rho_3^2 \text{var}(\ln(\delta'))} \right] \geq 0. \tag{1.13}$$

In this case, the omitted variables bias is non-negative because $\Lambda \geq 0$, $\rho_3 \geq 0$ (see equation (1.8)), $\rho_1^2 \geq 0$, $\rho_2^2 \geq 0$, and the variances are non-negative. All else equal, (1.13) will tend to be smaller if the variance of the labor supply shocks ($\text{var}(\ln(\beta'))$) is large relative to the variance of the productivity shocks ($\text{var}(\ln(\delta'))$). Furthermore, it can be shown that $\Lambda/(\rho_3\xi_1) = \left[\xi_3/(\rho_3\xi_1) - 1/\rho_1\right]$, so the relative bias is also small if the elasticity of production with respect to the productivity shocks (ξ_3) is small or if the elasticity of production with respect to the labor supply (ξ_1) is large.

The results from a simulation exercise that compares the relative magnitude of θ_1 to θ_2 under a range of structural parameter values can be found in Appendix 1.D. This exercise reveals that the two sets of potential values for θ_1 and θ_2 are not disjoint and, as a result, additional information about the structural parameter values and variances is required to determine whether the omitted variables bias (θ_2) is more of a concern than the employment-labor supply mismatch bias (θ_1).

1.3.2.3 Case III: Unobserved Productivity Shocks and Non-Labor Input Supply Shocks

Now I consider the more general case where there are unobserved productivity shocks (δ) and non-labor input supply shocks (γ). In this case, $\beta \neq 0, \delta \neq 0$, and $\gamma \neq 0$. Using the procedure outlined in section 1.3.2.1, one can derive the following production-employment equation under this scenario:

$$\ln(Q) = c + \frac{\xi_1}{\rho_1} \ln(A) + \Lambda \ln(\delta') + \Upsilon \ln(\gamma'),$$

where c is the constant of integration, and $\Upsilon \geq 0$ (see Appendix 1.C for proof) is defined as follows:

$$\Upsilon \equiv \left[\frac{\xi_2 \rho_1 - \xi_1 \rho_2}{\rho_1} \right] \geq 0.$$

The elasticity of labor-intensive crop production with respect to the labor supply can be estimated empirically using (1.11), but under this scenario, the error term (ν) contains the two terms $\Lambda \ln(\delta')$ and $\Upsilon \ln(\gamma')$, and $E[\nu | \ln(A)] \neq 0$ because productivity shocks and non-labor inputs supply shocks affect equilibrium employment via equation (1.8). In this case, the coefficient on the (log) equilibrium employment variable has a probability limit equal to

$$\Gamma_{OLS} = \xi_1 + \theta_1 + \underbrace{\left[\frac{\text{cov}(\ln(A), \Lambda \ln(\delta')) + \Upsilon \ln(\gamma'))}{\text{var}(\ln(A))}\right]}_{\theta_3},$$

where θ_3 represents the two sources of omitted variables bias. By integrating equation (1.8), one can obtain formulas for $\ln(A)$ and $\text{var}(\ln(A))$ and substitute them into θ_3. Even under the assumptions that the labor supply shocks (β') are uncorrelated with the productivity shocks (δ') and the non-labor input supply shocks (γ'), and that productivity shocks are uncorrelated with non-labor input supply shocks, there will be omitted variables bias because the model is estimated with an equilibrium employment variable (A) instead of a labor supply variable (β'). Under these assumptions $\text{cov}(\ln(\beta'), \ln(\gamma')) = \text{cov}(\ln(\beta'), \ln(\delta'))$ $= \text{cov}(\ln(\gamma'), \ln(\delta')) = 0$, and the formula for the relative bias from the two omitted variables can be expressed as

$$\frac{\theta_3}{\xi_1} = \frac{\Lambda}{\rho_3 \xi_1}\left[\frac{\rho_3^2 \text{var}(\ln(\delta'))}{E}\right] + \frac{\Upsilon}{\rho_2 \xi_1}\left[\frac{\rho_2^2 \text{var}(\ln(\gamma'))}{E}\right] \geq 0, \tag{1.14}$$

where

$$E = \rho_1^2 \text{var}(\ln(\beta')) + \rho_2^2 \text{var}(\ln(\gamma')) + \rho_3^2 \text{var}(\ln(\delta')).$$

Each source of omitted variables bias in this case is non-negative because $\Lambda \geq 0$, $\Upsilon \geq 0$, $\rho_2 \geq 0$ (see equation (1.8)), $\rho_3 \geq 0$, $\rho_1^2 \geq 0$, $\rho_2^2 \geq 0$, $\rho_1^2 \geq 0$, and the variances are non-negative. As one might expect, the magnitude of the relative bias from the unobserved productivity shocks in this case (the first term on the right hand side of (1.14)) depends on the same factors as those described in Case II, although a higher variance of the non-labor input supply shocks ($\text{var}(\ln(\gamma'))$) will also tend to reduce this source of bias. But a larger

var(ln(γ')) will also tend to increase the relative bias from the unobserved non-labor input supply shocks (the second term on the right hand side of (1.14)) counteracting the effect on the first term, so the impact is uncertain. However, both sources of omitted variables bias are unambiguously reduced when the variability of the labor supply shocks (var(ln(β'))) is higher. It can be shown that $\Upsilon/(\rho_2\xi_1) = \left[\xi_2/(\rho_2\xi_1) - 1/\rho_1)\right]$, so the relative bias from the unobserved non-labor input supply shocks will tend to be smaller if the elasticity of production with respect to the non-labor input supply shocks (ξ_2) is small or if the elasticity of production with respect to the labor supply shocks (ξ_1) is large.

The results from a simulation exercise that compares the relative magnitude of θ_1 to θ_3 under a range of structural parameter values can be found in Appendix 1.D. As in Case II, the simulation results indicate that additional information is required to determine whether the omitted variables bias (θ_3) or the employment-labor supply mismatch bias (θ_1) is larger. This result applies to each source of omitted variables bias considered independently and both sources of omitted variables bias considered jointly.

Of course, there are other cases one might want to consider in more detail, which are not covered in this study. For example, one might want to investigate the case where the local supply of labor is perfectly inelastic more thoroughly. As discussed in section 1.3.2.1, under this scenario, a change in the labor supply would cause an equivalent change in equilibrium employment (i.e., $e_A = 0 \implies dA^* = \beta$), so there would be no employment-labor supply mismatch bias (i.e., $\theta_1 = 0$). To the extent that labor supply shocks, and thus employment shocks, are uncorrelated with productivity and non-labor input supply shocks, there would also be no omitted variable bias (i.e., $\theta_2 = \theta_3 = 0$) because $\text{cov}(\ln(A), \ln(\delta')) = \text{cov}(\ln(A), \ln(\gamma')) = 0$. One might then want to relax the assumption that labor supply shocks are uncorrelated with productivity and non-labor input supply shocks and examine the implications. However, a scenario where $e_A = 0$ likely does not reflect the current state of local farm labor markets in California because higher farm wages still draw some immigrant workers to the farm (Charlton et al. 2019a). Although a more thorough investigation of the scenario where $e_A = 0$ may be warranted

in the future if local farm labor supplies become increasingly inelastic, I do not examine that case in more detail here.

1.3.3 Methodology

As discussed above, the main threats to identification when estimating the effect of a change in the farm labor supply on labor-intensive crop production are the use of an equilibrium employment variable in place of a labor supply variable and omitted variables bias. Even when the unobserved non-labor input supply shocks and productivity shocks are uncorrelated with the labor supply shocks, there will be bias because these unobserved variables impact equilibrium employment via equation (1.8). However, if an instrumental variable (Z) is available that generates labor supply driven variation in the employment variable but is uncorrelated with the unobserved non-labor supply and productivity shocks (i.e., if $E[\nu|Z] = 0$), using 2SLS to estimate Γ would remedy the omitted variables bias. Still, the employment-labor supply mismatch bias (θ_1) would remain. The two-stage least squares (2SLS) coefficient on the employment variable has the following probability limit:

$$\Gamma_{2SLS} = \Gamma + \frac{1}{\Phi}\left[\frac{\text{cov}(Z,\nu)}{\text{var}(Z)}\right] = \xi_1 + \theta_1,$$

where Γ is the OLS coefficient from equation (1.11) in the case where there are no omitted variables, and Φ is the coefficient on Z in the first stage regression (see Appendix 1.E for proof). Of course, instruments that are relevant and satisfy the exclusion restriction are rarely found in practice. Therefore, I rely upon the following fixed-effects panel regression model and attempt to mitigate the omitted variables bias through the use of fixed effects and control variables:

$$\ln(O_{ict}) = \Omega \ln(A_{ct}) + \phi_{ic} + \phi_t + \alpha_1^c t_c + \alpha_2^c t_c^2 + \sum_{m=1}^{12} (\tau^m Temp_{ct}^m + \pi^m Precip_{ct}^m) + \mu_{ict}, \quad (1.15)$$

where ln denotes the natural logarithm, i denotes a crop, c denotes a county, and t

denotes a year. The outcomes of interest $O_{ict} \in (Q_{ict}, H_{ict}, Y_{ict})$ are three measures of crop production, where Q_{ict} is the production of each labor-intensive crop in each county in each year, H_{ict} is the number of acres harvested, and Y_{ict} is the average yield (quantity harvested per acre). The main variable of interest $\ln(A_{ct})$ is the (log) number of crop workers employed at the county level during the county's peak employment quarter.

Table 1.1 contains a list of potential sources of bias and the corresponding variables that are used to plausibly mitigate each one. At a minimum, the model should include a set of year fixed effects (ϕ_t) and county fixed effects (ϕ_c). The year fixed effects (ϕ_t) should control for shocks that are common to all counties within a year, such as statewide droughts, which can induce non-labor input supply shocks, such as restricted access to water. The county fixed effects should control for time invariant factors that differ by county, such as senior water rights, geography, and soil quality, which could be correlated with production, harvested acreage, and labor use in hand-harvested crops.

Table 1.1 Sources of Bias and Mitigating Variables

Source of Bias	Example	Mitigating Variables
Productivity Shocks	Spring Freezes, Extreme Heat	$temp^m_{ct}$
	Rain During Pollination	$precip^m_{ct}$
	Technology Diffusion	$t_c + t_c^2$
Non-Labor Input Supply Shocks	Water Restrictions During Drought Years	ϕ_t
	Urban Expansion	$t_c + t_c^2$

Although it is plausible that Ω_{OLS} provides a significant improvement over Γ_{OLS} due to the model's bias mitigation strategy, it may still suffer from upward bias due to the employment-labor supply mismatch if the supply of labor is not perfectly inelastic (i.e., $e_A \neq 0$), the supply of the non-labor input is less than perfectly elastic (i.e., $e_B < \infty$), and the labor and non-labor inputs are imperfect substitutes (i.e., $\sigma < \infty$). As a result, it is plausible that $\xi_1 \leq \Omega_{OLS} < \Gamma_{OLS}$. Therefore, I interpret Ω_{OLS} as an upper bound for the elasticity of labor-intensive crop production with respect to the labor supply (ξ_1).

Because the employment variable varies at the county-year level, as opposed to the

crop-county-year level, if the panel of data is balanced, once the county fixed effects are included in the model the addition of crop fixed effects (ϕ_i) or crop-by-county fixed effects (ϕ_{ic}) may improve the explanatory power of the model but will not mitigate any source of bias. This result emerges because the employment variable does not vary by crop, so the crop fixed effects (or crop-by-county fixed effects) should only explain variation in the outcome variable but not in the employment variable once the county fixed effects are included in the model. However, to the extent that some crops are not produced in every year of the sample, including crop fixed effects or crop-by-county fixed effects will influence the coefficient of interest. The results in the following section reinforce this conclusion and demonstrate that, when the panel of data is balanced, the coefficients are indeed identical. My preferred specification includes the crop-by-county fixed effects because they explain a significant amount of variation in the outcome variable, thus providing a better fit to the data, and their inclusion does not prevent the identification of meaningful upper bounds even when panel of data is unbalanced. For example, in the unbalanced panel for the top 10 labor-intensive crop-producing counties (see bottom panel of Table 1.4), the R^2 from the production model without crop fixed effects or crop-by-county fixed effects (column 4) effects is 0.11, and the model with crop fixed effects in column 5 (respectively crop-by-county fixed effects in column 6) is 0.53 (respectively 0.91).

The second set of control variables are quadratic county trends, defined as $t_c + t_c^2$, where $t_c = \phi_c \times t$, $t_c^2 = \phi_c \times t^2$, and t is a continuous time variable. These trends help control for smooth, yet potentially nonlinear, changes in productivity that could be caused by local technology diffusion, such as smart irrigation technologies, and factors that impact the non-labor input supply, such as urban expansion, which affects the amount of land available for crop production.

The third set of variables includes 12 monthly county-level average temperature variables ($Temp_{ct}^m$) and 12 monthly county-level cumulative precipitation variables ($Precip_{ct}^m$), where m is an index for the month ($m = 1, ..., 12$). These 24 weather variables help control for local weather events, such as spring freezes, extreme heat, or rain during pollination,

which can affect the productivity of labor and non-labor inputs. And μ_{ict} is the error term.

In order to estimate the impact of shifts in the farm labor supply on the total value of production, I estimate the following model:

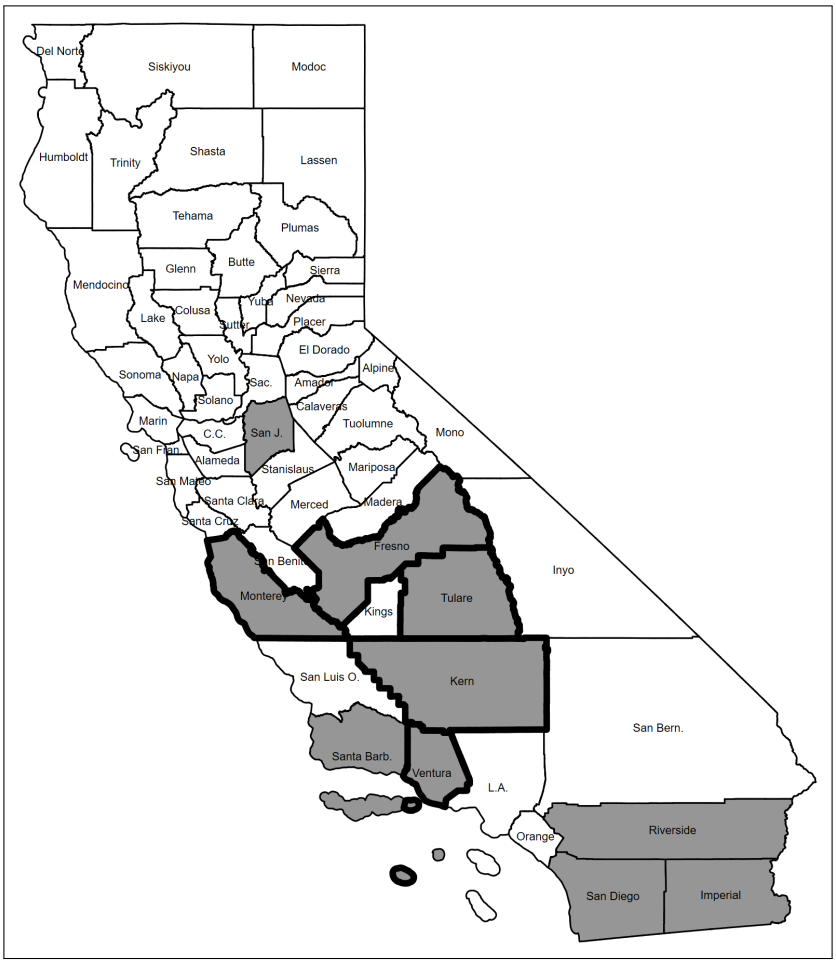
$$\ln(R_{ct}) = \Psi \ln(A_{ct}) + \phi_c + \phi_t + \alpha_1^c t_c + \alpha_2^c t_c^2 + \sum_{m=1}^{12} (\tau^m Temp_{ct}^m + \pi^m Precip_{ct}^m) + \mu_{ct}, \quad (1.16)$$

where c denotes a county, t denotes a year, and $R_{ct} = \sum_{i=1}^{I} \bar{p}_{ic} Q_{ict}$ denotes the average value of hand-harvested crops in a county in a given year, which is calculated by multiplying the production of each crop in each county by its average price (calculated separately for each crop in each county) and summing up over the crops in a county. The average crop price is defined as $\bar{p}_{ic} = 1/T_{ic} \times \sum_{t=1}^{T_{ic}} p_{ict}$, where T_{ic} identifies the number of years a crop was grown in a county. I use average crop prices, as opposed to annual crop prices, because decreases in output may be accompanied by price increases even at the county level. Since the ultimate concern is about the social value of crop production rather than farmer revenue, using annual crop prices would provide the wrong measure. Averaging a crop price over the entire sample period provides a measure of the social value of the crop and allows the estimates to better capture the socially relevant effects (Hagerty 2019). Unlike Hagerty (2019), I allow the price average to vary systematically by county in order to capture the permanent differences in crop value arising from geographic factors (e.g., closeness to processing/packing plants) and possibly harvesting time during the year. The coefficient of interest (Ψ) is interpreted as an upper bound for the effect on the total value of production. The remainder of the variables in equation (1.16) are defined above except for μ_{ct}, which is the error term.

In panel settings that have a natural regional clustering of observations (such as the crop-county level data used in this analysis), it is common to use standard errors that are clustered at the region level (Rogers 1993). In addition to correcting for heteroskedasticity, clustering standard errors at the region level also corrects for serial correlation within geographic regions, which, if present, renders inference based on White (1980)'s

heteroskedastic-robust standard errors invalid due to the violation of the error independence assumption. In order to conduct valid inference with clustered standard errors, the errors must not be correlated across clusters. However, the inter-cluster error independence assumption may be difficult to justify in certain settings, particularly when the clusters are geographic regions that are located within close proximity to each other, such as the 10 top producing counties used in this study (see Figure 1.6).[18]

Fig. 1.6 Geography of the Top 10 Labor-Intensive
Fruit and Vegetable Producing Counties in California



Note: The top 10 counties are shaded in gray. The top 5 counties are outlined with a thick black border. The calculations that determine the top 10 counties exclude all fruit and vegetable crops that have a viable option for mechanical harvest. After excluding those crops, these 10 counties produce about 86% of the value of all hand-harvested fruit and vegetable crops in the state.

[18] I exclude the more marginal counties because there is a greater potential for measurement error in the employment variable due to a larger share of workers who do not work in labor-intensive crop production.

When the inter-cluster error independence assumption is violated, clustered standard error estimates are biased, and inference is no longer valid. Using the Frees test, which is appropriate for use with static panel models when the number of years in the data is less than the number of observations in the cross sectional dimension and year fixed effects are included in the model (De Hoyos and Sarafidis 2006; Frees 1995), I test for cross-sectional dependence in the error term to determine if the use of clustered standard errors is appropriate. The Frees tests provide strong evidence of cross-sectional dependence, which likely results from the close geographic proximity of the 10 labor-intensive crop producing counties considered in the analysis, leading to spatial correlation across counties. Inference based on clustered standard errors is also not valid in my setting because the number of clusters (G) must be large in order for the standard error estimates to be consistent. My analysis is based on the largest 10 producing counties (i.e., $G = 10$), which is too small for clustered standard errors to be reliable. Cameron and Miller (2015) propose the use of the wild cluster bootstrap in cases where the number of clusters is small. However, inference based on the wild cluster bootstrap also requires the absence of error correlation across clusters, which does not hold in my setting. To conduct inference that is valid in the presence of cross-sectional dependence, I use Driscoll-Kraay standard errors (see Driscoll and Kraay 1998; Hoechle 2007), which are robust to general forms of cross-sectional dependence, heteroskedasticity, and error serial correlation up to a specified number of lags. I determine the number of degrees of serial correlation that the Driscoll-Kraay standard errors correct for by using the heteroskedastic-robust Cumby-Huizinga general test for serial correlation (Cumby and Huizinga 1990). For each table, I choose the most conservative degree of serial correlation across the entire set of results and allow the standard errors to correct for that degree of serial correlation in all specifications so that the standard errors are comparable. I also report clustered standard error estimates in the tables that follow for reference.

1.4 Data

The data used for the empirical analysis span the period 1990 to 2017 and cover 10 of the 44 fruit and vegetable crop producing California counties. There are a total of 74 labor-intensive crops used in the analysis, which are listed in Appendix 1.F. A selection of summary statistics for crops grown within a county in all sample years is displayed in Table 1.2.

The crop production data were obtained from the California County Agricultural Commissioners' reports, which are available in .pdf and .csv format on the website of the USDA's National Agricultural Statistics Service (NASS 2018). These data include the value (in U.S. dollars) and quantity (in U.S. tons) of production, the number of acres harvested, and the average yield per acre for each crop in each California county in each year. In a handful of cases, the source .csv data files contain apparent data entry errors, which were detected by conducting a visual examination of statewide production graphs for each commodity and investigating outliers. When possible, these errors were corrected by entering the values from the .pdf text reports. For example, in one instance, the .csv file mistakenly reported the number of 50 pound boxes instead of weight in tons. In a handful of cases, observations are consistent in the .csv and .pdf files but are an order of magnitude different from adjacent observations in the data set. I dropped those outliers out of an abundance of caution. Five data entry errors were updated with values from the .pdf files, and 13 were dropped from the analysis. Data on the value of production (in dollars) are available for all crops in all counties in all years. However, in some cases one or more of the production measures is missing for some crops in some years. In order to maintain a consistent set of observations throughout the study, observations are also dropped if at least one of the three production measures is missing. In order to determine the extent to which missing observations might affect the empirical results, I calculate the share of the total value that the missing production observations account for. The missing production observations account for less than 1% of the total value of the labor-intensive

Table 1.2 Summary Statistics for Hand-Harvested Fruit and Vegetable Crops Grown in All Sample Years in the Top 10 Counties

		Median	Mean	SD
Total Revenue (in $2017 millions)	Overall	769	906	550
	Between			513
	Within			261
Total Production (in millions of tons)	Overall	0.9	1.1	0.7
	Between			0.7
	Within			0.2
Total Harvested Acres (in 1000s)	Overall	70.0	93.8	53.9
	Between			55.5
	Within			12.5
Number of Workers Employed	Overall	19,194	25,513	16,216
	Between			16,033
	Within			5,551
Ave. County Temperature in Fahrenheit (Total County Precipitation in Inches)	January	50.8	50.2	5.1
		(1.9)	(3.1)	(3.3)
	February	52.0	52.1	5.1
		(2.4)	(3.1)	(3.0)
	March	55.4	55.8	5.5
		(1.4)	(2.2)	(2.5)
	April	58.3	59.0	5.9
		(0.6)	(1.0)	(1.2)
	May	63.4	64.9	6.6
		(0.2)	(0.5)	(0.7)
	June	69.9	71.2	7.4
		(0.0)	(0.1)	(0.3)
	July	75.3	76.3	7.9
		(0.0)	(0.1)	(0.2)
	August	74.7	76.1	7.7
		(0.0)	(0.1)	(0.2)
	September	71.2	72.7	6.6
		(0.1)	(0.2)	(0.3)
	October	63.9	64.9	5.5
		(0.3)	(0.7)	(1.0)
	November	55.7	56.0	5.3
		(0.8)	(1.1)	(1.1)
	December	50.3	49.4	4.9
		(1.7)	(2.3)	(2.3)
Obs.		279	279	279

Note: The statistics presented in this table are aggregated at the county-year level. When "Overall," "Between," or "Within" is not specified, the statistics reported are for "Overall."

crops in any given year, so it is unlikely that dropping observations significantly influences the estimates.

The county-level crop employment data were obtained from the public Quarterly Census of Employment and Wages (QCEW) data files (BLS 2018). These data include the average quarterly employment measures for each county based on the North American Industry Classification System (NAICS). To provide a measure of the hand harvest workforce, I include all crop workers directly hired by farmers (NAICS code 111) and those hired by farm labor contractors (NAICS code 115115). In certain counties, employment measures for one of the NAICS codes are suppressed in some years to protect employer anonymity. County-year observations with suppressed QCEW data are dropped from the analysis.[19] The employment measures used in the analysis identify the average employment during each county's peak employment quarter, assumed to be the period of time when the majority of the harvest activities take place. This "peak quarter" is identified by determining the quarter during which each county had its highest average employment over the time period 1990 to 2017. Once the peak quarter is defined for a county, the employment values that correspond to that quarter are assigned to the county for the entire sample period. For example, in Imperial county, where winter lettuce is grown, peak employment typically occurs during the first quarter of the calendar year, but in San Joaquin county, peak employment typically occurs during the second quarter. As a result, the employment measures from quarter 1 (respectively quarter 2) are assigned to Imperial county (respectively San Joaquin county). Assigning each county the employment measure during its peak employment quarter captures variation in local farm employment during the period of time when farmers are particularly susceptible to reduced labor availability because of the spike in harvest labor demand. The peak quarter is generally stable for most counties, although there are a handful of cases where a county's peak quarter fluctuates between adjacent quarters. However, when these values differ, the difference in

[19] Although the results are not shown here, I also estimated a set of results using imputed data for the suppressed observations. The results from both methods are qualitatively similar. I imputed values by estimating a quadratic trend with the non-suppressed observations (separately for each county and NAICS code that has suppressed data) and assigning the predicted values to the missing data.

employment between the assigned quarter and the county's actual peak quarter for that year is usually nominal. For example, when these values differ, the median (respectively mean) difference in employment is 4% (respectively 6%). The empirical estimates are nearly identical if I allow the peak employment quarter to vary from year to year.

Weather data were obtained from the National Oceanic and Atmospheric Administration Climate Data Online website (NOAA 2019). These data include information about temperature and precipitation provided by weather stations located in each county throughout the state. From these data, I generate 12 monthly county-level average temperature variables and 12 monthly county-level cumulative precipitation variables. There are some missing weather data points in Sutter and Colusa counties, which is only relevant for the falsification tests.

1.5 Results

1.5.1 Labor-Intensive Crops

1.5.1.1 Hand-Harvested Fruit and Vegetable Production

The results from equation (1.15) are shown in Tables 1.3 and 1.4. Table 1.3 shows the estimates from a balanced panel of crops that, for a given county, are grown in all sample years. For reference, Table 1.4 shows estimates from an unbalanced panel that includes the set of crops that, for a given county, are grown in at least half of the sample years (14 out of 28). Each table consists of two sets of results: one for the top 5 producing counties, and the other for the top 10 counties. The top 5 (respectively 10) counties produce 65% (respectively 86%) of the value of all labor-intensive fruit and vegetable crops produced in the state. Within each set of results is a subset of results for production, harvested acres, and average yield per acre. Each column in the table displays estimates from models that include a different set of control variables. With the exception of column (3), moving

Table 1.3 Effects of a Change in the Farm Labor Supply on Hand-Harvested Fruit and Vegetable Production for Crops Grown in All Sample Years

	(1)	(2)	(3)	(4)	(5)	(6)
			Top 5 Counties			
	0.460***	0.523***	0.499***	0.506***	0.506***	0.506***
Production	(0.147)	(0.095)	(0.101)	(0.083)	(0.084)	(0.084)
	[0.262]	[0.222]	[0.230]	[0.239]	[0.242]	[0.243]
R^2	0.084	0.084	0.084	0.085	0.739	0.909
	0.198**	0.264***	0.266***	0.290***	0.290***	0.290***
Harvested Acres	(0.079)	(0.072)	(0.053)	(0.074)	(0.075)	(0.075)
	[0.155]	[0.111]	[0.103]	[0.121]	[0.123]	[0.123]
R^2	0.119	0.119	0.119	0.120	0.766	0.933
	0.262***	0.259**	0.233***	0.216**	0.216**	0.216**
Yield	(0.081)	(0.098)	(0.068)	(0.099)	(0.100)	(0.100)
	[0.113]	[0.115]	[0.131]	[0.121]	[0.122]	[0.123]
R^2	0.119	0.121	0.120	0.122	0.721	0.753
N	1,932	1,932	1,932	1,932	1,932	1,932
			Top 10 Counties			
	0.416	0.414***	0.482***	0.463***	0.463***	0.463***
Production	(.)	(0.117)	(0.081)	(0.098)	(0.098)	(0.099)
	[0.174]	[0.160]	[0.150]	[0.143]	[0.144]	[0.145]
R^2	0.137	0.138	0.138	0.138	0.630	0.900
	0.282***	0.277***	0.320***	0.309***	0.309***	0.309***
Harvested Acres	(0.080)	(0.066)	(0.083)	(0.070)	(0.071)	(0.072)
	[0.115]	[0.105]	[0.110]	[0.098]	[0.099]	[0.100]
R^2	0.133	0.133	0.133	0.133	0.657	0.920
	0.133	0.137	0.162**	0.154**	0.154**	0.154**
Yield	(.)	(0.094)	(0.072)	(0.068)	(0.069)	(0.070)
	[0.119]	[0.103]	[0.138]	[0.112]	[0.113]	[0.114]
R^2	0.109	0.110	0.110	0.112	0.730	0.819
N	3,416	3,416	3,416	3,416	3,416	3,416
Year FE	X	X	X	X	X	X
Quadratic County Trends	X	X	X	X	X	X
Monthly Temp. Controls	–	X	–	X	X	X
Monthly Precip. Controls	–	–	X	X	X	X
County FE	X	X	X	X	X	–
Crop FE	–	–	–	–	X	–
Crop-by-County FE	–	–	–	–	–	X

Note: The results in this table are generated from a sample that only includes crops that were grown in a county for each of the 28 years in the sample. Significance levels are based on Driscoll-Kraay standard errors, which are reported in parentheses. In cases where Stata could not calculate Driscoll-Kraay standard errors because of a singular covariance matrix, the standard error is replaced with "(.)". Standard errors clustered at the county-level are reported in brackets for reference. * $p < .1$, ** $p < .05$, *** $p < .01$.

Table 1.4 Effects of a Change in the Farm Labor Supply on Hand-Harvested Fruit and Vegetable Production for Crops Grown in at Least 14 of the 28 Sample Years

	(1)	(2)	(3)	(4)	(5)	(6)
			Top 5 Counties			
	0.601***	0.649***	0.604***	0.622***	0.494***	0.467***
Production	(0.117)	(0.075)	(0.079)	(0.076)	(0.083)	(0.075)
	[0.315]	[0.266]	[0.250]	[0.256]	[0.218]	[0.206]
R^2	0.030	0.031	0.031	0.032	0.618	0.902
	0.338***	0.374***	0.346***	0.373***	0.329***	0.297***
Harvested Acres	(0.089)	(0.078)	(0.074)	(0.092)	(0.104)	(0.065)
	[0.183]	[0.157]	[0.115]	[0.155]	[0.123]	[0.115]
R^2	0.043	0.043	0.043	0.043	0.610	0.916
	0.263***	0.275***	0.258***	0.250***	0.165**	0.170**
Yield	(0.055)	(0.086)	(0.051)	(0.084)	(0.078)	(0.082)
	[0.145]	[0.111]	[0.142]	[0.103]	[0.101]	[0.096]
R^2	0.103	0.105	0.104	0.107	0.729	0.772
N	3,250	3,250	3,250	3,250	3,250	3,250
			Top 10 Counties			
	0.535***	0.545***	0.597***	0.593***	0.514***	0.366***
Production	(0.069)	(0.057)	(0.061)	(0.072)	(0.059)	(0.046)
	[0.200]	[0.175]	[0.175]	[0.158]	[0.143]	[0.125]
R^2	0.110	0.110	0.110	0.110	0.533	0.908
	0.374***	0.376***	0.402***	0.410***	0.369***	0.229***
Harvested Acres	(0.028)	(0.046)	(0.036)	(0.057)	(0.046)	(0.028)
	[0.117]	[0.112]	[0.098]	[0.097]	[0.123]	[0.092]
R^2	0.121	0.121	0.121	0.121	0.544	0.919
	0.161**	0.169***	0.195***	0.183***	0.145***	0.137***
Yield	(0.062)	(0.041)	(0.058)	(0.038)	(0.038)	(0.034)
	[0.102]	[0.075]	[0.106]	[0.074]	[0.054]	[0.062]
R^2	0.079	0.081	0.080	0.082	0.714	0.822
N	5,823	5,823	5,823	5,823	5,823	5,823
Year FE	X	X	X	X	X	X
Quadratic County Trends	X	X	X	X	X	X
Monthly Temp. Controls	–	X	–	X	X	X
Monthly Precip. Controls	–	–	X	X	X	X
County FE	X	X	X	X	X	–
Crop FE	–	–	–	–	X	–
Crop-by-County FE	–	–	–	–	–	X

Note: The results in this table are generated from a sample that only includes crops that were grown in a county in at least 14 of the 28 years in the sample. Significance levels are based on Driscoll-Kraay standard errors, which are reported in parentheses. Standard errors clustered at the county-level are reported in brackets for reference. * $p < .1$, ** $p < .05$, *** $p < .01$.

from left to right in the table, each column contains the results from a model specification that is progressively more demanding. The results from my preferred specification are presented in column (6), which includes the crop-by-county fixed effects.

When focusing on the production results for the top 5 counties in Table 1.3 (respectively Table 1.4), the upper bound indicates that a one percent decrease in the farm labor supply causes at most a 0.51% (respectively 0.47%) decrease in hand-harvested specialty crop production. Reduced production is primarily channeled through a decrease in the number of acres harvested, although there is a smaller effect on the average yield. One potential explanation for the yield effect is that farmers may be constrained by the number of times they can get harvest crews to harvest crops, such as strawberries, that do not ripen uniformly and require multiple rounds of harvest. In the top 5 counties, the results from Table 1.3 (respectively Table 1.4) indicate that a one percent decrease in the farm labor supply causes at most a 0.29% (respectively 0.30%) reduction in harvested acres and at most a 0.22% (respectively 0.17%) decrease in the average yield.

As the effects are estimated across a larger set of counties, the coefficients become smaller, although they remain significant. In the top 10 counties, the results from Table 1.3 (respectively Table 1.4) indicate that a one percent reduction in the farm labor supply causes at most a 0.46% (respectively 0.37%) reduction in production, at most a 0.31% (respectively 0.23%) reduction in harvested acres, and at most a 0.15% (respectively 0.14%) reduction in the average yield.

1.5.1.2 Value of Hand-Harvested Fruit and Vegetable Production

The results for the total value of hand-harvested crops, which are presented in Table 1.5, are generally consistent with the production results in section 1.5.1, although they suggest that the effects are concentrated in the top 5 counties. Moving from left to right in the table, each column provides a set of results from a model that is progressively more demanding. The results in column (4) indicate that a 1% decrease in the farm labor supply in the top 5 counties could cause as much as a 0.46% decrease in the total value of labor-intensive

crop production. The results for the top 10 counties are not statistically significant in the most demanding specification, but the coefficients in columns (3) and (4) are smaller than those for the top 5 counties, which is consistent with the production results.

Table 1.5 Effects of a Change in the Farm Labor Supply
on the Total Value of Hand-Harvested Fruit and Vegetable Crops

	(1)	(2)	(3)	(4)
		Top 5 Counties		
Revenue	0.077	0.321***	0.555***	0.458***
	(0.050)	(0.050)	(0.146)	(0.115)
	[0.250]	[0.148]	[0.246]	[0.218]
N	140	140	140	140
		Top 10 Counties		
Revenue	0.846***	0.538***	0.231*	0.137
	(0.039)	(0.099)	(0.124)	(0.136)
	[0.161]	[0.205]	[0.197]	[0.182]
N	279	279	279	279
Year F.E.	X	X	X	X
County F.E.	–	X	X	X
Quadratic County Trends	–	–	X	X
Temp. and Precip. Vars	–	–	–	X

Note: Significance levels are based on Driscoll-Kraay standard errors, which are reported in parentheses. Standard errors clustered at the county-level are reported in brackets for reference.
* $p < .1$, ** $p < .05$, *** $p < .01$.

1.5.2 Falsification Tests: Mechanically-Harvested Crops

1.5.2.1 Nut and Field Crop Production

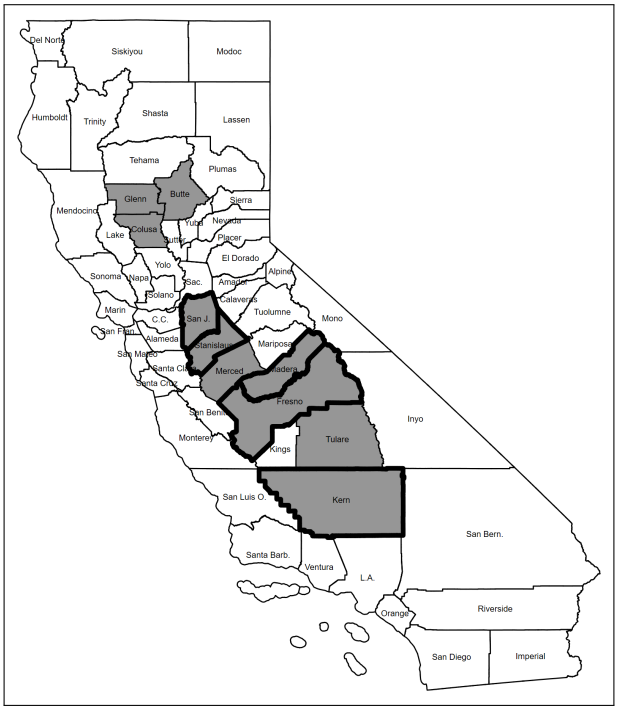
Nuts and field crops are harvested by machines and require much less labor than hand-harvested crops to produce. As a result, one should not expect to find effects of farm labor supply shocks on the production of mechanically-harvested crops as large as those found on labor-intensive crops. In order to test this hypothesis, I estimate the elasticity of production with respect to the farm labor supply for nuts and field crops, separately. I focus on the top 5 and 10 nut and field crop producing counties, which have some geographic overlap with the top labor-intensive crop producing counties but are not

produced in the coastal regions (see Figure 1.7).

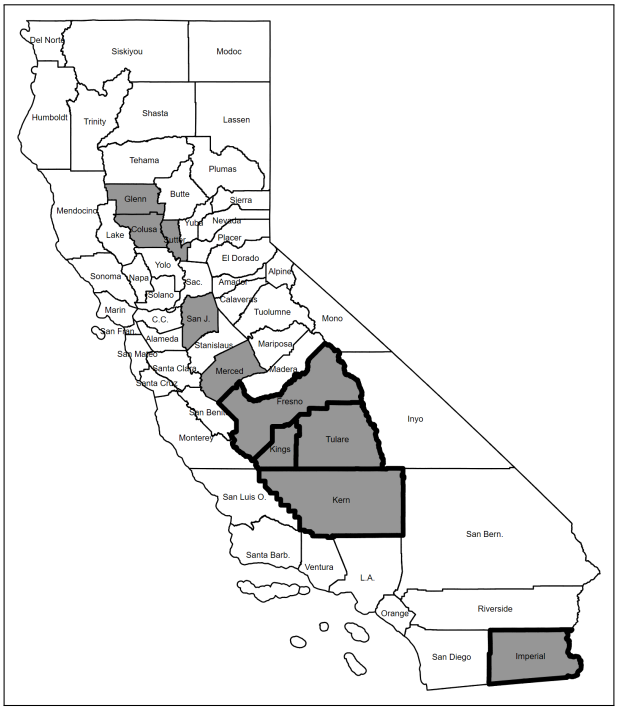
The estimated impacts for nuts (respectively field crops) are shown in Tables 1.6 and 1.7 (respectively Tables 1.8 and 1.9). As with the hand-harvested crops, the analysis for nuts and field crops is conducted separately using a balanced panel of data, as well as an unbalanced panel that includes crops grown within a county in at least half of the sample years. A review of these tables reveals two important facts. First, the magnitudes of the coefficients, when positive, are much smaller than those found in the fruit and vegetable crop analysis, and they are never positive and statistically significant. Second, in some cases, the production coefficients are negative, although when they are, they are only statistically significant for field crops in the balanced panel of data. A negative elasticity would imply that a decline in the farm labor supply is causing an increase in the production of mechanically-harvested crops. Although a negative elasticity would confirm anecdotal evidence, which suggests that some farmers are switching production out of hand-harvested crops into those that can be mechanically-harvested, the results presented here do not provide a uniform body of evidence to support it (e.g., Ryssdal 2017; Martin 2019; CFBF and UC Davis 2019; Rutledge and Taylor 2019a). Moreover, the estimated effects on the total value of production in section 1.5.2.2 fail to confirm the anecdotal evidence. Nevertheless, the evidence from this analysis is consistent with the hypothesis that negative farm labor supply shocks do not cause a significant decline in the production of mechanically-harvested crops but, instead, only cause significant negative effects in the sub-sector of agriculture that is particularly reliant upon labor inputs during harvest time.

Fig. 1.7 Geography of Top 5 and 10 Nut and Field Crop Producing Counties

A: Top Nut Producing Counties



B: Top Field Crop Producing Counties



Note: The top 10 counties are shaded in gray. The top 5 counties are outlined with a thick black border.

Table 1.6 Effects of a Change in the Farm Labor Supply on Nut Crop Production for Crops Grown in all Sample Years

	(1)	(2)	(3)	(4)	(5)	(6)
			Top 5 Counties			
	0.168	0.012	0.100	-0.011	-0.011	-0.011
Production	(0.213)	(0.110)	(0.232)	(0.146)	(0.146)	(0.148)
	[0.080]	[0.022]	[0.120]	[0.036]	[0.036]	[0.036]
R^2	0.278	0.279	0.278	0.279	0.627	0.946
	0.164	0.098	0.126	0.131	0.131	0.131
Harvested Acres	(.)	(0.093)	(0.118)	(0.116)	(0.116)	(0.118)
	[0.149]	[0.119]	[0.189]	[0.147]	[0.147]	[0.149]
R^2	0.209	0.209	0.209	0.209	0.723	0.973
	0.005	-0.086	-0.026	-0.142	-0.142	-0.142
Yield	(0.254)	(0.097)	(0.271)	(0.135)	(0.136)	(0.137)
	[0.119]	[0.113]	[0.183]	[0.117]	[0.118]	[0.119]
R^2	0.162	0.178	0.170	0.186	0.622	0.645
N	364	364	364	364	364	364
			Top 10 Counties			
	0.107	0.090	0.105	0.121	0.121	0.121
Production	(.)	(0.095)	(0.098)	(0.098)	(0.098)	(0.099)
	[0.076]	[0.066]	[0.081]	[0.061]	[0.062]	[0.062]
R^2	0.275	0.275	0.275	0.275	0.616	0.948
	0.057	0.035	0.061	0.043	0.043	0.043
Harvested Acres	(.)	(0.043)	(0.042)	(0.046)	(0.047)	(0.047)
	[0.043]	[0.058]	[0.042]	[0.058]	[0.058]	[0.059]
R^2	0.206	0.206	0.206	0.206	0.675	0.971
	0.049	0.055	0.044	0.078	0.078	0.078
Yield	(.)	(0.093)	(0.098)	(0.094)	(0.095)	(0.096)
	[0.058]	[0.065]	[0.055]	[0.055]	[0.055]	[0.055]
R^2	0.222	0.227	0.227	0.231	0.659	0.686
N	616	616	616	616	616	616
Year FE	X	X	X	X	X	X
Quadratic County Trends	X	X	X	X	X	X
Monthly Temp. Controls	–	X	–	X	X	X
Monthly Precip. Controls	–	–	X	X	X	X
County FE	X	X	X	X	X	–
Crop FE	–	–	–	–	X	–
Crop-by-County FE	–	–	–	–	–	X

Note: Significance levels are based on Driscoll-Kraay standard errors, which are reported in parentheses. In cases where Stata could not calculate Driscoll-Kraay standard errors because of a singular covariance matrix, the standard error is replaced with "(.)". Standard errors clustered at the county-level are reported in brackets for reference. * $p < .1$, ** $p < .05$, *** $p < .01$.

Table 1.7 Effects of a Change in the Farm Labor Supply on Nut Crop Production for Crops Grown in at Least 14 of the 28 Sample Years

	(1)	(2)	(3)	(4)	(5)	(6)
	Top 5 Counties					
	0.168	0.012	0.100	-0.011	-0.011	-0.011
Production	(0.213)	(0.110)	(0.232)	(0.146)	(0.146)	(0.148)
	[0.080]	[0.022]	[0.120]	[0.036]	[0.036]	[0.036]
R^2	0.278	0.279	0.278	0.279	0.627	0.946
	0.164	0.098	0.126	0.131	0.131	0.131
Harvested Acres	(.)	(0.093)	(0.118)	(0.116)	(0.116)	(0.118)
	[0.149]	[0.119]	[0.189]	[0.147]	[0.147]	[0.149]
R^2	0.209	0.209	0.209	0.209	0.723	0.973
	0.005	-0.086	-0.026	-0.142	-0.142	-0.142
Yield	(0.254)	(0.097)	(0.271)	(0.135)	(0.136)	(0.137)
	[0.119]	[0.113]	[0.183]	[0.117]	[0.118]	[0.119]
R^2	0.162	0.178	0.170	0.186	0.622	0.645
N	364	364	364	364	364	364
	Top 10 Counties					
	0.096	0.033	0.107	0.085	0.065	0.059
Production	(0.110)	(0.085)	(0.078)	(0.079)	(0.060)	(0.054)
	[0.133]	[0.105]	[0.142]	[0.127]	[0.101]	[0.060]
R^2	0.262	0.263	0.264	0.264	0.556	0.957
	0.070	0.024	0.085	0.053	0.035	0.022
Harvested Acres	(0.109)	(0.099)	(0.082)	(0.087)	(0.065)	(0.037)
	[0.126]	[0.101]	[0.113]	[0.094]	[0.072]	[0.049]
R^2	0.199	0.200	0.200	0.200	0.603	0.977
	0.026	0.009	0.021	0.032	0.030	0.037
Yield	(0.073)	(0.049)	(0.062)	(0.048)	(0.049)	(0.050)
	[0.053]	[0.056]	[0.060]	[0.056]	[0.054]	[0.054]
R^2	0.223	0.229	0.232	0.237	0.664	0.696
N	756	756	756	756	756	756
Year FE	X	X	X	X	X	X
Quadratic County Trends	X	X	X	X	X	X
Monthly Temp. Controls	–	X	–	X	X	X
Monthly Precip. Controls	–	–	X	X	X	X
County FE	X	X	X	X	X	–
Crop FE	–	–	–	–	X	–
Crop-by-County FE	–	–	–	–	–	X

Note: Significance levels are based on Driscoll-Kraay standard errors, which are reported in parentheses. In cases where Stata could not calculate Driscoll-Kraay standard errors because of a singular covariance matrix, the standard error is replaced with "(.)". Standard errors clustered at the county-level are reported in brackets for reference. * $p < .1$, ** $p < .05$, *** $p < .01$.

Table 1.8 Effects of a Change in the Farm Labor Supply on Field Crop Production for Crops Grown in All Sample Years

	(1)	(2)	(3)	(4)	(5)	(6)
			Top 5 Counties			
	0.566	0.448	0.185	0.163	0.163	0.163
Production	(0.355)	(0.356)	(0.315)	(0.285)	(0.288)	(0.290)
	[0.231]	[0.137]	[0.142]	[0.228]	[0.231]	[0.232]
R^2	0.148	0.150	0.150	0.151	0.930	0.945
	0.325	0.140	0.039	0.072	0.072	0.072
Harvested Acres	(0.334)	(0.314)	(0.317)	(0.275)	(0.278)	(0.280)
	[0.083]	[0.119]	[0.057]	[0.122]	[0.123]	[0.124]
R^2	0.106	0.109	0.110	0.112	0.837	0.875
	0.241**	0.309**	0.146**	0.091	0.091	0.091
Yield	(0.105)	(0.144)	(0.068)	(0.084)	(0.085)	(0.086)
	[0.151]	[0.161]	[0.118]	[0.120]	[0.121]	[0.122]
R^2	0.173	0.174	0.174	0.175	0.978	0.978
N	392	392	392	392	392	392
			Top 10 Counties			
	-0.339	-0.260	-0.447**	-0.427***	-0.427***	-0.427**
Production	(0.200)	(0.177)	(0.184)	(0.152)	(0.153)	(0.155)
	[0.292]	[0.233]	[0.298]	[0.233]	[0.235]	[0.237]
R^2	0.206	0.206	0.206	0.207	0.876	0.953
	-0.240	-0.158	-0.318*	-0.305**	-0.305**	-0.305**
Harvested Acres	(.)	(0.154)	(0.165)	(0.141)	(0.141)	(0.143)
	[0.266]	[0.174]	[0.259]	[0.183]	[0.184]	[0.186]
R^2	0.211	0.212	0.212	0.213	0.744	0.902
	-0.099*	-0.102*	-0.129**	-0.123*	-0.123*	-0.123*
Yield	(0.050)	(0.055)	(0.059)	(0.064)	(0.064)	(0.065)
	[0.063]	[0.087]	[0.059]	[0.062]	[0.063]	[0.063]
R^2	0.134	0.134	0.134	0.134	0.969	0.973
N	1,092	1,092	1,092	1,092	1,092	1,092
Year FE	X	X	X	X	X	X
Quadratic County Trends	X	X	X	X	X	X
Monthly Temp. Controls	–	X	–	X	X	X
Monthly Precip. Controls	–	–	X	X	X	X
County FE	X	X	X	X	X	–
Crop FE	–	–	–	–	X	–
Crop-by-County FE	–	–	–	–	–	X

Note: Significance levels are based on Driscoll-Kraay standard errors, which are reported in parentheses. In cases where Stata could not calculate Driscoll-Kraay standard errors because of a singular covariance matrix, the standard error is replaced with "(.)". Standard errors clustered at the county-level are reported in brackets for reference. * $p < .1$, ** $p < .05$, *** $p < .01$.

Table 1.9 Effects of a Change in the Farm Labor Supply on Field Crop Production for Crops Grown in at Least 14 of the 28 Sample Years

	(1)	(2)	(3)	(4)	(5)	(6)
			Top 5 Counties			
	0.041	-0.076	0.106	-0.101	0.084	-0.000
Production	(0.262)	(0.317)	(0.237)	(0.293)	(0.166)	(0.158)
	[0.251]	[0.262]	[0.179]	[0.164]	[0.107]	[0.122]
R^2	0.027	0.027	0.028	0.028	0.837	0.904
	0.030	-0.125	0.103	-0.131	0.064	0.026
Harvested Acres	(0.218)	(0.250)	(0.178)	(0.215)	(0.134)	(0.127)
	[0.222]	[0.250]	[0.158]	[0.187]	[0.109]	[0.136]
R^2	0.058	0.059	0.059	0.061	0.660	0.799
	0.011	0.050	0.003	0.031	0.020	-0.026
Yield	(0.099)	(0.132)	(0.094)	(0.126)	(0.061)	(0.063)
	[0.098]	[0.106]	[0.090]	[0.118]	[0.057]	[0.056]
R^2	0.031	0.032	0.032	0.032	0.966	0.974
N	1,337	1,337	1,337	1,337	1,337	1,337
			Top 10 Counties			
	-0.192	-0.180	-0.189	-0.203	-0.122	-0.126
Production	(0.149)	(0.171)	(0.133)	(0.146)	(0.115)	(0.110)
	[0.123]	[0.154]	[0.135]	[0.167]	[0.184]	[0.162]
R^2	0.099	0.099	0.099	0.099	0.794	0.924
	-0.005	-0.039	0.010	-0.057	-0.047	-0.035
Harvested Acres	(0.102)	(0.137)	(0.088)	(0.114)	(0.100)	(0.099)
	[0.057]	[0.086]	[0.062]	[0.106]	[0.133]	[0.114]
R^2	0.156	0.156	0.157	0.157	0.606	0.853
	-0.187**	-0.141*	-0.199**	-0.147**	-0.074**	-0.091**
Yield	(0.087)	(0.075)	(0.076)	(0.065)	(0.033)	(0.033)
	[0.078]	[0.081]	[0.076]	[0.074]	[0.053]	[0.051]
R^2	0.042	0.043	0.043	0.043	0.960	0.973
N	2,537	2,537	2,537	2,537	2,537	2,537
Year FE	X	X	X	X	X	X
Quadratic County Trends	X	X	X	X	X	X
Monthly Temp. Controls	–	X	–	X	X	X
Monthly Precip. Controls	–	–	X	X	X	X
County FE	X	X	X	X	X	–
Crop FE	–	–	–	–	X	–
Crop-by-County FE	–	–	–	–	–	X

Note: Significance levels are based on Driscoll-Kraay standard errors, which are reported in parentheses. Standard errors clustered at the county-level are reported in brackets for reference.
* $p < .1$, ** $p < .05$, *** $p < .01$.

1.5.2.2 Value of Nut and Field Crop Production

The results for the total value of nut and field crop production, which can be found in Table 1.10, are generally consistent with the production results in section 1.5.2. The more demanding specifications show very small coefficients, none of which are statistically significant. Even the coefficients in column (2), which are from a model specification that arguably contains the least amount of control variables that should be considered for this type of analysis, are never large and positive. These results indicate that decreases in the farm labor supply likely do not have a significant impact on the total value of

Table 1.10 Effects of a Change in the Farm Labor Supply on the Total Value of Nut and Field Crops

	(1)	(2)	(3)	(4)
		Top 5 Counties		
Nut Revenue	0.188***	-0.156*	-0.054	-0.064
	(0.040)	(0.088)	(0.193)	(0.167)
	[0.162]	[0.351]	[0.156]	[0.142]
N	140	140	140	140
Field Crop Revenue	0.061**	0.036	0.053	-0.028
	(0.029)	(0.073)	(0.120)	(0.126)
	[0.041]	[0.316]	[0.158]	[0.135]
N	137	137	137	137
		Top 10 Counties		
Nut Revenue	0.378***	0.075	0.070	0.054
	(0.010)	(0.058)	(0.087)	(0.078)
	[0.077]	[0.257]	[0.108]	[0.071]
N	277	277	277	271
Field Crop Revenue	0.231***	0.117**	-0.059	-0.073
	(0.019)	(0.057)	(0.079)	(0.091)
	[0.038]	[0.248]	[0.100]	[0.123]
N	276	276	276	242
Year F.E.	X	X	X	X
County F.E.	–	X	X	X
Quadratic County Trends	–	–	X	X
Temp. and Precip. Vars	–	–	–	X

Note: Significance levels are based on Driscoll-Kraay standard errors, which are reported in parentheses. Standard errors clustered at the county-level are reported in brackets for reference.
* $p < .1$, ** $p < .05$, *** $p < .01$.

mechanically-harvested crop production.

1.6 Conclusion

Recent studies have found a decline in the U.S. farm labor supply, which has been driven by demographic and structural changes in Mexico, increased U.S. border security measures, and a decline in the number of farm workers willing to engage in follow-the-crop migration, which has reduced the geographic reach of local farm labor markets (Zahniser et al. 2011; Passel et al. 2012; Hertz and Zahniser 2012; Taylor et al. 2012; Fan et al. 2015; Charlton and Taylor 2016; Zahniser et al. 2018, 2019). A smaller farm labor supply has the potential to reduce the nation's access to safe and healthy produce, put upward pressure on food prices, or create significant economic losses for farmers. In order to examine the extent to which changes in the farm labor supply affect crop production, I estimate fixed-effects panel regressions using a rich set of production data from California, the leading agricultural producer and employer in the U.S. I use an equilibrium displacement model to gain insight into the bias of my empirical estimates under different market scenarios. This exercise indicates that my point estimates should be interpreted as upper bounds. The empirical results reveal statistically significant upper bounds for the effects on the production of hand-harvested fruit and vegetable crops but not on mechanically-harvested nuts or field crops.

Although the bounds for fruit and vegetable production are economically meaningful, they indicate that the impacts of the declining farm labor supply will likely be limited into the foreseeable future. The effects are perhaps best exemplified by focusing on the top 5 producing counties, which produce 65% of the value of all labor-intensive crops in the state. According to Charlton and Taylor (2016), the U.S. farm labor supply is shrinking by about one percent each year. A decline in the farm labor supply of that magnitude in the top 5 counties could cause a loss of an additional 60,000 tons (or 0.47%) of fruits and vegetables each year (i.e., 60,000 tons in the first year, 60,000 + 60,000 = 120,000 tons in

the second year, etc.). Over a decade, production value losses of 0.46% per year in those counties could add up to as much as $3.4 billion (or roughly 2.5% of the total value of these crops).[20] These results suggest that the negative trend in the farm labor supply could generate economically meaningful losses for some farmers, but it will likely not devastate the aggregate production of fruits and vegetables over the next decade.

The results from this study also reveal that there is not a one-to-one relationship between labor-intensive crop production and farm labor, which is an indication that, to some extent, labor inputs can be substituted for other inputs. There are at least two factors that likely contribute to this result. First, the farm labor supply in this study only considers the number of workers and does not account for adjustments on the intensive margin, such as changes in the number of weeks worked per year or hours of work per week. Data from the National Agricultural Workers Survey indicate that, on average, farm workers are supplying more units of labor each year. Over the past few decades, farmers have become increasingly reliant upon farm labor contractors to reduce frictions in the farm labor market (Thilmany and Blank 1996; Thilmany 1996; BLS 2018). The use of farm labor contractors helps reduce the burden associated with finding harvest workers, which can increase the number of employee-employer matches throughout the year. An increase in the number of these matches can translate to intensive margin adjustments as workers find more employment. As a result, a 1% decrease in the supply of workers may correspond to a smaller decrease in the supply of labor units, and estimates based on the units of labor could potentially be larger than those uncovered by this analysis.

Second, farmers are increasingly making use of labor-saving technologies (CFBF and UC Davis 2019; Rutledge and Taylor 2019a). Although mechanical harvesters are currently not available for the vast majority of fruit and vegetable crops, other technologies, such as hydraulic platforms in tree orchards and conveyor belts in lettuce fields, are

[20]These amounts are present discounted values that assume 2% inflation and a discount rate of 1%. These figures are also calculated under the assumption that the production of hand-harvested fruits and vegetables is not replaced with the production of other crops. These values also do not include crops that are both hand harvested and mechanically harvested (such as wine and raisin grapes). Wine grapes, for example, generate about $3.5 billion in farm gate revenue each year, and a significant proportion (perhaps 50%) is harvested by hand.

readily available and can help stretch the remaining workforce by increasing efficiency and reducing the physical difficulty associated with harvesting, enabling workers to work longer shifts and maintain employment at an older age.

The declining farm labor supply is a valid concern for California farmers who are unable to hire all the workers they need during harvest time. However, the fact that the region is so well suited to produce high-value labor-intensive crops, coupled with the fact that farmers are actively making adjustments to mitigate issues stemming from reduced labor availability, means that farmers will likely continue producing these high-value crops into the foreseeable future even if fewer workers are available.

Chapter 2: The Short-Run Impacts of Immigration on Native Workers: A Sectoral Approach[1][2]

2.1 Introduction

Labor economists have long been interested in the impacts of immigration on the labor market outcomes of native-born workers in the United States (US), starting with the seminal work of Grossman (1982), followed by influential contributions by Card (1990), Altonji and Card (1991), Friedberg and Hunt (1995), Borjas et al. (1997), Card (2001), Borjas (2003), Card (2009), Peri and Sparber (2009), Ottaviano and Peri (2012), and Dustmann et al. (2017), to name a few.

While the wealth of estimates produced by this literature has failed to paint a consensual picture of immigration effects (Basso and Peri 2015b), recent work by Dustmann et al. (2016) helps rationalize some of the empirical discrepancies found across wage studies, elucidating how different sources of variation in fact identify different structural parameters. While the "national skill-cell approach" of Borjas (2003) and the "mixture approach" of Card (2001) identify *relative* wage effects (across education-experience groups and ed-

[1]The research presented in this chapter was conducted in conjunction with co-author Dr. Pierre Mérel, Professor of Agricultural and Resource Economics at the University of California, Davis.

[2]This chapter will appear (perhaps with slight modifications) in the book *World Enclyclopaedia of Global Migration* edited by Dr. Robert Sauer.

ucation groups, respectively), the "pure spatial approach" potentially identifies a total effect, and is therefore the method of choice according to Dustmann et al. (2016).

Building upon the spatial approach, this paper proposes new estimates of the short-run impacts of immigration on the employment conditions of US-born workers based on a fixed-effects panel regression of US metropolitan areas spanning the years 1990-2011, a period during which the US experienced a remarkable increase in immigration. We use a novel partial identification strategy that has not been exploited in the related literature to date. Our approach requires estimating immigration impacts at a sectoral, rather than economy-wide, level. While this restriction may be seen by some as a weakness, it allows us to paint a contrasted picture of immigration impacts across nine sectors of the US economy with high immigrant worker penetration: construction, transportation, manufacturing, maintenance, food service, personal services, computers, engineering, and science. Not counting agriculture, these nine sectors are the ones with highest immigrant shares over the period 1990-2011.[3] Taken together, they have employed 34.9% of the total native workforce and 50.3% of the low-skilled native workforce.[4] While the first six sectors employ predominantly low-skilled workers, the last three employ predominantly high-skilled workers. Table 2.1 shows the percentages of workers having no more than a high-school diploma and those having at least a Bachelor's degree in each of these nine sectors.

The sectoral approach delivers upper bounds for the short-run impacts of immigration on native workers' earnings, occupational levels, and sectoral employment rate. In the personal service, food service, and construction sectors, upper bounds on earnings are typically negative, statistically significant, and of much larger magnitude than recently published estimates for the US economy as a whole, suggesting that there exist transitory costs to immigration for part of the native population. We find that a 10 percentage point

[3]We do not look at agriculture because our data set focuses on workers located in large metropolitan areas, and because the share of immigrants in that sector calculated using our data set grossly understates the actual prevalence of immigrants, as inferred from other sources like the National Agricultural Workers Survey, which uses a nationally representative sample of agricultural workers.

[4]Here we define "low-skilled" as having no more than a high-school diploma or equivalent. Our empirical analysis includes native workers with any level of educational achievement.

Table 2.1 Educational Attainment of the Native-born Workforce by Sector

Sector	High-School or Less (%)	Bachelor's Degree or More (%)
Food Service	65.4	7.1
Maintenance	72.3	5.1
Personal Service	50.8	13.4
Construction	67.2	5.6
Manufacturing	67.1	5.7
Transportation	68.2	6.2
Computers	8.7	59.4
Engineering	6.7	69.8
Science	9.9	73.2

Note: The native-born workforce is defined by individuals who were between the ages of 18-64, were not in school, were not living in group quarters, and who were jointly in the labor force at the time of the survey and worked a positive number of weeks during the previous year.

increase in the share of immigrant workers in personal services (respectively food service, respectively construction), which is less than the increases that occurred over the period of investigation, causes at least a 6.6%, (respectively 6.0%, respectively 2.9%) drop in the annual earnings of natives. These effects are generally more pronounced for occupations within these sectors most exposed to immigrant inflows. For example, when focusing on workers within personal services (respectively construction) in occupations with the highest immigrant shares, we find effects about twice as large as those for the sector as a whole.

Earnings results in the remaining sectors are more nuanced. Although our point estimates suggest negative effects on the annual earnings of natives in maintenance and transportation, these effects are only statistically significant once we focus on more disaggregated portions of these sectors. In the maintenance sector, we find significantly negative effects for natives in occupations related to landscaping; interestingly, the effect on these workers is of comparable magnitude as that found in the immigrant-exposed occupations of the construction sector such as roofers or painters. In the transportation sector, which includes such varied occupations as aircraft pilots, boat operators, or garbage collectors, we find significantly negative effects in immigrant-exposed occupations like drivers or

loaders. We do not find significant earnings effects in the manufacturing sector, likely due to the traded nature of the goods produced. Nor do we find effects in the three higher-skilled sectors considered, perhaps due to complementarities between native and immigrant labor in these sectors (Ottaviano and Peri 2012; Manacorda et al. 2012).

An important insight of our analysis is that annual earnings effects, where present, may be partly driven by reductions in the occupational levels of natives, i.e., fewer weeks worked per year. This is particularly true for construction occupations, as well as immigration-exposed personal service occupations such as child and personal care. In these occupations, income is often earned "per job" and workers compete for jobs, sometimes through a formal bidding process. Such occupations also have high rates of self-employment, and work may be undeclared. To the extent that immigrant workers, some of whom work illegally, have a preference for work unreported to the government or are willing to accept lower pay, they cost less to employers and may be in a position to outcompete natives.[5]

In line with earlier literature, our results on earnings are derived conditioning on workers earning a strictly positive income (among other criteria). Our analysis shows that immigration has also had sizable effects on native workers' employment rate in the six low-skilled sectors considered, including manufacturing. Again, the largest effects are found in the construction, food service, and personal services sectors where a 10 percentage point increase in the share of immigrants causes at least a 2.3 (respectively 1.8, respectively 1.7) percentage point decrease in the employment rate of natives when considering the entire sector and a 3.6 (respectively 1.8, respectively 3.0) percentage point decrease when focusing on immigrant-exposed occupations. In the three higher-skilled sectors, there is no discernible effect of immigration on sectoral employment, suggesting that natives are either not displaced, or are displaced but find employment in another

[5]The construction sector, in particular, is notorious for having high rates of "under the table" employment (i.e., workers paid in cash without reporting employment to the government). Contractors engaging in this type of employment have a competitive advantage over others who strictly employ workers "on the books" because of the additional cost of workman's compensation, unemployment insurance, and other payroll taxes, which must be factored into bids (Fishman 2013).

sector.

Although the definition of economic sectors we use is quite broad and accommodates within-sector mobility, immigration in one sector could plausibly cause some natives to shift to other sectors, raising concerns that our sectoral estimates could be partly driven by compositional effects. To address this concern, we pool lower-skilled sectors into a composite sector and estimate effects at the level of the composite sector. The results confirm negative effects of immigration on annual earnings, occupational levels, and the employment rate.

The sectoral approach proposed here relies on comparisons of immigration shocks across regions and thus belongs to the "spatial approach" literature pioneered by Card (1990) and Altonji and Card (1991). As explained by Dustmann et al. (2016), the spatial approach can, from a structural perspective, capture the total effect of immigration on native outcomes, at least under the assumption that immigration into one city does not indirectly affect outcomes in others, e.g., through the displacement of natives.[6] Similarly, applying the spatial approach at the sectoral level implicitly assumes that native workers within a sector are principally affected by immigration into that sector, and not into others.

A critical issue facing analyses that rely on spatial comparisons is identification. Immigrants sort into locations, supposedly following employment opportunities. Locations with better opportunities for immigrant workers are plausibly those where demand for labor is higher, potentially confounding the effect of immigration on native wages or employment. The literature has resorted to instrumental variables approaches in order to address this issue, the most popular instrument being a shift-share instrument constructed by interacting the fraction of immigrants from a country who are observed living in a city in a prior reference period by the *national* inflow of immigrants from that country in the current period (and then summing up across origin countries). The instrument thus represents the total influx of immigrants in the current period that would be obtained if

[6]In addition to outmigration of natives, movements of goods and capital across cities can cause the spatial approach to fail to detect the effects of immigration on native outcomes (Borjas 2003). See Appendix 2.A for a formal argument.

new influxes were perfectly correlated with the geographical distribution in the reference period. Nonetheless, many authors have questioned the validity of the shift-share instrument due to the possible spatial correlation between initial immigrant settlement patterns and subsequent growth in employment opportunities (e.g., Reed and Danziger (2007); Borjas (2014); Basso and Peri (2015b), and more recently Goldsmith-Pinkham et al. (2019)).

Perhaps more importantly, a recent paper by Jaeger et al. (2018b) demonstrates that estimates obtained from the shift-share instrument conflate short-run (negative) impacts with long-run recovery processes whenever there is limited change in the composition of immigrant inflows at the national level over time, as has been the case in the US since the 1980s. According to the authors, the only time period in the US where the shift-share instrumental variable approach—or the improved strategy they propose—may be successfully leveraged is the decade 1970-1980, which saw a considerable shift in the country-of-origin composition of US immigration inflows due to the enactment of the Immigration and Nationality Act of 1965. Although they find evidence of negative short-run impacts of immigration on natives' wages based on this earlier period, it is not clear whether these impacts can be extrapolated to current conditions, due, for example, to the secular increase in immigration and the fact that effects may not be globally linear.

In this paper, we leverage a novel partial identification method formalized by Nevo and Rosen (2012) to address the effect of increased immigration on the employment opportunities of native-born workers in the context of the spatial correlation approach. Our partial identification strategy relies on the use of a series of so-called "imperfect instruments:" instruments for the sectoral immigrant share in a given city and year that, although still potentially correlated with the error term (unobserved demand shocks about city and year averages), are plausibly less correlated with it than the regressor itself, albeit in the same direction. In this sense, they represent imperfect instrumental variables or IIVs. Because of the remaining correlation, which violates the exclusion restriction, the IIV estimate is biased. However, Nevo and Rosen (2012) show that under certain conditions,

the IIV estimate can be used as a lower or upper bound to the coefficient of interest.[7] We use their insights to derive upper bounds for the negative effects of immigration on native employment conditions over the period 1990-2011. Because our approach relies on spatial differences, it delivers estimates that are also possibly subject to spatial-arbitrage bias. But since both sources of bias (imperfect instrument and spatial arbitrage) work in the same direction, our estimates are conservative in nature. Nonetheless, we find that in the food service and personal service sectors, immigration impacts are negative, statistically significant, and larger in magnitude than comparable estimates derived in the US context for recent decades. Once we focus on immigrant-exposed occupations, we also find evidence of large negative effects in the construction sector. In these sectors our estimates for earnings effects are consistent with the latest figures derived by Jaeger et al. (2018b) for the earlier decade 1970-1980. Our estimate for the construction sector appears consistent with that derived by Bratsberg and Raaum (2012) for the Norwegian constructor sector.[8]

In terms of empirical implementation, the dual requirement that the correlation between the IIV and the error term be of the same sign as, but of a lower magnitude than, the correlation between the regressor and the error term does have a cost.[9] Our approach focuses on one sector of the economy at a time in order to use as an IIV for the sectoral share of immigrants the share of immigrants *across all sectors*, or *across all other sectors*. These instruments are plausibly correlated with demand pulls that affect native employment/earnings in the sector of interest in the same direction as the sectoral immigrant share: economic booms attract immigrants across all sectors, and they increase employment op-

[7]They also show how one may derive two-sided bounds, but for reasons highlighted below, our setting does not allow such derivation.

[8]Bratsberg and Raaum (2012) rely on differences in immigrant shares *across* construction trades, rather than intercity comparisons. As such, the interpretation of their estimate differs from ours: whereas our estimate can be interpreted as capturing the *total* sectoral effect of immigration on native outcomes—assuming away spatial arbitrage—theirs represents a *relative* effect across construction trades. Following the argument of Dustmann et al. (2016), our estimate is closer to the relevant effect because it encompasses effects of immigration that are common to all construction trades.

[9]One may argue that the shift-share instrument discussed above already constitutes an IIV. In some studies, like Dustmann et al. (2005), the use of the shift-share instrument actually results in a less negative impact of immigration. In others (e.g., Reed and Danziger (2007); Basso and Peri (2015b); Jaeger et al. (2018b)) the estimate becomes more negative but the change is minimal, suggesting that the IIV correlation with the error term remains high in comparison with that of the regressor.

portunities for natives in any given sector. However, since the immigrant share pertains to the entire economy (or the rest of the economy), it is likely less correlated with the sectoral demand pulls than the sectoral immigrant share itself. Our IIV estimates, which are typically much more negative than the OLS estimates, confirm this intuition. Our finding that the immigrant share has a negative effect on natives' employment rate across all six low-skilled sectors also provides some evidence that labor may not be completely mobile across sectors in the short run, underscoring the relevance of a sectoral approach.

Our paper contributes to the literature on the impacts of immigration on the employment conditions of natives in several ways. First, we deploy a novel instrumental variable strategy that represents an alternative to the much criticized shift-share instrument in the context of the spatial correlation approach. Our strategy acknowledges the inherent remaining correlation between our instrument and unobserved sectoral demand shocks but leverages it to derive an upper bound on the negative impacts of immigration on natives' employment conditions. Second, we are able to produce estimates of immigration impacts for a relatively recent period; Jaeger et al. (2018b) show that the shift-share instrument approach may only produce reliable impacts for the period 1970-1980 in the US context. Third, in spite of the fact that spatial correlation estimates may mask larger national effects (Borjas 2003), several of our estimated effects are larger in magnitude than most recent estimates for the US, suggesting that natives can be hurt by immigration in the short run. Fourth, the sectoral approach allows us to provide a nuanced picture of immigration impacts across the economic sectors most exposed to immigration.[10] We relate our findings to critical differences across sectors and occupations regarding goods tradability, immigrant penetration, and skill requirements.

The rest of the article is organized as follows. Section 2.2 discusses recent immigration trends in the sectors we investigate. Section 2.3 describes our data sources. Section 2.4

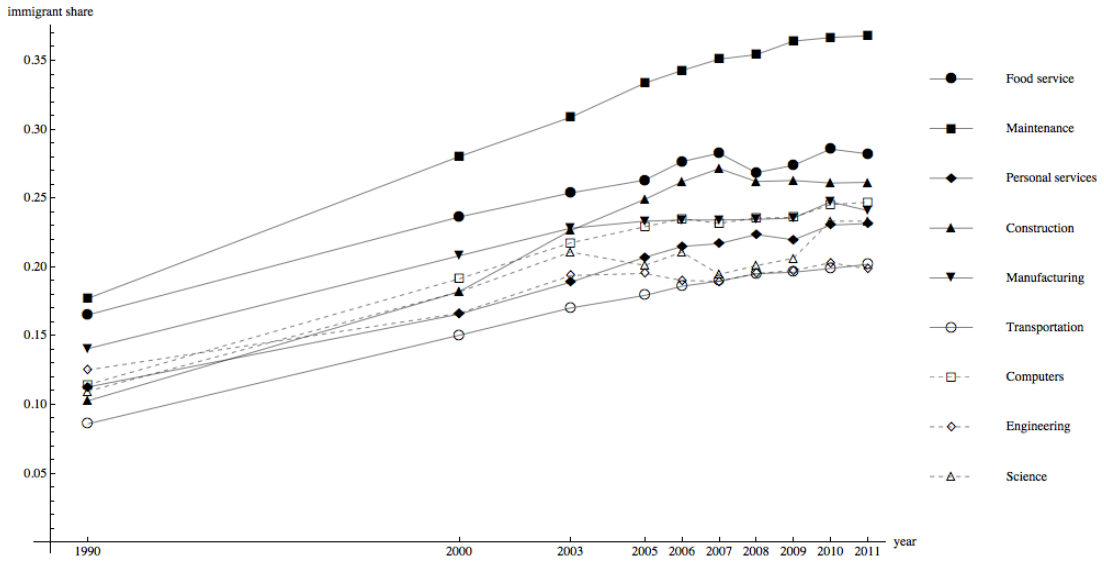
[10]The idea that the effects of immigration on native workers may be more pronounced in industries with high immigrant share was recently explored by Dustmann et al. (2017) in their study of a commuting policy along the German-Czech border. Their main analysis considers all industries together as the spatial variation in immigrant inflows that they exploit cannot address the selection of immigrant workers into industries experiencing positive labor demand shocks. Nonetheless, the results they report in Appendix V.D suggest larger negative effects on wages and employment in immigrant-exposed industries.

describes the IIV strategy we deploy, building upon the work of Nevo and Rosen (2012). Section 2.5 discusses our results, and Section 2.6 concludes.

2.2 Background

Since the enactment of the Immigration and Nationality Act of 1965, the US has experienced a remarkable increase in immigration, with the share of foreign-born individuals in the total population increasing from 4.7% in 1970 to 13.4% in 2015 (López and Bialik 2017). There were 27,400,000 foreign-born (immigrant) individuals in the US labor force in 2017, representing 17.1% of the total labor force (U.S. Department of Labor 2018). Construction and extraction occupations attracted 9.3% of employed immigrant workers, making these occupations the single category with the highest number of immigrants, and one with an immigrant share of 30.4%. Building and grounds cleaning and maintenance occupations

Fig. 2.1 Evolution of the Immigrant Share by Sector



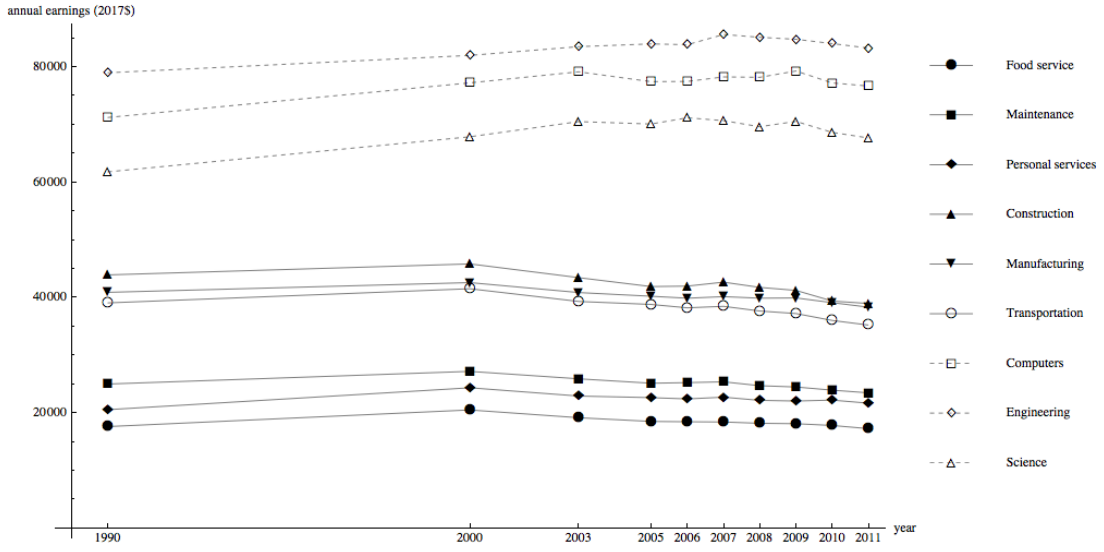
Source: IPUMS data processed by the authors.

Note: Figure was created in the program Mathematica. The immigrant share is calculated over individuals aged 18-64, not living in group quarters, not in school, and in the labor force. Individuals are considered not to be in the labor force if they report being out of the labor force at the time of the survey and working zero weeks during the previous year.

employed 8.4% of immigrant workers, and the immigrant share in that sector reached 37.4%.

Low-skilled sectors of the economy with high immigrant penetration, as defined in this paper, have seen a remarkable increase in the share of immigrant workers in the sectoral workforce (Figure 2.1). According to our data (see Section 2.3), between 1990 and 2011 the share of immigrants in the construction sector has increased from 10% to 26% while that in the maintenance sector has increased from 18% to 37%. Other sectors have seen comparable trends. To the extent that the increase in the sectoral share of immigrant workers has not been uniform across geographical labor markets, this pronounced trend represents an opportunity to empirically identify immigration impacts on the employment conditions of native workers while controlling for common national shocks such as recessions or business cycles.

Fig. 2.2 Evolution of Annual Earnings of Natives by Sector, All Occupations

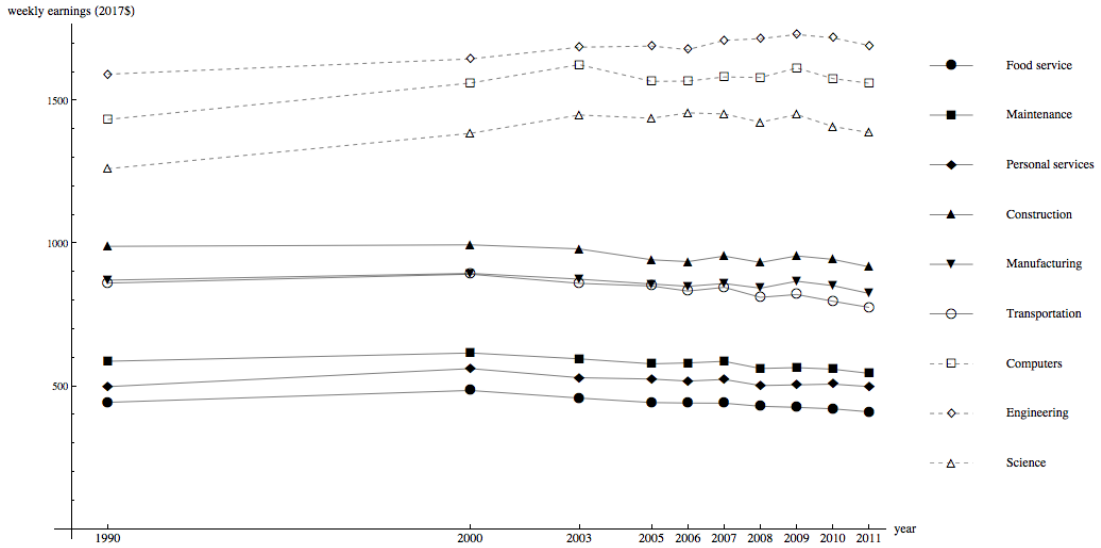


Source: IPUMS data processed by the authors.

Note: Figure was created in the program Mathematica. Earnings are calculated over natives aged 18-64, not living in group quarters, not in school, in the labor force, and with annual earnings above zero and below $300,000 (in 2017$). Individuals are considered not to be in the labor force if they report being out of the labor force at the time of the survey and working zero weeks during the previous year. Annual earnings include wage income and income from a person's own business or farm.

The evolution of annual earnings of native workers at the national level is depicted in Figure 2.2 for the period 1990-2011. The figure shows a clear clustering of earnings across the nine sectors considered. Of the nine sectors we analyze, the workers in the computers, engineering, and science sectors are at the top of the income distribution. Their average annual earnings fall between the $60,000 and $80,000 range. Workers in the construction, transportation, and manufacturing sectors have annual earnings that cluster around $40,000. At the bottom of the income distribution are the maintenance, food service, and personal service workers, who have annual earnings that cluster around $20,000. As depicted in Figure 2.3, trends in estimated weekly earnings tell a very similar story.

Fig. 2.3 Evolution of Weekly Earnings of Natives by Sector, All Occupations



Source: IPUMS data processed by the authors.
Note: Weekly earnings are calculated over natives aged 18-64, not living in group quarters, not in school, in the labor force, and with weekly earnings above $50 and below $5,769.23 (in 2017$).

Note that it is difficult to relate the immigrant share to native earnings or employment by simply looking at national-level aggregates. The immigrant share shows a clear upward trend over the period. Earnings appear relatively stable while employment seems to be mostly driven by macroeconomic factors. Indeed, Figure 2.4 shows the unemployment

rate over time for the entire US economy from the Bureau of Labor Statistics. Consistent with our sectoral data, unemployment increased after 2001, and then again after 2008. Importantly, our empirical strategy nets out any common national effects through year fixed effects and relies on differences across MSAs in the evolution of the immigrant share about the MSA average.

Fig. 2.4 Evolution of the National Unemployment Rate

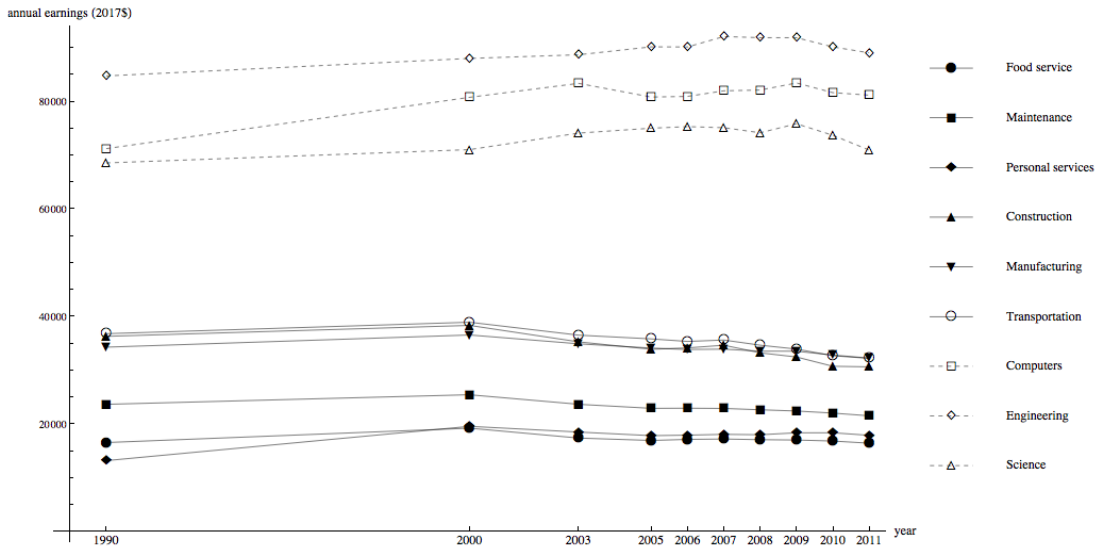
Source: Bureau of Labor Statistics, seasonally adjusted monthly unemployment rate for individuals 16 years and older.

Note: Figure was created in the program Mathematica.

Figure 2.5 depicts annual earnings for native workers in occupations with the highest immigrant shares, by sector. When compared to Figure 2.2, the figure shows that immigrants tend to select into lower-paying occupations within each of the low-skill sectors. The opposite holds for high-skill sectors.

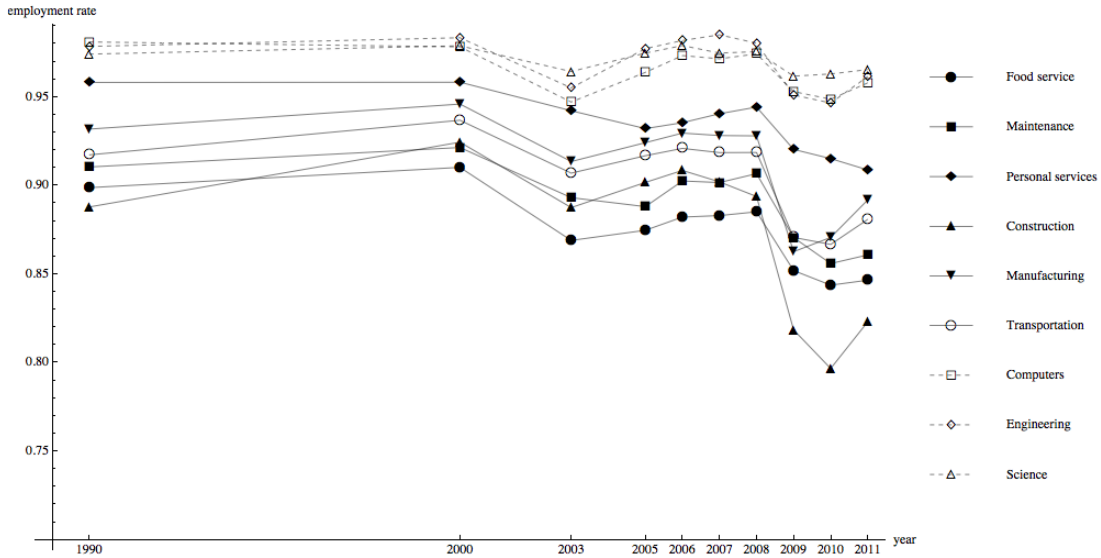
Figure 2.6 depicts the evolution of employment by sector over the period 1990-2011. The figure shows the early effects of the US subprime mortgage crisis on employment in the construction sector, followed by cascading effects on employment in other sectors during the Great Recession.

Fig. 2.5 Evolution of Annual Earnings of Natives by Sector, Immigration-exposed Occupations



Source: IPUMS data processed by the authors.

Fig. 2.6 Evolution of the Native Employment Rate by Sector



Source: IPUMS data processed by the authors.

Note: The employment rate is calculated over natives aged 18-64, not living in group quarters, not in school, and in the labor force at the time of the survey.

2.3 Data

The data used for this analysis were obtained from the Integrated Public Use Microdata Series (IPUMS) provided by the University of Minnesota (Ruggles et al. 2017). These data include US Census data from the 1990 5% State sample and 2000 5% sample as well as American Community Survey (ACS) data between the years 2001 and 2011. Due to a missing geographic variable ("metarea") used to assign a location to workers, the years 2001, 2002, and 2004 are excluded from our data set, and our analysis does not extend beyond 2011. Our analysis is conducted separately for several sectors of the economy as identified by the Census Bureau's 2010 classification using the variable "occ2010." This variable provides a "consistent, long-term classification of occupations" (Ruggles et al. 2017), which identifies sectors of the labor market as well as individual occupations within each sector.

Our analysis is conducted on the following sectors of the US economy: Food Preparation and Serving ("food service"), Building and Grounds Cleaning and Maintenance ("maintenance"), Personal Care and Service ("personal services"), Construction, Production ("manufacturing"), Transportation and Material Moving ("transportation"), Computers and Mathematical ("computers"), Life, Physical, and Social Science ("science"), and Architecture and Engineering ("engineering"). These sectors are selected because, excluding agriculture, they have the highest immigrant worker penetration across all the economic sectors in the U.S. (see Figure 2.1).

In order to reduce attenuation bias caused by measurement error, our analysis is conducted on the largest 150 MSAs in terms of population. Smaller MSAs are likely to include only a few surveyed individuals from a given sector in a given year, which can lead to noisy measures of our regressor of interest (the sectoral share of immigrants working in each MSA).

The data we use is a repeated cross-section of individual-level data that includes the annual earnings of the individual during the preceding year, the number of weeks worked

in the preceding year, the employment status (employed/unemployed/out of the labor force), the MSA where the individual lives (taken to be the relevant labor market), their birthplace, as well as information about educational attainment, race, gender, and marital status. This last set of variables is used to construct "residualized" dependent variables that are purged of potentially confounding demographic factors (see Section 2.4.1). The birthplace variable is used to select natives and to construct the sectoral and multi-sectoral immigrant shares. Between 2008 and 2011, the variable identifying the number of weeks worked ("wkswork2") is only available as a categorical variable that assigns individuals to time intervals (e.g., 50-52 weeks). We transform this variable into a continuous one by assigning the midpoint of the relevant interval as the number of weeks worked.

The income amounts reported in the surveys are nominal values. We convert these values to constant 2017 dollars using the CPI provided by the Bureau of Labor Statistics for all items (US city average, all urban consumers) at `https://www.bls.gov/data`. The income values for the 1990 and 2000 Census years represent income from the previous calendar year, and the ACS data between 2001 and 2011 report income from the previous 12 months. We adjust the 1990 (respectively 2000) income values using the corresponding 1989 (respectively 1999) CPI values, but we use the CPI values corresponding to the sample years for the ACS samples.[11] We define an individual's annual earnings as the sum of their wage income and the income from their own business or farm.[12] We compute weekly earnings by dividing annual earnings by the number of weeks worked.

To avoid outliers, when generating regional averages of the earnings variable we exclude workers reporting annual earnings of $300,000 or more. Since our dependent variable is the logarithm of earnings, we also exclude individuals reporting zero earnings and those for which the reported value is $1 (a code for "breaking even"). When generating

[11] Because the ACS is administered throughout the year, income amounts reported by individuals surveyed in January will represent mostly income generated in the previous calendar year, and they will represent income generated mostly during the current year for those surveyed in December. Although the Census Bureau provides a variable that attempts to adjust for this, the adjustments are imperfect, and Ruggles et al. (2017) find that the adjusted and unadjusted income values are essentially perfectly correlated. As a result, Ruggles et al. (2017) does not recommended using the adjustment variable, thus we refrain from using it.

[12] That is: we use the variable "incearn" provided in the IPUMS data set.

the weekly earnings variable, we exclude workers making at least $300,000/52 = $5,769.23 per week and those making $50 per week or less.

To make sure that we capture the effect of immigration on individuals who are

Table 2.2 Summary Statistics

			All occ.		Exposed occ.	
Sector	Variable	Unit	Mean	S.D.	Mean	S.D.
Personal Service	Annual earnings (natives)	2017$	21,433	4,666	16,756	4,927
	Weekly earnings (natives)	2017$	511	96	430	105
	Employment Rate (natives)		0.934	0.048	0.919	0.073
	Share of immigrant workers		0.157	0.142	0.190	0.173
	Share of immigrant workers in all other industries		0.137	0.113		
Food Service	Annual earnings (natives)	2017$	17,909	3,843	16,763	3,866
	Weekly earnings (natives)	2017$	444	82	424	86
	Employment Rate (natives)		0.877	0.055	0.871	0.064
	Share of immigrant workers		0.218	0.172	0.243	0.188
	Share of immigrant workers in all other industries		0.134	0.112		
Construction	Annual earnings (natives)	2017$	40,261	7,504	32,412	7,189
	Weekly earnings (natives)	2017$	944	158	811	155
	Employment Rate (natives)		0.871	0.069	0.839	0.087
	Share of immigrant workers		0.194	0.176	0.246	0.211
	Share of immigrant workers in all other industries		0.134	0.111		
Maintenance	Annual earnings (natives)	2017$	24,136	4,840	22,003	4,622
	Weekly earnings (natives)	2017$	581	101	543	98
	Employment Rate (natives)		0.891	0.057	0.884	0.061
	Share of immigrant workers		0.262	0.222	0.273	0.230
	Share of immigrant workers in all other industries		0.132	0.109		
Transportation	Annual earnings (natives)	2017$	36,474	4,972	33,297	4,758
	Weekly earnings (natives)	2017$	825	103	764	98
	Employment Rate (natives)		0.904	0.047	0.892	0.053
	Share of immigrant workers		0.152	0.149	0.165	0.161
	Share of immigrant workers in all other industries		0.137	0.112		
Manufacturing	Annual earnings (natives)	2017$	40,050	6,735	33,370	6,053
	Weekly earnings (natives)	2017$	884	136	764	127
	Employment Rate (natives)		0.912	0.048	0.894	0.062
	Share of immigrant workers		0.217	0.180	0.264	0.209
	Share of immigrant workers in all other industries		0.132	0.111		
Computers	Annual earnings (natives)	2017$	70,659	11,937	73,577	13,439
	Weekly earnings (natives)	2017$	1,450	237	1,508	267
	Employment Rate (natives)		0.966	0.043	0.968	0.045
	Share of immigrant workers		0.148	0.116	0.166	0.129
	Share of immigrant workers in all other industries		0.137	0.114		
Engineering	Annual earnings (natives)	2017$	79,959	11,846	85,681	13,884
	Weekly earnings (natives)	2017$	1,630	235	1,736	278
	Employment Rate (natives)		0.970	0.042	0.976	0.049
	Share of immigrant workers		0.144	0.119	0.155	0.144
	Share of immigrant workers in all other industries		0.137	0.114		
Science	Annual earnings (natives)	2017$	65,118	14,197	67,607	18,958
	Weekly earnings (natives)	2017$	1,343	272	1,389	370
	Employment Rate (natives)		0.972	0.052	0.970	0.076
	Share of immigrant workers		0.149	0.123	0.201	0.175
	Share of immigrant workers in all other industries		0.138	0.114		

Note: These figures are based on the 150 most populous MSAs and are representative of an average MSA in our sample during the period 1990-2011. Figures will differ from national averages due to differences in MSA populations. The statistics for the Personal Service, Food Service, Construction, Maintenance, Transportation, and Manufacturing sectors are calculated with 1,387 observations. The number of observations used to generate the statistics in the Computers, Engineering, and Science sectors varies between 1,363 and 1,386 depending on the variable and sector. To define immigrant-exposed occupations ("Exposed occ."), we select occupations with the highest immigrant shares within a sector, until the total number of native workers in those occupations exceeds 50% of the total native workforce in the sector.

actually in the workforce, we follow the literature by including only working-age adults (18 to 64 years of age) who are not in school and do not live in group quarters (e.g., jails or other institutions). Because we perform our analysis at the MSA-year level, our analysis only considers individuals that are identified in the data as living in a specific MSA.[13] In addition, we exclude individuals whose birthplace is not identified and those who jointly report being out of the labor force at the time of the survey and working zero weeks during the previous year.

The individual-level data samples used in our analysis are "weighted," and as such IPUMS recommends using weighted averages to construct variables that are representative at the regional level. We follow this recommendation by applying the personal weights (variable "perwt") provided in the data sample to generate our immigrant share regressors and instruments. The resulting MSA-panel data sets in each sector are unbalanced as some MSAs are not represented in the year 2003.[14]

Table 2.2 summarizes our data. Note that the mean and standard deviations are calculated across MSAs and years. Since MSAs have different population sizes, the mean values are representative of an average MSA included in our analysis rather than national averages.

2.4 Methodology

The main difficulty in measuring the effect of immigration on the labor market outcomes of native-born workers in a sector $s = 1, \ldots, S$ of the economy is that increases in immigration in sector s are likely correlated with unobserved demand-pull factors in that sector that also affect natives' earnings and employment. We use the imperfect instrumental variable approach described below to partially identify the effect of immigration on the labor market outcomes of native-born workers by economic sector.

[13]That is, we ignore individuals for which the MSA identifier is missing in the data.
[14]Our results are robust to removing the year 2003 from the sample.

2.4.1 Model Specification and Instrument Choice

We estimate a sectoral regression of the form

$$y_{it}^{s} = \beta p_{it}^{s} + \alpha_i + \phi_t + \epsilon_{it}^{s} \tag{2.17}$$

where i denotes a metropolitan statistical area, t denotes a year, α_i and ϕ_t are fixed effects, p_{it}^{s} is the immigrant share in sector s, and y_{it}^{s} is the outcome of interest.

Our outcome variables include the natural logarithm of the annual or weekly earnings of native-born workers, the proportion of native-born workers working full time, and the proportion of natives in the labor force who are employed—our measure of the native sectoral employment rate.[15] In order to identify the effect of immigration on the distribution of native workers across occupational levels, we use several definitions of "full-time" workers: workers who worked 48 weeks or more, 40 weeks or more, 27 weeks or more, 14 weeks or more, and 1 week or more.[16]

In order to address MSA-specific changes in the composition of the native sectoral workforce over time, for each choice of dependent variable (say the logarithm of annual earnings) we first regress individual-level observations around a set of observable individual characteristics related to gender, marital status, race, education, and work experience as well as a full set of MSA-year fixed effects. The estimated MSA-year fixed effects are then used as the dependent variable in the IIV regression on immigrant shares (see Appendix 2.C).[17]

Our main regressor is a measure of immigration defined as the fraction of foreign-born workers in sector s relative to the total workforce in that sector in each MSA and year,

[15]The proportion of employed workers is calculated by dividing the number of respondents indicating being employed in the previous week by the number of respondents indicating either being employed in the previous week or having been looking for a job in the previous four weeks.

[16]These cutoffs are chosen to match the categories defined by the variable "wkswork2," which identifies a range of weeks worked for each individual (instead of the actual number of weeks worked) in the IPUMS data starting in 2008.

[17]This technique mirrors that used by Reed and Danziger (2007) in a cross-sectional context. We get very similar results if instead we average the dependent variable over observations in each MSA-year cell using the personal weights provided in IPUMS.

p^s_{it}. Apart from the fact that we focus on one sector at a time and do not differentiate by skill, this is the same regressor as that used by Altonji and Card (1991), Borjas (2003), Borjas (2014), or Llull (2017), and it is directly related to the one used in Dustmann et al. (2005) (the ratio of immigrant to native workers) or Bratsberg and Raaum (2012) (a transformation thereof).

In a recent review of George Borjas' *Immigration Economics*, Card and Peri (2016) criticize the use of the immigrant share regressor on the grounds that due to possible *native* inflows correlated with demand pulls that affect native wages, the regressor might be negatively correlated with the error term, resulting in *negative* bias on the correlation of interest. If both immigrants *and* natives are attracted to areas with positive demand pulls, whether the immigrant share is positively or negatively correlated with the error term ultimately depends on whether natives or immigrants are more responsive to these pulls. It seems reasonable to believe that the immigrant population would respond more promptly to local demand shocks than natives, so that the net bias, in fact, remains positive. A basic reason why the immigrant population would be more responsive is that in any given period, part of this population is migrating from abroad (current inflow), i.e., it is already mobile (Borjas 2001). In addition, Card (2001)'s results suggest little migratory response of natives to immigration shocks, while Cadena and Kovak (2016) show that low-skilled Mexican-born immigrants respond much more strongly to local labor demand shocks than natives, even after arrival. Card and Peri (2016)'s preferred regressor, used in a regression where the dependent variable is the growth in wages rather than the current wage, is defined as the ratio of the *current inflow* of immigrants to the *previous workforce* (including natives and previously arrived immigrants). Our specification reflects the idea that it is the stock of foreign-born workers, rather than the current inflow, that may affect native wages.[18]

[18]Card and Peri (2016) also argue that their regressor better captures Borjas' "relevant wage elasticity," defined as the derivative of the log wage of a given skill group with respect to the "immigration-induced percent increase in the labor supply of (the) group." Borjas defines the immigration-induced percent increase in the labor supply as the ratio of *current foreign-born workers* to *current US-born workers*. With this definition, the relevant wage elasticity can be directly deduced from the estimate of the coefficient on the immigrant

Our instruments are variables that measure the proportion of immigrants across many sectors of the economy, including (respectively excluding) sector s itself (p_{it}^{S}, respectively p_{it}^{S-s}). The instrument p_{it}^{S} is calculated across all economic sectors and corresponds to the regressor used in a spatial correlation approach that considers all sectors of the economy rather than one in isolation. The instrument p_{it}^{S-s} removes the contribution of sector s itself to the immigrant share and it is our preferred instrument. We also use two alternative instruments in our analysis for comparison. The first one, denoted p_{it}^{10}, is a variant of the instrument p_{it}^{S} constructed using the ten economic sectors with the highest proportions of immigrants. The second one, denoted p_{it}^{S-sPop}, is a variant of the instrument p_{it}^{S-s} that uses the share of immigrants in the entire sectoral population, which includes individuals who are not part of the active workforce at the time of the survey but identify as belonging to the sector.

Although p_{it}^{S} and p_{it}^{S-s} are likely correlated with unobservable labor demand shocks in sector s, perhaps due to macroeconomic shocks that affect all sectors,[19] they are plausibly less correlated with the sectoral error term than the sectoral regressor p_{it}^{s}, making them good candidates for an imperfect instrument. Still, the ability of imperfect instruments constructed using information from other economic sectors to improve on OLS estimates of the immigrant share-outcome relationship partially hinges on whether labor demand shocks about MSA (α_i) and year (ϕ_t) effects are sufficiently heterogeneous across sectors. That is, if labor demand shocks were perfectly correlated across sectors, there would be no reason to expect much bias reduction from the use of the imperfect instrument. Fortunately, this seems not to be the case in our data. For instance, a regression of the log

share (Borjas 2003), whereas it cannot be deduced from Card and Peri (2016)'s regression (unless there are only two periods, no immigrants in the first period, and no change in the native workforce between periods, see Appendix 2.B). An important difference between the two specifications is that Borjas' specification considers that immigrants affect native outcomes in levels irrespective of the timing of their arrival, whereas Card and Peri (2016)'s specification identifies effects from changes in outcomes using the most recent inflow, measured relative to the previous workforce irrespective of its immigrant-native composition. The fact that we exploit year-to-year variation to identify short-run effects, coupled with the fact that our panel is missing some years, makes the latter approach less justifiable in our context.

[19] Another reason why p_{it}^{S} may be correlated with sectoral demand pulls is that sector s itself is included in the calculation of the immigrant share.

annual earnings of natives at the MSA by year by sector level (using the six low-skilled economic sectors that are the main focus of this study) on a set of sector and MSA-by-year fixed effects has a R-squared of 0.73, meaning that a significant amount of variation remains in the outcome after netting out common shocks.

In addition to the fact that common macroeconomic shocks may result in a correlation between the overall immigrant share (say p^S_{it}) and sectoral labor demand shocks (ϵ^s_{it}), overall immigration may have a direct effect on native labor demand in a given sector. In particular, one may worry that although sectoral immigration may hurt natives in that sector because native and immigrant labor are substitutable, immigration as a whole may affect the economy in ways that improve natives' employment conditions.[20] In that case, our instrument would be an omitted variable of the following underlying immigration-native outcome relationship:

$$y^s_{it} = b_1 p^s_{it} + b_2 p^S_{it} + a_i + f_t + e^s_{it}.$$

While our framework does not allow identification of a causal relationship between *overall* immigration and *sectoral* outcomes, it can speak to the sign of the bias that would be caused on the estimate of the sectoral effect b_1. Denoting by $\tilde{p}^s_{it}$ the residuals of a regression of p^s_{it} on location and time fixed effects, our IIV estimate of b_1 has a probability limit equal to

$$\beta^{\text{IV}}_{p^s} = b_1 + b_2 \frac{\text{var}(p^S)}{\text{cov}(\tilde{p}^s, p^S)} + \frac{\text{cov}(p^S, e^s)}{\text{cov}(\tilde{p}^s, p^S)}.$$

It is natural to assume that $\text{cov}(\tilde{p}^s, p^S) > 0$. It is then clear that if overall immigration improves sectoral native labor outcomes ($b_2 > 0$), then our estimate of b_1 will be biased upwards. That is, our IIV strategy provides a conservative estimate of the sectoral impact of sectoral immigration.

We provide results for all IPUMS occupations within each sector, as well as results

[20] This situation can be thought of as a variant of Ottaviano and Peri (2012)'s argument that immigration-induced shocks to a skill group have effects on wages in other skill groups.

pertaining to what we call "immigrant-exposed occupations" within a sector. To define immigrant-exposed occupations, we select occupations with the highest immigrant shares within a sector, until the total number of native workers in those occupations exceeds 50% of the total native workforce in the sector. Notably, we do not redefine our regressor of interest (or the imperfect instruments) when focusing on immigrant-exposed occupations, that is, we look at the effect of the overall sectoral immigrant share on outcomes for workers in occupations with the highest immigrant penetration.

2.4.2 The IIV strategy

We use the results contained in Proposition 2 of Nevo and Rosen (2012). This proposition provides us with a one-sided bound given by the IIV estimate.

For the purpose of this section, let us adopt the same notation as Nevo and Rosen (2012). We write the DGP underlying model (2.17) as

$$Y = X\beta + \boldsymbol{W}\boldsymbol{\delta} + U \tag{2.18}$$

where Y is the dependent variable, X is the sectoral immigrant share, $\boldsymbol{W}$ is a row vector of covariates comprising dummy variables for each MSA and dummy variables for each year, and U is the error term, which satisfies $\mathbb{E}[\boldsymbol{W}'U] = 0$. We denote by Z (or Z_1, when necessary to avoid confusion) our preferred instrument, p_{it}^{S-s}. We denote by Z_2 an alternative instrument, for instance the instrument constructed as the share of immigrant workers in sectors with the ten highest shares of immigrant workers. For two random variables, say X and Y, σ_{xy} denotes the covariance between X and Y. We use σ_x to denote the standard deviation of X. We denote the correlation between X and Y as ρ_{xy}. We further denote by β^{OLS} (respectively β_z^{IV}) the probability limits of the OLS estimator (respectively the IV estimator using instrument Z) of parameter β in Equation (2.18).

We denote by $\tilde{X}$ (respectively $\tilde{Y}$) the residuals from the OLS regression of X (respec-

tively Y) on $\boldsymbol{W}$, that is,

$$\begin{cases} \tilde{X} = X - \boldsymbol{W}\mathbb{E}[\boldsymbol{W}'\boldsymbol{W}]^{-1}\mathbb{E}[\boldsymbol{W}'X] \\ \tilde{Y} = Y - \boldsymbol{W}\mathbb{E}[\boldsymbol{W}'\boldsymbol{W}]^{-1}\mathbb{E}[\boldsymbol{W}'Y] \end{cases} . \tag{2.19}$$

$\tilde{X}$ and $\tilde{Y}$ thus represent the residualized regressor and the residualized outcome variable about location and year effects, respectively. Nevo and Rosen (2012) show that $\tilde{Y} = \tilde{X}\beta + U$. Using the Frisch-Waugh-Lovell theorem (Frisch and Waugh 1933; Lovell 1963) and its extension to IV estimation (Giles 1984), it is straightforward to show that

$$\begin{cases} \beta^{\text{OLS}} = \beta + \frac{\sigma_{\tilde{x}u}}{\sigma^2_{\tilde{x}}} \\ \beta^{\text{IV}}_z = \beta + \frac{\sigma_{zu}}{\sigma_{\tilde{x}z}} \end{cases} . \tag{2.20}$$

To fix ideas, consider the case where the dependent variable represents the annual earnings of natives, which implies that $\sigma_{\tilde{x}u} = \sigma_{xu} > 0$ since unobserved demand pulls would tend to increase native earnings and the immigrant share. Given Equation (2.20), we would expect the OLS estimate to be asymptotically biased upwards. That is, $\beta \leq \beta^{\text{OLS}}$. We now make the following two-part assumption, referred to as Assumptions 3 and 4 in Nevo and Rosen (2012):

Assumption 1 $0 \leq \rho_{zu} \leq \rho_{xu}$.

Assumption 1 implies that the direction of correlation with the error term in (2.18) is the same for the regressor and the instrument, but the "intensity" of the correlation is lessened when using the instrument. In that sense, the instrument is "less endogenous" than the regressor. It is also natural in our setting (and we systematically test this condition) to expect that $\sigma_{\tilde{x}z} = \sigma_{\tilde{x}\tilde{z}} > 0$, that is, the shocks in the immigrant share about city and year means are positively correlated across a given sector and the rest of the economy.[21] Because $\sigma_{zu} \geq 0$ from Assumption 1, Equation (2.20) implies that the IV estimate is also asymptotically biased, in the same direction as the OLS estimate, that is, $\beta \leq \beta^{\text{IV}}_z$. In

[21] $\tilde{Z}$ denotes the residual from the regression of Z on $\boldsymbol{W}$.

addition, $\beta_z^{\text{IV}} < \beta^{\text{OLS}} \Leftrightarrow \sigma_{zu}\sigma_{\tilde{x}}^2 - \sigma_{\tilde{x}u}\sigma_{\tilde{x}z} < 0 \Leftrightarrow \rho_{zu} < \rho_{\tilde{x}u}\rho_{\tilde{x}z} = \rho_{xu}\rho_{\tilde{x}z}$. Importantly, the fact that the instrument be less endogenous than the regressor in the sense of Assumption 1 is necessary, but not sufficient, for the IV estimate to improve on the OLS estimate. In particular, if the correlation between the residualized sectoral immigrant share and its economy-wide counterpart is positive but weak, it could be the case that $\beta^{\text{OLS}} < \beta_z^{\text{IV}}$ even if Assumption 1 holds.

Nevo and Rosen (2012)'s analysis suggests that under our Assumption 1, the verified assumption that $\sigma_{\tilde{x}z} > 0$, and the additional assumption that $\sigma_{\tilde{x}x}\sigma_z - \sigma_x\sigma_{\tilde{x}z} > 0$ (which is also satisfied in our setting), one may be able to improve on the upper bound β_z^{IV} by using a combined instrument defined as $V(1) = \sigma_x Z - \sigma_z X$.[22] The probability limit of the corresponding IV estimator can be derived as

$$\beta_{V(1)}^{\text{IV}} = \beta + \frac{\sigma_x\sigma_{zu} - \sigma_z\sigma_{xu}}{\sigma_x\sigma_{\tilde{x}z} - \sigma_z\sigma_{\tilde{x}x}}. \tag{2.21}$$

Under the above assumptions, it turns out that $\beta_{V(1)}^{\text{IV}} < \beta_z^{\text{IV}} \Leftrightarrow \beta^{\text{OLS}} < \beta_z^{\text{IV}} \Leftrightarrow \beta_{V(1)}^{\text{IV}} < \beta^{\text{OLS}}$. Therefore, the use of $V(1)$ as an instrument does not improve on either β_z^{IV} or even β^{OLS} when $\beta_z^{\text{IV}} < \beta^{\text{OLS}}$. In cases where $\beta^{\text{OLS}} < \beta_z^{\text{IV}}$ however, $\beta_{V(1)}^{\text{IV}}$ improves on β^{OLS}.

Finally, Nevo and Rosen (2012)'s analysis suggests a way to derive a lower bound for our effect of the immigrant share on annual income. The idea, developed in Proposition 5 and Lemma 2 of their paper, is that if the analyst has not only one, but two IIVs, say Z_1 and Z_2, he or she may be able to construct a weighted difference, say $\omega(\gamma) = \gamma Z_2 - (1-\gamma)Z_1$, with $\gamma \in (0,1)$, that satisfies $\sigma_{\omega(\gamma)u} \geq 0$ and $\sigma_{\tilde{x}\omega(\gamma)} < 0$. That is, by differencing the two IIVs, one may be able to obtain a new IIV that is still positively correlated with the error term, but is now negatively correlated with the regressor. The probability limit of the corresponding IV estimator is

$$\beta_{\omega(\gamma)}^{\text{IV}} = \beta + \frac{\sigma_{\omega(\gamma)u}}{\sigma_{\tilde{x}\omega(\gamma)}} \tag{2.22}$$

[22]This instrument $V(1)$ is a limit value of the set of instruments $V(\lambda) = \sigma_x Z - \lambda\sigma_z X$. The authors show that for $\lambda = \lambda^* = \frac{\rho_{zu}}{\rho_{xu}}$, a value unknown to the analyst, the instrument $V(\lambda)$ is valid. Assumption 1 essentially implies that $\lambda^* \in [0,1]$, which is used to derive bounds for β.

implying that $\beta^{\text{IV}}_{\omega(\gamma)}$ constitutes a lower bound for β. Nevo and Rosen (2012) even provide a testable sufficient condition for $\omega(\gamma)$ to meet these requirements for some unknown value $\gamma^* \in (0,1)$: it must be that $\sigma_{\tilde{y}z_1}\sigma_{\tilde{x}z_2} - \sigma_{\tilde{y}z_2}\sigma_{\tilde{x}z_1} < 0$. Even if this condition, which guarantees the existence of a value γ^* from which a lower bound can be derived, is satisfied in our analysis, we have no guidance as to what this value of γ^* should be. In fact, without an additional assumption on γ^* (besides $\sigma_{\tilde{x}\omega(\gamma)} < 0$, which, given $\sigma_{\tilde{x}z_j} > 0$, is equivalent to $\gamma < \bar{\gamma} \equiv \frac{\sigma_{\tilde{x}z_1}}{\sigma_{\tilde{x}z_1}+\sigma_{\tilde{x}z_2}}$), one can only deduce that $-\infty = \beta^{\text{IV}}_{\omega(\bar{\gamma})} < \beta$, that is, the lower bound is uninformative. In what follows, we therefore only report the values of β^{OLS} and the upper bound $\beta^{\text{IIV}} = \min(\beta^{\text{IV}}_z, \beta^{\text{IV}}_{V(1)})$. We report $\beta^{\text{IIV}} = \beta^{\text{IV}}_z$ because the IV estimate does improve on the OLS estimate in the vast majority of regressions.

2.5 Results

We begin by presenting results pertaining to the short-run impact of immigration on the earnings of natives, organized by sector of the economy. We report earnings results for the manufacturing sector and the three higher-skilled sectors in Appendix 2.D, as none of them are statistically significant. We then report short-run effects of immigration on natives' employment rate across all six low-skilled sectors. Employment effects for the three higher-skilled sectors are generally negative, but never significant. They are reported in Appendix 2.E. In order to address the possibility that the negative effects we uncover in low-skilled sectors may be driven by compositional effects (i.e., more productive workers leaving a sector in response to the immigration shock), we then report results for a composite sector defined as the aggregate of these sectors. All our first-stage partial F-statistics for our preferred instrument have the correct sign (that is, $\sigma_{\tilde{x}z} > 0$), are larger than 70, and are not reported.

2.5.1 Earnings Effects

2.5.1.1 Personal Services, Food Service, and Construction

We first report results for the effect of the immigrant share on the annual earnings of native-born workers. We provide results for all occupations within a sector, as well as results pertaining to immigrant-exposed occupations.

Table 2.3 shows that the annual earnings of native workers in the personal service, food service, and construction sectors are negatively affected by the sectoral share of immigrants. Although the OLS estimate is never statistically significant, the IIV estimates often are, and they are much larger in magnitude.

Importantly, the move from the IIV constructed from the share of immigrants in immigration-exposed sectors (IIV-10) to that constructed from the share of immigrants across all sectors (IIV-All) and across all other sectors (IIV-All but) has the expected effect on the point estimate: the effect systematically becomes more negative as the correlation between the IIV and the error term is attenuated. Personal service, food service, and construction all belong to the 10 sectors with the highest immigrant shares and each is therefore included in the calculation of the IIV-10.[23] The attenuation in the correlation between the error term and the series of imperfect instruments likely comes from the fact that in sectors less prone to immigration, a positive shock in labor demand (which we assume would be positively correlated with a positive shock in the demand for labor in the sector of interest, say construction) may not correlate as much with an increase in the share of immigrant workers as in sectors with larger immigrant shares.[24] In addition, since the correlation of interest is with sectoral demand pulls, the fact that the share of immigrants is calculated across a broader set of industries mechanically "dilutes" the correlation with any sectoral-specific shock in labor demand, and completely eliminates it when the

[23] The other sectors are the other three low-skill sectors we study, plus agriculture, computers, engineering, and science.

[24] An alternative explanation may be that labor-demand shocks are more correlated among immigration-exposed sectors than between immigration-exposed and immigration-poor sectors.

Table 2.3 Effect of Immigration on the Annual Earnings of Native-born Workers

		(1) OLS	(2) IIV-10	(3) IIV-All	(4) IIV-All but	(5) IIV-All but Pop
Personal services	All occ.	-0.011	-0.114	-0.499**	-0.658**	-0.591*
		(0.127)	(0.254)	(0.234)	(0.287)	(0.340)
	Exposed occ.	-0.105	-0.597*	-0.955***	-1.233***	-1.231***
		(0.177)	(0.358)	(0.307)	(0.375)	(0.425)
Food service	All occ.	0.056	-0.280**	-0.386***	-0.598***	-0.510***
		(0.066)	(0.136)	(0.131)	(0.202)	(0.196)
	Exposed occ.	0.122	-0.249*	-0.347**	-0.585***	-0.401**
		(0.093)	(0.129)	(0.139)	(0.218)	(0.199)
Construction	All occ.	0.019	-0.029	-0.158	-0.293*	-0.294*
		(0.068)	(0.098)	(0.116)	(0.160)	(0.157)
	Exposed occ.	-0.162**	-0.308***	-0.487***	-0.689***	-0.699***
		(0.082)	(0.116)	(0.135)	(0.194)	(0.200)
Observations		1,387	1,387	1,387	1,387	1,387

Note: All regressions include MSA and year fixed effects. Standard errors are clustered at the MSA level. Column (1) reports the OLS estimate of β in Equation (2.17). Column (2) reports the IV estimate obtained by using the immigrant share in the 10 sectors with the highest immigrant shares. Column (3) reports the IV estimate obtained by using the immigrant share across all sectors. Column (4) reports the IV estimate obtained by using the immigrant share across all other sectors. Columns (2) - (4) use instruments constructed from individuals who were between the ages of 18-64, were not in school, were not living in group quarters, and who were jointly in the labor force at the time of the survey and worked a positive number of weeks during the previous year. Column (5) reports the IV estimate obtained by using the immigrant share of the entire population across all other sectors of the economy. All occ. refers to the analysis conducted on all occupations within the sector. Exposed occ. refers to the analysis conducted on the immigrant-exposed occupations within the sector. * (respectively **, respectively ***) denotes statistical significance at the 10% (respectively 5%, respectively 1%) level.

immigrant share is calculated for all other sectors.

As suggested by Equation (2.20), the tightness of the upper bound afforded by the IIV estimate relative to the OLS estimate is also inversely related to the correlation between the sectoral immigrant share and the instrument. Our results suggest that this effect either reinforces, or at least does not supersede, the changing strength of correlation between the error term and the various instruments. In Appendix 2.F, we formally derive a testable condition that can be used to determine the relative endogeneity of the imperfect

instruments used in our analysis. We use this test to show that the IIV-10 instrument (p_{it}^{10}) is, in fact, the most endogenous instrument used in our analysis and that our preferred instrument, the IIV-All but variable (p_{it}^{S-s}), is the least endogenous.

As explained in Section 2.4.2, the preferred IIV estimate should be interpreted as an upper bound. That is, the true underlying parameter is likely more negative. Our preferred estimates, given by the "IIV-All but" estimate, imply that a 10 percentage point increase in the share of immigrants is associated with at least a 6.6% (respectively 6.0%, respectively 2.9%) decrease in the annual earnings of native workers in the personal service (respectively food service, respectively construction) sector. Table 2.3 further shows that in the personal service and construction sectors, the effect is almost doubled for workers in occupations where the share of immigrants is higher.

On balance, these upper bounds appear large relative to recent econometric estimates reported in the literature. Estimates obtained from location-year or location-year-skill comparisons of average wages across *all occupations* range from -0.22 (Borjas 2003) to positive values (Basso and Peri 2015b). Borjas (2014) reports an estimate of -0.21 for the period 1990-2010 (-0.24 for males) using the same data source as ours, an MSA-year-skill regression and a shift-share instrumental variable approach. Card (2001) reports that city comparisons typically estimate the effect of a 10 percentage point increase in the fraction of immigrants to correlate with a less than 1% decrease in native wages.[25]

There are two essential channels by which the annual earnings of native-born workers may be affected by immigration flows: their wage rate may decrease and/or they may work fewer weeks per year (none in the extreme). The second channel is particularly relevant for the construction sector because construction workers are typically paid per "job." A year's worth of earnings is made up of earnings from a potentially large number of jobs. If workers have difficulty filling in their schedule due to increased competition from cheaper, and perhaps illegal immigrant labor, they may end up with lower annual

[25] Admittedly, our upper bounds fall short of the larger effect on lower-skilled natives' earnings found in Altonji and Card (1991), a 12% decrease for each 10 percentage point increase in the immigrant share. On balance, they are also less negative than the estimate derived for the decade 1970-1980 by Jaeger et al. (2018b).

earnings even if their weekly earnings (annual earnings divided by the number of weeks worked) do not change. A similar remark may hold in certain personal service occupations with high immigrant penetration, like child and personal care, where self-employment is high.

Table 2.4 Effect of Immigration on the Weekly Earnings of Native-born Workers

		(1) OLS	(2) IIV-10	(3) IIV-All	(4) IIV-All but	(5) IIV-All but Pop
Personal services	All occ.	-0.034	-0.036	-0.275*	-0.353*	-0.277
		(0.067)	(0.076)	(0.149)	(0.186)	(0.209)
	Exposed occ.	-0.069	-0.163	-0.409**	-0.525**	-0.496*
		(0.117)	(0.208)	(0.203)	(0.249)	(0.278)
Food service	All occ.	0.052	-0.162*	-0.190**	-0.296**	-0.137
		(0.048)	(0.083)	(0.085)	(0.129)	(0.130)
	Exposed occ.	0.085	-0.167*	-0.191*	-0.325**	-0.086
		(0.063)	(0.091)	(0.103)	(0.158)	(0.143)
Construction	All occ.	0.026	0.020	0.002	-0.037	-0.053
		(0.045)	(0.061)	(0.074)	(0.100)	(0.096)
	Exposed occ.	-0.080	-0.153**	-0.217**	-0.300**	-0.298***
		(0.059)	(0.075)	(0.090)	(0.125)	(0.115)
Observations		1,387	1,387	1,387	1,387	1,387

Note: All regressions include MSA and year fixed effects. Standard errors are clustered at the MSA level. Column (1) reports the OLS estimate of β in Equation (2.17). Column (2) reports the IV estimate obtained by using the immigrant share in the 10 sectors with the highest immigrant shares. Column (3) reports the IV estimate obtained by using the immigrant share across all sectors. Column (4) reports the IV estimate obtained by using the immigrant share across all other sectors. Columns (2) - (4) use instruments constructed from individuals who were between the ages of 18-64, were not in school, were not living in group quarters, and who were jointly in the labor force at the time of the survey and worked a positive number of weeks during the previous year. Column (5) reports the IV estimate obtained by using the immigrant share of the entire population across all other sectors of the economy. All occ. refers to the analysis conducted on all occupations within the sector. Exposed occ. refers to the analysis conducted on the immigrant-exposed occupations within the sector. * (respectively **, respectively ***) denotes statistical significance at the 10% (respectively 5%, respectively 1%) level.

Table 2.4 reports weekly earnings effects. Weekly earnings are constructed by dividing annual earnings by the number of weeks worked. Weekly (or hourly) effects partially mask annual earnings effects insofar as one margin of response to increased immigration

may be the reduction in the quantity of labor supplied by natives. Indeed, a general rule here is that point estimates for weekly effects are smaller in magnitude than for annual earnings effects. For instance, we find that a 10 percentage point increase in the share of immigrant workers causes at least a 3.5% (respectively 3.0%) decrease in the weekly earnings of native personal service (respectively food service) workers. Effects are again more pronounced in the immigration-exposed occupations. Weekly earnings effects are not statistically significant for the construction sector as a whole, although they are for immigration-exposed occupations, with an estimated effect of at least minus 3.0%.

Looking at impacts on occupational levels confirms a redistribution of natives away from full-time and high-time work towards part-time work. Table 2.5 shows the effects on the share of native construction workers having worked at least a certain number of weeks in the past year. Effects are shown for all construction occupations and for immigrant-exposed construction occupations. Preferred IIV estimates (based on p^{S-s}) are statistically significant, and the pattern of increasingly negative effect as the instrument becomes likely less endogenous is maintained. Overall, the estimates suggest that immigration has a negative effect on the occupational level of native construction workers. For instance, a 10 percentage point increase in the share of immigrants is predicted to result in at least a 2.1 percentage point decrease in the share of natives construction workers working at least 40 weeks. For exposed construction trades, the effects are more pronounced (3.9 percentage points).

To get a better idea of the effect of immigration on the occupational level of natives, Figure 2.7 uses the estimates reported in Table 2.5 to depict the shift in the distribution of native construction workers across occupation levels, from unemployed to full-time workers, induced by a 20 percentage point increase in the share of immigrant workers in construction.[26] The initial distribution is constructed by using occupational shares averaged across sample years and MSAs.

We also find that occupational levels of native workers are affected negatively by the

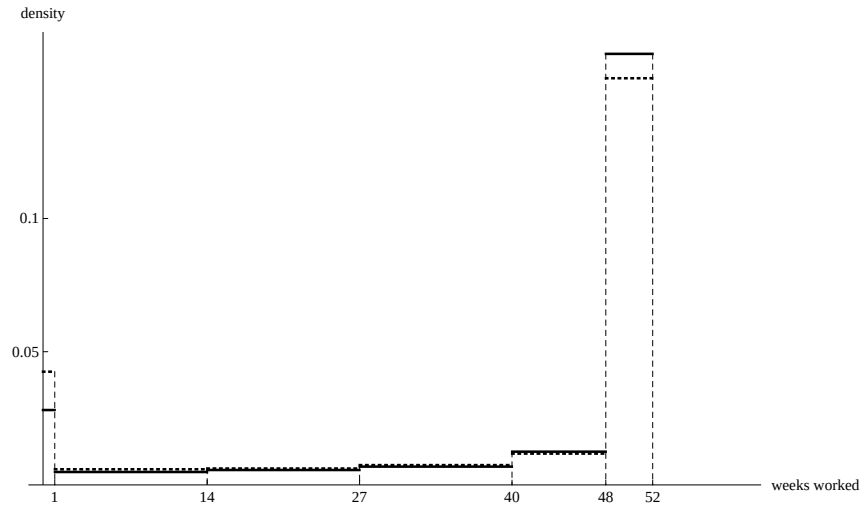
[26] We choose 20 percent rather than 10 percent so that the change in the distribution is more legible.

Table 2.5 Effect of Immigration on the Distribution of Weeks Worked
Among Native-born Construction Workers

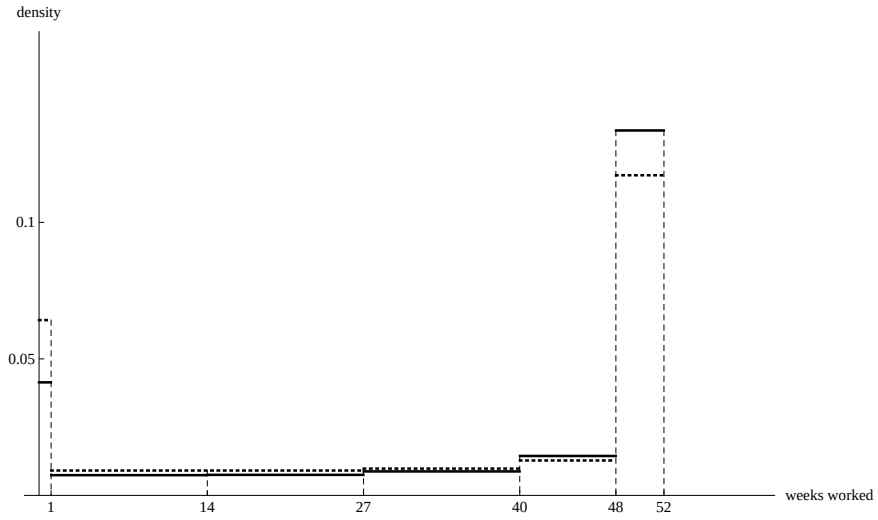
		OLS	IIV-10	IIV-All	IIV-All but	IIV-All but Pop
All occ.	48 weeks or more	0.021	-0.024	-0.098*	-0.183**	-0.181**
		(0.034)	(0.041)	(0.051)	(0.073)	(0.075)
	40 weeks or more	0.034	-0.032	-0.113**	-0.212***	-0.208***
		(0.029)	(0.037)	(0.048)	(0.072)	(0.073)
	27 weeks or more	0.025	-0.020	-0.094**	-0.175***	-0.157***
		(0.024)	(0.029)	(0.038)	(0.058)	(0.057)
	14 weeks or more	0.005	-0.023	-0.079***	-0.138***	-0.122***
		(0.018)	(0.021)	(0.026)	(0.041)	(0.039)
	one week or more	0.010	-0.014	-0.037**	-0.072***	-0.070***
		(0.012)	(0.014)	(0.016)	(0.024)	(0.026)
Exposed occ.	48 weeks or more	-0.033	-0.121**	-0.211***	-0.328***	-0.323***
		(0.047)	(0.059)	(0.076)	(0.109)	(0.104)
	40 weeks or more	-0.034	-0.151***	-0.254***	-0.394***	-0.372***
		(0.038)	(0.051)	(0.069)	(0.102)	(0.099)
	27 weeks or more	-0.007	-0.096**	-0.202***	-0.327***	-0.263***
		(0.032)	(0.043)	(0.057)	(0.087)	(0.083)
	14 weeks or more	-0.035	-0.085***	-0.147***	-0.224***	-0.213***
		(0.029)	(0.031)	(0.038)	(0.058)	(0.051)
	one week or more	0.007	-0.035*	-0.064***	-0.114***	-0.101***
		(0.016)	(0.019)	(0.021)	(0.032)	(0.033)
Observations		1,387	1,387	1,387	1,387	1,387

Note: All regressions include MSA and year fixed effects. Standard errors are clustered at the MSA level. Column (1) reports the OLS estimate of β in Equation (2.17). Column (2) reports the IV estimate obtained by using the immigrant share in the 10 sectors with the highest immigrant shares. Column (3) reports the IV estimate obtained by using the immigrant share across all sectors. Column (4) reports the IV estimate obtained by using the immigrant share across all other sectors. Columns (2) - (4) use instruments constructed from individuals who were between the ages of 18-64, were not in school, were not living in group quarters, and who were jointly in the labor force at the time of the survey and worked a positive number of weeks during the previous year. Column (5) reports the IV estimate obtained by using the immigrant share of the entire population across all other sectors of the economy. All occ. refers to the analysis conducted on all occupations within the sector. Exposed occ. refers to the analysis conducted on the immigrant-exposed occupations within the sector. * (respectively **, respectively ***) denotes statistical significance at the 10% (respectively 5%, respectively 1%) level.

Fig. 2.7 Effect of a 20 Percentage Point Increase in the Immigrant Share on Native Workers' Occupational Levels in the Construction Sector



(a) All Occupations



(b) Immigrant-exposed Occupations

Note: Figure was created in the program Mathematica. The solid (respectively dashed) line represents the distribution of native workers across occupational levels before (respectively after) the increase in immigration.

Table 2.6 Effect of Immigration on the Distribution of Weeks Worked
Among Native-born Food Service Workers

		OLS	IIV-10	IIV-All	IIV-All but	IIV-All but Pop
All occ.	48 weeks or more	-0.010	-0.083	-0.109	-0.161	-0.221**
		(0.038)	(0.057)	(0.072)	(0.111)	(0.091)
	40 weeks or more	-0.010	-0.134**	-0.164**	-0.239**	-0.241***
		(0.034)	(0.056)	(0.065)	(0.096)	(0.091)
	27 weeks or more	-0.015	-0.099**	-0.138**	-0.208**	-0.182**
		(0.027)	(0.048)	(0.061)	(0.096)	(0.087)
	14 weeks or more	0.001	-0.002	-0.015	-0.035	-0.047
		(0.018)	(0.035)	(0.040)	(0.055)	(0.057)
	one week or more	-0.005	-0.007	-0.007	-0.006	-0.006
		(0.014)	(0.024)	(0.019)	(0.019)	(0.017)
Exposed occ.	48 weeks or more	0.004	-0.060	-0.091	-0.143	-0.178*
		(0.036)	(0.055)	(0.066)	(0.100)	(0.094)
	40 weeks or more	-0.017	-0.131***	-0.159***	-0.230***	-0.199**
		(0.038)	(0.050)	(0.058)	(0.086)	(0.088)
	27 weeks or more	-0.012	-0.076	-0.088	-0.139	-0.113
		(0.027)	(0.057)	(0.073)	(0.108)	(0.098)
	14 weeks or more	0.017	0.012	0.014	-0.002	-0.022
		(0.024)	(0.045)	(0.058)	(0.081)	(0.082)
	one week or more	-0.001	-0.001	-0.004	-0.002	-0.003
		(0.021)	(0.035)	(0.033)	(0.031)	(0.049)
Observations		1,387	1,387	1,387	1,387	1,387

Note: All regressions include MSA and year fixed effects. Standard errors are clustered at the MSA level. Column (1) reports the OLS estimate of β in Equation (2.17). Column (2) reports the IV estimate obtained by using the immigrant share in the 10 sectors with the highest immigrant shares. Column (3) reports the IV estimate obtained by using the immigrant share across all sectors. Column (4) reports the IV estimate obtained by using the immigrant share across all other sectors. Columns (2) - (4) use instruments constructed from individuals who were between the ages of 18-64, were not in school, were not living in group quarters, and who were jointly in the labor force at the time of the survey and worked a positive number of weeks during the previous year. Column (5) reports the IV estimate obtained by using the immigrant share of the entire population across all other sectors of the economy. All occ. refers to the analysis conducted on all occupations within the sector. Exposed occ. refers to the analysis conducted on the immigrant-exposed occupations within the sector. * (respectively **, respectively ***) denotes statistical significance at the 10% (respectively 5%, respectively 1%) level.

Table 2.7 Effect of Immigration on the Distribution of Weeks Worked
Among Native-born Personal Service Workers

		OLS	IIV-10	IIV-All	IIV-All but	IIV-All but Pop
All occ.	48 weeks or more	0.006	-0.022	-0.007	-0.008	-0.008
		(0.040)	(0.047)	(0.043)	(0.043)	(0.044))
	40 weeks or more	0.012	0.010	-0.001	-0.001	-0.005
		(0.035)	(0.044)	(0.078)	(0.099)	(0.114)
	27 weeks or more	0.006	-0.046	-0.065	-0.087	-0.102
		(0.031)	(0.066)	(0.068)	(0.085)	(0.093)
	14 weeks or more	0.009	-0.046	-0.072	-0.099	-0.115
		(0.030)	(0.047)	(0.046)	(0.061)	(0.079)
	one week or more	-0.007	-0.036	-0.035	-0.044	-0.057
		(0.012)	(0.028)	(0.028)	(0.036)	(0.044)
Exposed occ.	48 weeks or more	-0.031	-0.101	-0.077	-0.086	-0.176
		(0.063)	(0.155)	(0.167)	(0.210)	(0.225)
	40 weeks or more	-0.028	-0.255**	-0.229*	-0.290*	-0.362**
		(0.056)	(0.130)	(0.123)	(0.153)	(0.174)
	27 weeks or more	-0.021	-0.298**	-0.294**	-0.382***	-0.461***
		(0.054)	(0.122)	(0.116)	(0.145)	(0.161)
	14 weeks or more	0.021	-0.145*	-0.144*	-0.200**	-0.289**
		(0.046)	(0.081)	(0.077)	(0.099)	(0.131)
	one week or more	-0.012	-0.086*	-0.089*	-0.115*	-0.193**
		(0.021)	(0.052)	(0.049)	(0.063)	(0.084)
Observations		1,387	1,387	1,387	1,387	1,387

Note: All regressions include MSA and year fixed effects. Standard errors are clustered at the MSA level. Column (1) reports the OLS estimate of β in Equation (2.17). Column (2) reports the IV estimate obtained by using the immigrant share in the 10 sectors with the highest immigrant shares. Column (3) reports the IV estimate obtained by using the immigrant share across all sectors. Column (4) reports the IV estimate obtained by using the immigrant share across all other sectors. Columns (2) - (4) use instruments constructed from individuals who were between the ages of 18-64, were not in school, were not living in group quarters, and who were jointly in the labor force at the time of the survey and worked a positive number of weeks during the previous year. Column (5) reports the IV estimate obtained by using the immigrant share of the entire population across all other sectors of the economy. All occ. refers to the analysis conducted on all occupations within the sector. Exposed occ. refers to the analysis conducted on the immigrant-exposed occupations within the sector. * (respectively **, respectively ***) denotes statistical significance at the 10% (respectively 5%, respectively 1%) level.

immigrant share in the food service sector (Table 2.6) and, for exposed occupations, in the personal services sector (Table 2.7), thereby contributing to the negative annual earnings effects reported above.

2.5.1.2 Maintenance and Transportation

The results from the maintenance and transportation sectors are much less clear-cut than those in the sectors analyzed in Section 2.5.1.1, at least when considering all occupations within each sector together. Although we find small negative effects on the employment rate of natives (see Section 2.5.2), we do not find significant effects on annual or weekly earnings. However, once we focus on occupations within these sectors with higher immigrant penetration and/or lower skill requirements, we uncover significant negative effects that were masked when these occupations were grouped with higher-skill occupations. For example, the transportation sector includes aircraft pilots as well as laborers who load freight trucks. One would not expect to find low-skilled immigrants competing with aircraft pilots, so including pilots in the analysis is not very informative.

In the transportation sector, which includes many occupations, we select occupations with a high immigrant share. The occupations selected include taxi drivers, truck drivers, vehicle cleaners, packers, etc. In IPUMS, the maintenance sector as a whole only includes four broad occupations: janitors (and supervisors), landscapers (and supervisors), housekeepers, and pest control workers. Among those, the occupations with highest immigrant penetration are landscapers (34.9%) and housekeepers (44.7%). The next high-immigrant occupation is janitorial workers (25.8%). Our data indicates that landscapers and housekeepers also have the lowest average educational attainment in the maintenance sector. We explore immigration impacts for each of the four maintenance occupations, but, perhaps surprisingly, only find significant effects for landscapers. While housekeeping has a high immigrant penetration, housekeepers have by far the lowest annual and weekly earnings of all occupations in the maintenance sector.[27] Therefore, it is plausible that the

[27]For instance, our data indicates that housekeepers' weekly earnings are about one third lower than

absence of an effect is due to earnings having reached a floor below which native workers would stop supplying labor. It is also plausible that as we focus on more narrowly defined occupations, too few individuals are left in the IPUMS data set to construct the MSA averages of immigrant penetration, causing attenuation bias.

Table 2.8 Landscaping, Housekeeping, and Exposed Transportation Occupations

		(1) OLS	(2) IIV-10	(3) IIV-All	(4) IIV-All but	(5) IIV-All but Pop
Landscaping	Annual earnings	-0.011	-0.544**	-0.586**	-0.750**	-0.673*
		(0.106)	(0.251)	(0.289)	(0.376)	(0.365)
	Weekly earnings	-0.039	-0.208	-0.193	-0.221	-0.140
		(0.064)	(0.143)	(0.156)	(0.214)	(0.174)
	Employment rate	0.036	-0.130**	-0.181***	-0.264***	-0.151**
		(0.027)	(0.057)	(0.054)	(0.079)	(0.069)
	Observations	1,386	1,386	1,386	1,386	1,386
Housekeeping	Annual earnings	-0.074	-0.177	-0.106	-0.099	-0.099
		(0.157)	(0.182)	(0.166)	(0.166)	(0.166)
	Weekly earnings	0.025	-0.004	0.006	0.013	0.013
		(0.081)	(0.095)	(0.094)	(0.093)	(0.092)
	Employment rate	0.045	-0.080	-0.071	-0.098	-0.088
		(0.028)	(0.070)	(0.073)	(0.090)	(0.081)
	Observations	1,382	1,382	1,382	1,382	1,382
Exposed transportation	Annual earnings	-0.187*	-0.201	-0.261*	-0.274	-0.212
		(0.098)	(0.145)	(0.142)	(0.167)	(0.165)
	Weekly earnings	-0.060	-0.107	-0.081	-0.084	-0.076
		(0.067)	(0.087)	(0.085)	(0.100)	(0.082)
	Employment rate	-0.053	-0.075*	-0.104***	-0.119***	-0.154***
		(0.032)	(0.038)	(0.036)	(0.044)	(0.044)
	Observations	1,387	1,387	1,387	1,387	1,387

Note: All regressions include MSA and year fixed effects. Standard errors are clustered at the MSA level. Column (1) reports the OLS estimate of β in Equation (2.17). Column (2) reports the IV estimate obtained by using the immigrant share in the 10 sectors with the highest immigrant shares. Column (3) reports the IV estimate obtained by using the immigrant share across all sectors. Column (4) reports the IV estimate obtained by using the immigrant share across all other sectors. All occ. refers to the analysis conducted on all occupations within the sector. Exposed occ. refers to the analysis conducted on the immigrant-exposed occupations within the sector. * (respectively **, respectively ***) denotes statistical significance at the 10% (respectively 5%, respectively 1%) level.

those of landscapers.

Table 2.8 shows selected results for landscaping, housekeeping, and immigrant-exposed transportation activities.[28] We find that a 10 percentage point increase in the share of immigrants causes at least a 7.5% (respectively 2.6%, column (3)) decrease in the annual earnings of landscaping (respectively exposed transportation) workers. This earnings effect appears to be channeled through lower rates of employment in both sectors, as well as a reduced incidence of full-time employment in immigration-exposed transportation occupations (see Table 2.9). For example, in exposed transportation occupations a 10 percentage point increase in the share of immigrants leads to at least a 1.2 percentage

Table 2.9 Effect of Immigration on the Distribution of Weeks Worked in Immigration-exposed Transportation Occupations

	(1) OLS	(2) IIV-10	(3) IIV-All	(4) IIV-All but	(5) IIV-All but Pop
48 weeks or more	-0.092**	-0.119*	-0.109*	-0.119*	-0.140**
	(0.036)	(0.062)	(0.058)	(0.069)	(0.061)
40 weeks or more	-0.078**	-0.123**	-0.112**	-0.122*	-0.151***
	(0.032)	(0.056)	(0.052)	(0.063)	(0.057)
27 weeks or more	-0.072**	-0.086*	-0.073*	-0.075*	-0.102**
	(0.032)	(0.049)	(0.043)	(0.043)	(0.051)
14 weeks or more	-0.060**	-0.063	-0.063	-0.060	-0.060*
	(0.028)	(0.039)	(0.043)	(0.051)	(0.031)
One week or more	-0.006	-0.010	-0.021	-0.025	-0.038
	(0.018)	(0.023)	(0.021)	(0.026)	(0.029)
Observations	1,387	1,387	1,387	1,387	1,387

Note: All regressions include MSA and year fixed effects. Standard errors are clustered at the MSA level. Column (1) reports the OLS estimate of β in Equation (2.17). Column (2) reports the IV estimate obtained by using the immigrant share in the 10 sectors with the highest immigrant shares. Column (3) reports the IV estimate obtained by using the immigrant share across all sectors. Column (4) reports the IV estimate obtained by using the immigrant share across all other sectors. Columns (2) - (4) use instruments constructed from individuals who were between the ages of 18-64, were not in school, were not living in group quarters, and who were jointly in the labor force at the time of the survey and worked a positive number of weeks during the previous year. Column (5) reports the IV estimate obtained by using the immigrant share of the entire population across all other sectors of the economy. * (respectively **, respectively ***) denotes statistical significance at the 10% (respectively 5%, respectively 1%) level.

[28] Results for other maintenance occupations are not statistically significant and are available upon request.

point decrease in the share of workers working at least 40 weeks/year and a 1.2 percentage point decrease in the employment rate.[29] For landscapers, the same change in the share of immigrants leads to a 2.6 percentage point decrease in the native employment rate. This evidence suggests that immigrants are displacing native workers, causing them to work less in some of the occupations in the maintenance and transportation sectors, which is likely attributable to low skill requirements and, subsequently, a high substitutability between natives and immigrants.

2.5.2 Effects on the Sectoral Employment Rate

Table 2.10 reports estimates of the effect of immigration on natives' self-reported employment status. The sectoral employment rate is defined as the share of the active population (those in the sector reporting working the previous week or having been in search of a job for the previous four weeks) who reported working the previous week. It is therefore equal to one minus the sectoral unemployment rate. Results show that the immigrant share has a negative effect on natives' employment rate. In some sectors, these effects are relatively large. For instance, in construction (respectively food service, respectively personal service), a 10 percentage point increase in the share of immigrants causes at least a 2.3 (respectively 1.8, respectively 1.7) percentage point decrease in the employment rate, and larger effects amongst workers in exposed construction and personal service occupations. The mean unemployment rate in our sample of MSA-years lies between 7% for native personal service workers and 13% for native construction workers, so these effects are not trivial. Importantly, we find negative and statistically significant effects in all six sectors, including manufacturing. Overall, these employment effects contrast with the zero to positive correlations reported by Basso and Peri (2015b) for the period 1970–2010.

[29] For a precise definition of the employment rate, see Section 2.5.2.

Table 2.10 Effect of Immigration on the Employment Rate of Native-born Workers

		(1) OLS	(2) IIV-10	(3) IIV-All	(4) IIV-All but	(5) IIV-All but Pop
Personal services	All occ.	-0.043	-0.136***	-0.139***	-0.168***	-0.197**
		(0.028)	(0.050)	(0.050)	(0.064)	(0.078)
	Exposed occ.	-0.070	-0.217***	-0.239***	-0.296***	-0.397***
		(0.046)	(0.072)	(0.077)	(0.101)	(0.153)
Food service	All occ.	0.012	-0.082**	-0.113**	-0.179**	-0.195**
		(0.027)	(0.039)	(0.047)	(0.074)	(0.076)
	Exposed occ.	0.014	-0.075	-0.106*	-0.176*	-0.203**
		(0.033)	(0.049)	(0.061)	(0.095)	(0.094)
Construction	All occ.	-0.054**	-0.099***	-0.158***	-0.234***	-0.223***
		(0.025)	(0.034)	(0.039)	(0.058)	(0.057)
	Exposed occ.	-0.099***	-0.173***	-0.249***	-0.355***	-0.320***
		(0.033)	(0.045)	(0.052)	(0.078)	(0.074)
Transportation	All occ.	-0.037	-0.061**	-0.076***	-0.087**	-0.112***
		(0.028)	(0.031)	(0.028)	(0.034)	(0.035)
	Exposed occ.	-0.053	-0.075*	-0.104***	-0.119***	-0.154***
		(0.032)	(0.038)	(0.036)	(0.044)	(0.044)
Maintenance	All occ.	-0.011	-0.035	-0.049*	-0.062*	-0.067*
		(0.019)	(0.026)	(0.027)	(0.034)	(0.035)
	Exposed occ.	-0.017	-0.044	-0.062**	-0.075**	-0.089**
		(0.020)	(0.028)	(0.028)	(0.035)	(0.038)
Manufacturing	All occ.	-0.036	-0.078**	-0.098**	-0.118**	-0.146***
		(0.024)	(0.036)	(0.041)	(0.050)	(0.048)
	Exposed occ.	-0.045	-0.082*	-0.109**	-0.124*	-0.101*
		(0.028)	(0.048)	(0.055)	(0.067)	(0.059)
Observations		1,387	1,387	1,387	1,387	1,387

Note: All regressions include MSA and year fixed effects. Standard errors are clustered at the MSA level. Column (1) reports the OLS estimate of β in Equation (2.17). Column (2) reports the IV estimate obtained by using the immigrant share in the 10 sectors with the highest immigrant shares. Column (3) reports the IV estimate obtained by using the immigrant share across all sectors. Column (4) reports the IV estimate obtained by using the immigrant share across all other sectors. Columns (2) - (4) use instruments constructed from individuals who were between the ages of 18-64, were not in school, were not living in group quarters, and who were jointly in the labor force at the time of the survey and worked a positive number of weeks during the previous year. Column (5) reports the IV estimate obtained by using the immigrant share of the entire population across all other sectors of the economy. All occ. refers to the analysis conducted on all occupations within the sector. Exposed occ. refers to the analysis conducted on the immigrant-exposed occupations within the sector. * (respectively **, respectively ***) denotes statistical significance at the 10% (respectively 5%, respectively 1%) level.

2.5.3 Results for Pooled Low-Skilled Sectors

Results for the six low-skilled sectors suggest that native workers are affected by immigration through various channels. One concern when focusing on sectoral effects is that native workers may sort across sectors in response to immigration shocks, raising concerns that the estimates may be partly driven by compositional effects. There are at least two reasons to believe that this should not be too much of a concern in our case. First, our immigration effects are derived after residualizing outcomes on a series of observables that include education, race, work experience, gender, and marital status. To the extent that mobility is driven by these observables, our estimates should reflect average net effects.[30] Second, mobility within a sector is already implicitly accounted for in our sectoral estimates, and the definition of our sectors is quite broad. Because individuals who change occupations tend to seek employment in occupations that require a set of skills similar to the one they already have, it is likely that movement between occupations occurs intra-sectorally. Our sectoral definitions are also broad enough to account for potential task specialization and upgrading (Peri and Sparber 2009). For example, our construction sector includes first-line supervisors of construction trades and our maintenance sector includes first-line supervisors of housekeeping and janitorial workers.

Nonetheless, to further address cross-sectoral mobility of natives, we pool the six low-skilled sectors into a composite sector and re-estimate immigration impacts for that sector. Given that these six sectors are all low-skill, in addition to being high-immigrant, it seems unlikely to us that natives working in these sectors could be displaced outside of the composite sector. Therefore, compositional effects should be less of a concern for that sector. In the IPUMS classification, there are two additional sectors that employ primarily workers with a high-school degree or less: extraction (mining and oil drilling) and technical maintenance (mechanics, electrical or equipment repairers, etc.). Although the occupations within these sectors typically appear to require some field knowledge

[30] We also ran the sectoral analysis without residualizing outcomes on observables. The estimates were comparable to those reported above.

and technical training,[31] which may help explain why their immigrant share is smaller,[32] it is plausible that natives working in our six low-skilled sectors could be displaced into them. Therefore, we also report results for a composite sector that includes extraction and technical maintenance, in addition to the other six low-skilled sectors.

Table 2.11 Pooled Low-Skilled Sector Results

	Six Sectors		Eight Sectors	
	(1)	(2)	(3)	(4)
	OLS	IIV-All but-6	OLS	IIV-All but-8
Annual earnings	0.057	-0.302**	0.086	-0.294*
	(0.084)	(0.149)	(0.082)	(0.152)
Weekly earnings	0.100*	-0.059	0.118**	-0.054
	(0.056)	(0.098)	(0.057)	(0.099)
Employment rate	-0.058***	-0.154***	-0.062***	-0.148***
	(0.019)	(0.038)	(0.019)	(0.038)
48 weeks or more	-0.004	-0.105*	0.006	-0.110*
	(0.028)	(0.055)	(0.027)	(0.058)
40 weeks or more	-0.023	-0.140***	-0.020	-0.144***
	(0.025)	(0.048)	(0.025)	(0.050)
27 weeks or more	-0.016	-0.126***	-0.015	-0.134***
	(0.020)	(0.040)	(0.020)	(0.042)
14 weeks or more	-0.001	-0.114***	0.000	-0.119***
	(0.017)	(0.031)	(0.017)	(0.033)
One week or more	0.001	-0.047**	-0.001	-0.048**
	(0.010)	(0.021)	(0.010)	(0.020)
N	1,387	1,387	1,387	1,387

Note: All regressions include MSA and year fixed effects. Standard errors are clustered at the MSA level. Columns (1) and (2) pool the six low-skilled, high-immigrant, sectors previously analyzed in the sectoral results. Columns (3) and (4) additionally include the extraction and technical maintenance sectors into the pooled analysis. Column (1) (respectively (3)) reports the OLS estimate of β in Equation (2.17) while pooling the six (respectively eight) low-skilled sectors together. Column (2) (respectively (4)) reports the IV estimate obtained by using the immigrant share across all sectors of the economy excluding the six (respectively eight) low-skilled sectors used in the analysis. * (respectively **, respectively ***) denotes statistical significance at the 10% (respectively 5%, respectively 1%) level.

[31] In the technical maintenance sector, the share of workers having an AA degree is about twice as high as the average share in the six high-immigrant, low-skill sectors.

[32] In particular, the extraction sector has an immigrant share of only 7.8%.

Table 2.11 reports the upper bounds for the effects of immigration on native outcomes for the composite low-skilled sectors. All upper bounds are negative and statistically significant (except for the weekly earnings effect, which is negative but not significant). When considering the six low-skilled, high-immigrant, sectors, the estimates reveal that a 10 percentage point increase in the share of immigrants causes at least a 3.0% decrease in the annual earnings of low-skilled native workers. The results also demonstrate that the decrease in the annual earnings is channeled through lower employment rates and fewer weeks of work per year. Each 10 percentage point increase in the share of immigrants causes at least a 1.5 percentage point decrease in the employment rate and a 1.1 (respectively 1.4) percentage point decrease in the proportion of low-skilled natives who work at least 48 (respectively 40) weeks per year. When including the extraction and technical maintenance sectors, we find very similar results. If the negative impacts presented in the sectoral analysis were entirely driven by compositional effects due to inter-sectoral mobility, we would expect them to disappear when aggregating the low-skilled sectors. Instead, the results in Table 2.11 suggest that significant negative effects persist even after allowing for plausible occupational mobility.

2.6 Conclusion

Economists have long sought to understand the impacts of immigration on the employment conditions of natives. Recently, there has been renewed interest in this question by policy makers and the general public. While most economists would agree that in the long run, any wage effects of immigration-induced labor supply shocks will be buffered by capital adjustments, there has been disagreement in the empirical literature about whether short-run impacts should even be of concern. Admittedly, the question is difficult to answer. Exogenous labor supply shocks rarely happen in practice, and the use of observational data on wages and employment limits the range and usefulness of the effects that can be estimated empirically (Borjas 2003; Ottaviano and Peri 2012; Dustmann

et al. 2016) while potentially affecting the reliability of the estimates (Jaeger et al. 2018b).

The present study does not purport to completely resolve these fundamental trade-offs. However, it offers a novel approach to the problem—a sectoral analysis that relies on imperfect instruments—as well as meaningful bounds on short-run immigration impacts in important sectors of the US economy for a recent period. We find negative effects of immigration on native earnings in sectors where we would most expect to find them: low-skilled sectors that produce non-traded goods where immigrant penetration has been high in recent decades. The negative effects that we find are perhaps best exemplified by looking at the construction sector, which employs a sizable share of the native and immigrant workforce over the period of our study (an average of 5.9% and 9.4%, respectively, according to our data). In that sector, we find that a 10 percentage point increase in the share of immigrants, which falls short of the historical increase over the period 1990-2011, causes at least a 6.9% (respectively 3.6 percentage point, respectively 3.9 percentage point) decline in the annual earnings (respectively employment rate, respectively full-time occupational rate) of native workers when we focus on immigration-exposed occupations such as painters and roofers. We find qualitatively similar results in other low-skilled sectors of the US economy such as personal services and food service. Our estimated impacts should be interpreted as upper bounds (meaning that the true effect is larger in magnitude) for at least two reasons: first, our strategy does not entirely correct for endogeneity bias, and second, the area-year variation we exploit may mask larger effects due to spatial arbitrage by native workers or capital flows across areas.

Chapter 3: Economic Assimilation of Mexicans and Central Americans in the United States[1]

3.1 Introduction

The economic assimilation of immigrants, usually measured by economists by how their income and employment status compares to that of similarly skilled natives, is a crucial outcome for immigrants and the native-born population (natives). On the one hand, it affects the material and psychological well-being of immigrants. For example, their gains from migration are typically larger if they achieve earnings comparable to those of the receiving country's residents (Clemens et al. 2016). On the other hand, it may contribute to a more favorable opinion about immigrants, which can reduce discrimination and promote equality (Alesina et al. 2018). The United States has historically been a place where immigrants, attracted by economic opportunities, have been able to overcome initial difficulties and succeed economically (Chiswick 1978). While differences among national groups exist, the narrative relative to immigrants who arrived in the U.S. before the 80's is that most of them were able to assimilate economically. Similarly, the evidence on earlier immigrants suggests that they also assimilated economically and, when compared

[1]The research presented in this chapter was conducted in conjunction with co-author Dr. Giovanni Peri, Professor of Economics at the University of California, Davis.

to similar natives, did not have a significant initial earnings gap (Abramitsky et al. 2014).

Several studies have pointed out that the recent history of immigrant assimilation has changed. These studies suggest that there has been a deterioration in the initial immigrant-native earnings gap for the cohorts that arrived in the 80's and 90's, which is putting their economic assimilation at risk. Recently, Borjas (2015) argues that not just the initial gap, but the rate of economic assimilation, measured as the average wage convergence of immigrants over time, has been declining for the cohorts that arrived in the 80's and 90's. These studies paint a picture of a progressive increase in the initial earnings gap and a decline in the "catching up" of more recent immigrants. A decline in economic assimilation would be a cause for worry, as it implies that immigrants arrive with a larger initial disadvantage and are unable to make up for it. As different groups of immigrants are present in very different proportions, depending on the cohort of arrival, the changing wage gap and trajectory of the average immigrant relative to the average native is due, in part, to a composition effect. Borjas (2015) shows that this is partly the case. As migrants of different nationalities have different education levels, age, and initial skills, the changing composition may give a false impression of a deteriorating initial gap and convergence rate. A situation in which immigrants are able to earn wages comparable to similar natives, but where their composition has changed over time in terms of education, age, and place of origin is very different from a scenario in which all immigrants are increasingly lagging behind at arrival and whose assimilation rates are declining. The first scenario is consistent with stable levels of assimilation for each group even if the composition of immigrants is changing. The second implies a decrease in economic assimilation, which could lead to discrimination and create barriers to immigrant participation in the labor market. If some groups of immigrants are assimilating while others are not, an in-depth investigation into the causes of such deterioration for the latter groups would be in order. Although our analysis only focuses on the three largest groups of immigrants in the U.S., we fail to find evidence of deteriorating assimilation in each case.

In this paper, we focus on Mexicans and Central Americans, and we update the

existing picture of assimilation to very recent cohorts. Mexicans and Central Americans represent the largest, and the least economically affluent, group of immigrants in the U.S., hence a documentation of their recent labor and income dynamics is important. We follow the labor market assimilation of different arrival cohorts over time, starting with the cohort arriving between 1965 and 1969 and ending with the one arriving between 2005 and 2011. We first ask whether these immigrants, who usually have low educational attainment and are often employed in manual and low-paying jobs, have performed poorly in terms of their employment probability and earnings *relative to natives of the same age, and then relative to natives of the same age and education*. By focusing on this group, we zoom into the economic assimilation of low-skilled immigrants, and we examine whether it has deteriorated over time. The second contribution to this literature is that we look at the employment probability, in addition to earnings. Mexicans and Central Americans have been employed in many low-skilled jobs, and the general perception is that they work at high rates. Rarely, however, has the employment probability been the focus of analysis in the United States. By following various cohorts of arrival constructed with the U.S. Census and American Community Survey (ACS) data, we also examine how large the attrition of each cohort is, which is often attributed to return migration.

While we do not identify causal determinants of economic assimilation, we determine whether the sector of employment, the location, and the local economic environment are related to the initial gap and assimilation rate of these immigrants. We also discuss the potential role of the changing composition, in terms of education and language proficiency, of the more recent cohorts. Finally, for comparison, we analyze the earnings convergence behavior of two other immigrant groups, which have been quite different from Mexicans and Central Americans in terms of skills, and whose numbers have been growing at a faster rate in the last decade, namely Chinese and Indian immigrants.

Our analysis uncovers four main findings. First, upon arrival, Mexicans and Central Americans had an initial earnings gap of about 40 percent and only cut it in half during the first two to three decades in the U.S., with not much progress after that. Second, we

find that both the initial gap and the speed of convergence have not worsened for recent cohorts of arrival. In fact, the most recent cohorts (the ones that arrived between 1995 and 1999 and between 2005 and 2011) have fared quite well relative to similar natives both in terms of the initial gap and in terms of convergence. However, given that natives at low levels of education have done poorly in the U.S. labor market, and because Mexicans and Central Americans tend to have relatively low educational attainment, recent Mexicans and Central Americans have not performed very strongly overall in terms of earnings. It is important to note that our empirical strategy does not determine the extent to which economic assimilation is driven by a decline in native earnings or employment. It is possible that the economic convergence observed in the data is partially driven by a deterioration of the labor market conditions for natives as a result of increased immigration.

Third, when looking at the employment probability, the picture is more positive. Recent Mexicans and Central Americans have almost no employment rate gap at arrival. In addition, they have consistently surpassed U.S. natives in terms of their likelihood of being employed. Moreover, the employment probability of this immigrant group has become higher for recent cohorts relative to previous ones. This superior performance of low-skilled immigrants in terms of employment rate distinguishes the U.S. from Europe and most other countries, where the reverse is true (see Battisti et al. 2018). When decomposing Mexicans and Central Americans by sector of employment, we find that the initial gaps are smaller and the assimilation rates are faster for immigrants in the construction sector, while their performance is worst in the agricultural sector. The poor assimilation in the agricultural sector is driven partly by the fact that we include farmers, farm supervisors, and farm workers in our analysis, and immigrants are concentrated in the lower paying non-supervisory occupations, where wage growth has been stagnant in recent decades. We also find a somewhat smaller initial earnings gap in urban (rather than rural) areas. By looking at the evolution of observable characteristics of recent arrival cohorts, especially the ones arriving between 1995 and 1999 and between 2005 and 2011, we notice that they are comparable to previous cohorts in terms of the share of Central Americans and English

language skills. Although they have a slightly higher education level and a larger share of nonwhite individuals, these differences do not seem particularly large and are unlikely to drive the significant improvement in the performance of the most recent cohorts.

Finally, by analyzing the other two largest groups of immigrants in the U.S., the Chinese and Indians, who show a much higher average educational attainment than Mexicans and Central Americans, we see that the relative performance of the most recent cohorts (i.e., the ones that arrived between 1995 and 1999 and between 2005 and 2011) was better than the performance of those who arrived in the 70's and 80's. We conclude that the impression of a worse initial gap and declining convergence rate presented in Borjas (2015) is likely an artifact of the changing composition of immigrants, rather than a deterioration in the actual assimilation of immigrants.

The rest of the paper is developed as follows. Section 3.2 frames this paper in the existing literature on assimilation of immigrants. Section 3.3 introduces the data and some aggregate statistics. Section 3.4 covers the empirical models we estimate. Section 3.5 describes the main results on the earnings and employment rate assimilation of Mexican and Central American immigrants. Section 3.6 describes some differences between immigrants by sector of employment and location. Section 3.7 analyzes the economic assimilation of the other two largest groups of immigrants, the Chinese and Indians, and Section 3.8 provides some concluding remarks.

3.2 Economic Assimilation of Immigrants in the Literature

Since the seminal work of George Borjas (Borjas 1985), who showed that in order to analyze the earnings convergence of immigrants, one has to follow a cohort of arrival over time and differentiate across arrival cohorts, the economic literature has followed such an approach. This approach is a significant improvement over the cross-sectional analysis of immigrants first explored by Chiswick (1978), which compares different groups who have been in the country for different periods of time and confounds changes in the initial gap

and changes in assimilation rates across cohorts.

However, even the cohort analysis must be considered with caution. As subsequent cohorts of immigrants in the U.S. have been quite different in terms of their composition (by origin and education), the initial average gap in earnings has changed, and the average convergence in earnings may have varied over time due to changes in composition. Typically, this literature looks at the aggregate set of immigrants and compares them to the average native. If the composition of immigrants and the performance of different groups of natives changes over time, wage dynamics relative to all workers of a certain skill group can be confounded with changes in assimilation rates. Moreover, as the cohort approach does not use longitudinal data, changes in the cohort composition due to attrition from the return migration of individuals to their countries of origin can also be a relevant concern. Using longitudinal data from SIPP (Survey of Income and Program Participation) linked to tax records, Villareal and Tamborini (2018) show that recent arrival cohorts of immigrants have not performed worse than earlier ones and that the race of immigrants affects their assimilation, with black and Hispanic immigrants at a disadvantage. In their study, the authors can actually follow individuals over time, capturing more closely the individual wage dynamics. However, the small size of their sample, the fact that they consider all immigrant groups together, and the fact that they do not compare immigrants to natives of a similar age and education makes their study less informative about the economic assimilation of economically disadvantaged groups of immigrants, such as those that are considered in our analysis.

In this paper we use an approach similar to Borjas (2015), but we focus on a specific and more homogeneous group of immigrants: Mexicans and Central Americans. Moreover, we focus on a comparison with similarly aged and educated natives so that the income and employment dynamics of native groups do not get confused with changes in assimilation rates. While the recent literature on immigrants' income convergence in the U.S. has raised questions about the ability of recent cohorts to assimilate, the literature on the assimilation of immigrants in Europe, which is more recent, has emphasized the

employment gap of immigrants, especially refugees, and their slow convergence.

Evidence from the U.K. (Clark and Lindley 2006), Norway (Bratsberg et al. 2017), and from a set of 13 E.U. countries (Ho and Turk-Ariss 2018) finds a significant initial employment gap of immigrants relative to natives, especially when considering refugees and immigrants from low-income source countries. While some convergence is usually observed, it is far from complete even after 20 years. Several recent papers have looked at what policies have been successful in promoting a more complete and rapid convergence. Using causal inference through regression discontinuity and quasi-experimental evidence on assignment to policies, some recent papers have established that language training (Lochmann et al. 2019), active labor market policies (Sarvimäki and Hämäläinen 2016), and improvements in the processing time of asylum requests (Hainmueller et al. 2016) have improved the labor market assimilation and performance of immigrants. Overall, however, the recent literature expresses concern about the assimilation of recent immigrants, especially refugees in Europe (Fasani et al. 2018). Our paper looks at the economic assimilation of the most vulnerable group of immigrants in the U.S. (Mexicans and Central Americans) and examines whether their initial gap has widened and if their convergence rates have slowed down over time. While we will not provide causal evidence on the effect of policies, we will identify some factors, such as occupation and location, as important correlates of the economic assimilation of this group, and we will discuss characteristics of immigrants as potentially correlated with their assimilation.

3.3 Data and Earnings Gap-Convergence for All Immigrants

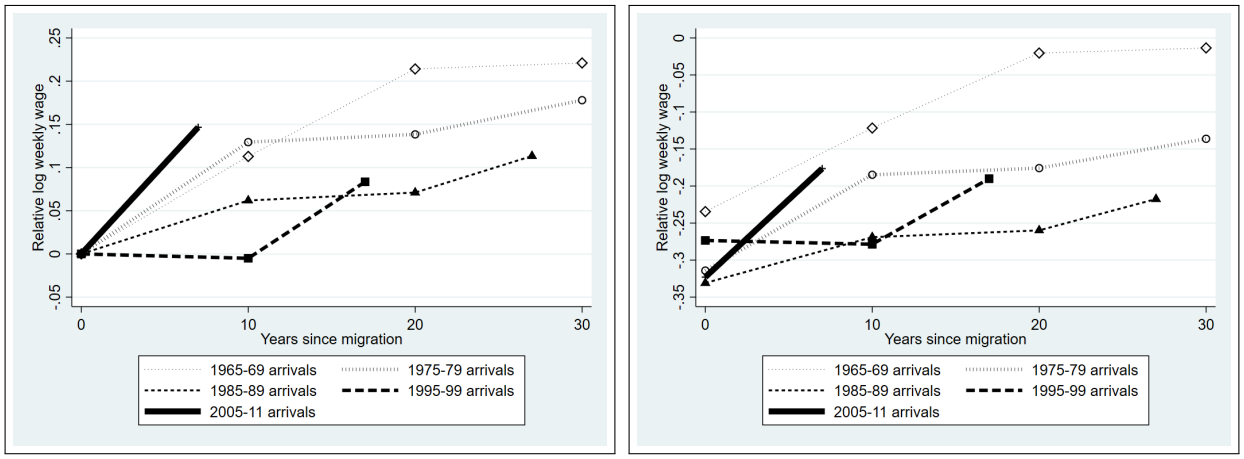
The data we use were obtained from IPUMS (Ruggles et al. 2019) and contain samples similar to those used in Borjas (2015). However, we update our analysis to the year 2017 and document for the first time assimilation in the most recent seven years for which IPUMS data are available. These data include the decennial U.S. Census samples spanning the

period between 1970 and 2000, as well as the pooled 2009-2011 and the 2017 ACS samples[2].

The sample of individuals used in the earnings analysis only includes males between the ages of 25 and 64 who have between 1 and 40 years of potential work experience, worked at least one week during the previous year, were not living in group quarters or attending school at the time of the survey, and arrived in the U.S. at age 18 or older. For the employment rate analysis, the same criteria are used, but individuals who did not work and those who did not generate earnings are also included in the sample as we are constructing the employment rate (employment probability) for this group.[3] We classify individuals as employed if they worked at least one week in the previous year.

Fig. 3.1 Age-Adjusted Convergence for the Relative Weekly Earnings of Immigrant Cohorts from All Countries of Origin

A: Normalized Convergence **B:** Inital Gap and Convergence



Note: The wage differentials presented in this figure are calculated from regressions that are estimated separately for each cross section. The dependent variable in these regressions identifies the log weekly earnings of each individual, and the explanatory variables include a third-order polynomial for age and a set of fixed effects: one for each immigrant cohort. The omitted group is comprised of native-born workers such that the coefficients on the cohort fixed effects represent each cohort's log weekly earnings relative to native-born workers in a given survey year, which are used to construct the data points in the graphs. In Figure 3.1A, the relative log weekly earnings for each cohort is normalized to zero at the time of the entry.

[2]Our samples differ slightly from those used by Borjas (2015) because of errata in the 2009-2011 ACS sample that had not been corrected at the time of Borjas (2015)'s analysis. In particular, on July 1, 2015 (which is after the date that Borjas (2015) had conducted his analysis), IPUMS adjusted the CPI on the source variables (inwage and incbus00) that are used to construct the main outcome variable (incearn) used in the analysis. In addition, on May 25th, 2017, IPUMS made another adjustment to the source variable incwage. Nevertheless, replication exercises using these corrected data reveal coefficients that are either identical or very close (all are within .01) to those reported by Borjas (2015), so we are confident that the updated samples we use will reflect estimates that are comparable to his analysis.

[3]We use the terms "employment rate" and "employment probability" interchangeably throughout this paper.

Table 3.1 Age-Adjusted Relative Log Weekly Earnings of Immigrant Cohorts from All Countries of Origin by Census Cross Section

Cohort	1970	1980	1990	2000	2010	2017
1950-59 arrivals	0.037**	0.032**	0.100**	0.147**	...	...
	(0.000)	(0.002)	(0.003)	(0.010)		
1960-64 arrivals	-0.058**	-0.041**	0.046**	0.074**	0.594**	...
	(0.001)	(0.001)	(0.004)	(0.004)	(0.019)	
1965-1969 arrivals	-0.235**	-0.122**	-0.020**	-0.014*	0.196**	...
	(0.001)	(0.000)	(0.003)	(0.005)	(0.010)	
1970-74 arrivals	...	-0.223**	-0.124**	-0.128**	-0.057**	0.161**
		(0.001)	(0.002)	(0.006)	(0.004)	(0.012)
1975-1979 arrivals	...	-0.314**	-0.185**	-0.176**	-0.136**	-0.118**
		(0.001)	(0.000)	(0.005)	(0.004)	(0.007)
1980-84 arrivals	...	...	-0.285**	-0.236**	-0.206**	-0.188**
			(0.001)	(0.002)	(0.006)	(0.010)
1985-1989 arrivals	...	...	-0.331**	-0.269**	-0.260**	-0.218**
			(0.001)	(0.002)	(0.005)	(0.011)
1990-94 arrivals	...	...	...	-0.269**	-0.271**	-0.168**
				(0.003)	(0.003)	(0.010)
1995-1999 arrivals	...	...	...	-0.273**	-0.279**	-0.190**
				(0.004)	(0.001)	(0.006)
2000-04 arrivals	...	...	...	...	-0.349**	-0.224**
					(0.003)	(0.003)
2005-2011 arrivals	...	...	...	...	-0.323**	-0.176**
					(0.004)	(0.003)
2012-17 arrivals	...	...	...	...	...	-0.103**
						(0.005)
N	945,579	2,002,074	2,373,285	2,708,438	1,653,425	557,077

Note: The wage differentials presented in this table are calculated from regressions that are estimated separately for each cross section, which are identified by the year displayed in the column heading. The dependent variable identifies the log weekly earnings of each individual, and the explanatory variables include a third-order polynomial for the age of the individual and a set of fixed effects: one for each immigrant cohort, including one (not shown in the table) for the cohort that arrived in the U.S. prior to 1950. The omitted group is comprised of native-born workers such that the coefficients in a column each represent a separate cohort's log weekly earnings relative to native-born workers in that survey year. The "2010" cross section is generated from the pooled 2009-11 American Community Surveys. Standard errors in parentheses are clustered at the cohort level. † $p < .1$, * $p < .05$, ** $p < .01$

For the earnings analysis, all dollar amounts have been adjusted to real 1999 dollars using the consumer price index (CPI) for "Current, not seasonally-adjusted, U.S. city average for all items for all urban consumers."[4]

[4]This CPI can be found using the "Multi-Screen Data Search" tool at https://www.bls.gov/cpi/data.htm. Since the census samples report earnings from the previous year, we also use the CPI from the previous year to adjust earnings reported in the census samples. However, the ACS surveys reflect information about the previous 12 months (not the previous calendar year). Following Borjas (2015), we also use the previous

Figure 3.1 and Table 3.1 update the stylized facts shown in Borjas (2015) relative to all immigrants, adding in the cohort that arrived in 2005-2011 and expanding the sample to 2017. Table 3.1 shows the estimates of the log earning gap relative to U.S. natives in the same age group for each cohort of entry in each Census year 1970, 1980, 1990, 2000, 2010, and we add the year 2017 from the ACS data, which allows one more cohort and a longer period of analysis for previous cohorts. Figure 3.1 shows those gaps in a chart, connecting each entry cohort over 30 years of stay in the U.S. We first standardize the initial gap to 0 in Panel A and then show the actual estimated initial gap in log points in Panel B. These initial graphs and table provide a benchmark for the average immigrant in a cohort in terms of the earnings gap upon arrival to the U.S. and the average convergence rate over time. Panel B of Figure 3.1 also reveals that there is a progressively larger initial gap and slower convergence rate for more recent cohorts. In particular, the cohorts that arrived between 1985 and 1989 and between 1995 and 1999, which are the two most recent cohorts considered in Borjas (2015), show large initial gaps and slow convergence rates relative to the previous two cohorts. These figures, however, compare cohorts of immigrants that changed drastically in country of origin and education levels over time. That is, these results compare the average immigrant to the average U.S. native and do not account for education or country of origin, so they only provide a limited understanding of economic assimilation as it relates to the more vulnerable immigrants.

3.4 Methodology and Empirical Specification

In order to estimate the rate of wage and employment convergence for Mexicans and Central Americans, we start by estimating the following model separately for each cross section (τ) while restricting the sample to include only native-born and immigrant workers from Mexico and Central America:

year's CPI to adjust the reported earnings from the ACS samples.

$$Y_{\ell\tau} = \beta_0 + \boldsymbol{\beta} \boldsymbol{C}_\ell + \boldsymbol{\Gamma} \boldsymbol{X}_{\ell\tau} + \epsilon_{\ell\tau}. \tag{3.23}$$

In equation (3.23), $Y_{\ell\tau} \in \big(\ln(w_{\ell\tau}), Emp_{\ell\tau}\big)$ represents the measure of labor market performance of interest. The variable $\ln(w_{\ell\tau})$ is the natural log of the weekly earnings of individual ℓ measured in year (cross section) τ, $Emp_{\ell\tau}$ is a dummy variable that identifies whether an individual was employed during the previous year, $\boldsymbol{X}_{\ell\tau}$ is a third-order polynomial for the individual's age, $\boldsymbol{C}_\ell$ is a vector of fixed effects representing each immigrant cohort in the sample being considered and one fixed effect, omitted in the regression, representing native workers, and $\epsilon_{\ell\tau}$ is the error term.[5] With this notation and convention, the coefficients $\boldsymbol{\beta}$ for the vector of fixed effects $\boldsymbol{C}_\ell$ capture the log weekly earnings or employment rate differential for each immigrant cohort group relative to native workers of the same age after allowing for nonlinear age effects.[6]

Then we pool the information in all cross sections and allow for the comparison of different cohorts of immigrants to similarly aged and educated natives. We estimate the following model, including natives and immigrants from Mexico and Central America:

$$Y_{\ell\tau} = \beta_0 + \boldsymbol{\Gamma} \boldsymbol{X}_{\ell\tau} + \boldsymbol{\alpha} \boldsymbol{y}_{\ell\tau} + \boldsymbol{\beta} \boldsymbol{C}_\ell + \boldsymbol{\theta}(\boldsymbol{y}_{\ell\tau} \boldsymbol{C}_\ell) + \boldsymbol{S}_{\ell\tau} + \epsilon_{\ell\tau}. \tag{3.24}$$

In equation (3.24), $\boldsymbol{X}_{\ell\tau}$ is third order polynomial for the age of each individual, $\boldsymbol{y}_{\ell\tau}$ is a third order polynomial that identifies the number of years in the U.S. capturing the potentially nonlinear effect of U.S. work experience, $\boldsymbol{C}_\ell$ is a vector of dummy variables identifying each immigrant cohort and $\boldsymbol{y}_{\ell\tau} \boldsymbol{C}_\ell$ identifies a cohort-specific additional experience trend. The term $\boldsymbol{S}_{\ell\tau}$ is a vector of age-education-survey year fixed effects.[7] The introduction of such a rich set of skill-by-year effects implies that we are comparing immigrants to natives of the same age in the same age-education group. The estimated

[5] We define employed as working at least one week during the previous year.

[6] All regressions that use equation (3.23) are weighted by the individual sample weights using the variable "perwt."

[7] We include four education groups (high school dropouts, high school graduate, some college, and college diploma) and eight age groups broken into five year intervals between the ages of 25 and 64 years old.

coefficients $\boldsymbol{\beta}$ capture the (log earnings or employment rate) gap of a specific cohort at arrival and the coefficients $\boldsymbol{\theta}$ capture the average decennial growth of that specific cohort of immigrants relative to natives.[8][9]

All the tables that show results from equation (3.24) report the cohort-of-arrival specific initial gap and the 10-year estimated relative growth. The coefficients are estimated first without the age-education-year effects ($\boldsymbol{S}_{\ell\tau}$), so as to capture the earning gap and growth of Mexicans and Central Americans relative to the average native of the same age, and then with the set of age-education-year fixed effects ($\boldsymbol{S}_{\ell\tau}$), so as to capture the gap and convergence relative to natives of a similar age and education level. The difference between those two specifications captures the part of the gap and convergence explained simply by the composition of immigrants across education groups and the different performance of those groups over time, common to natives and immigrants.

3.5 Empirical Findings: Earnings and Employment Convergence for Mexicans and Central Americans

3.5.1 Adjusted Earnings Gaps and Convergence

Figure 3.2 below shows the convergence of log earnings for Mexicans and Central Americans relative to U.S. natives of a similar age, either normalizing the initial gap to 0 (Panel A) or starting from the estimated initial gap (Panel B). This figure is generated from the $\boldsymbol{\beta}$ coefficients in equation (3.23) above. We use dotted lines for the early cohorts, dashed for the intermediate cohorts, and a solid line for the most recent cohort, and we emphasize each subsequent cohort with a line that increases in thickness. Several things are worth

[8]Our model deviates from the one used in Borjas (2015) by constraining the age effects to be equal for natives and immigrants. This allows us to conveniently compare the results from equation (3.24) to equation (3.23), which uses the same constraint for age.

[9]All regressions that use equation (3.24) are weighted by the variable "perwt" divided by the population of the cross section that the observation belongs to.

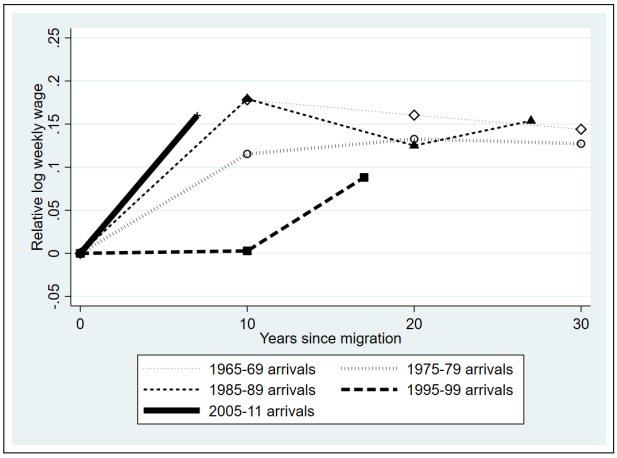
noting. First, while the initial earnings gap is somewhat smaller for the very early cohort (1965-69), the difference is small, and the convergence rate seems roughly comparable across cohorts. Second, the initial gap is substantial (-60 to -70 log points), and it is only reduced by 15 to 20 log points in the first twenty years. Third, the great recession seems to have produced one lost decade of convergence for all cohorts. In Panel B we indicate which segments in the convergence of three cohorts coincide with the period 2000-2010, which is when the great recession took place. Each of those segments is either flat or downward sloping, implying a zero or negative convergence rate in that decade for all cohorts. Finally, for the most recent cohort that arrived between 2005 and 2011, the convergence rate seems to be quite good, with an initial gap comparable to the cohort that arrived in the 70's and a more rapid rate of convergence than any of the previous cohorts. In fact, this cohort closed the earnings gap by 17 log points within 10 years. It may be early to evaluate the economic success of this cohort, but these results are encouraging.

Although Figure 3.2 presents the relative gap and convergence over time, it does not account for the fact that the population of Mexicans and Central Americans in the U.S. has a large concentration of individuals with relatively low educational attainment. If the wages of the less educated have not grown at a rate comparable to the average native-born worker, there can be an appearance of slower assimilation while the reason for slow convergence to the mean could result from increased earnings inequality for both natives and immigrants. In order to clean our analysis from this issue, in Table 3.2 below, we show a comparison of the initial earnings gap and ten year relative earnings growth of each cohort. We first compare immigrants to the average U.S. native of similar age in column 1 and then compare immigrants to U.S. natives with the same age and education level in column 2, reflecting the inclusion of the age-education-year fixed effects in equation (3.24). The table shows three important differences between columns 1 and 2. First, relative to column 1, the results in column 2 show initial gaps that are reduced by one fourth to one third for each cohort. Most cohorts have a gap of 42-43 log points

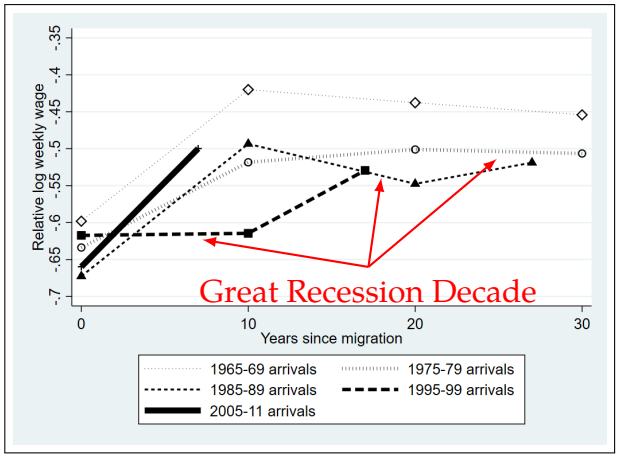
Fig. 3.2 Age-Adjusted Convergence for the Relative Weekly Earnings and Employment Rate of Mexican and Central American Cohorts

Earnings

A: Normalized Convergence

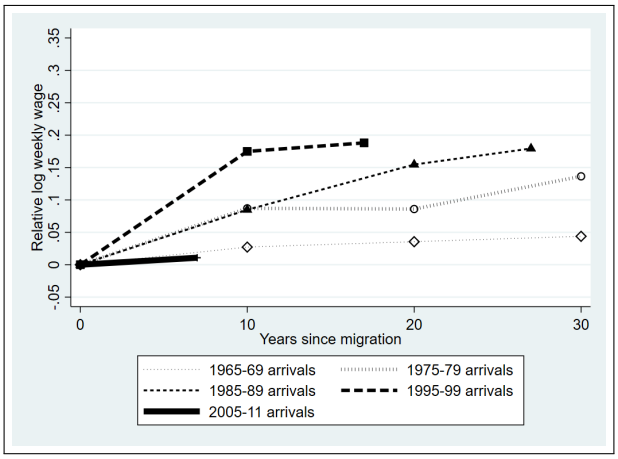


B: Inital Gap and Convergence

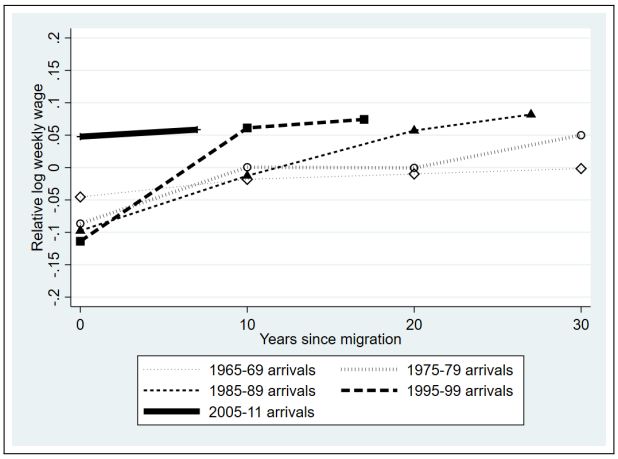


Employment

C: Normalized Convergence



D: Initial Gap and Convergence

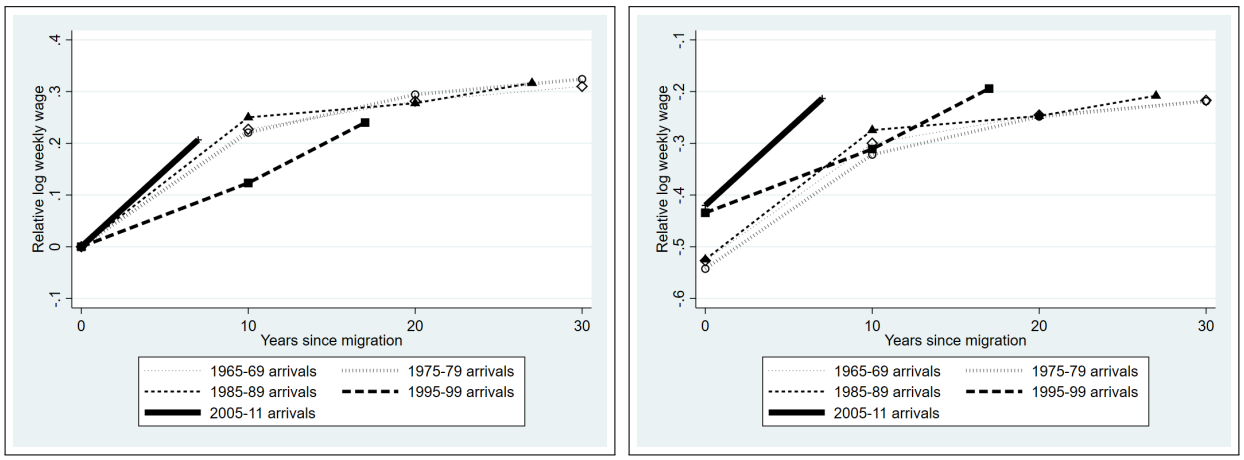


Note: The log weekly earnings and employment rate differentials presented in this figure are calculated from regressions that are estimated separately for each cross section. The dependent variable in the earnings regressions identifies the log weekly earnings of each individual. The dependent variable in the employment regressions identifies whether each individual was employed for at least one week during the previous year. The explanatory variables for both the earnings and the employment regressions include a third-order polynomial for age and a set of fixed effects: one for each immigrant cohort. The omitted group is comprised of native-born workers such that the coefficients on the cohort fixed effects represent each cohort's employment rate relative to native-born workers in a given survey year, which are used to construct the data points in the graphs. In Figure 3.2A, the relative log weekly earnings for each cohort is normalized to zero at the time of entry. In Figure 3.2C, the relative employment rate for each cohort is normalized to zero at the time of the entry.

(about 34-34 percent) when compared to similarly educated natives (see column 2). Second, the convergence is faster, equal to 20 log points (about 19 percent) in the first decade for most cohorts. Third, the two most recent cohorts (the ones that arrived between 2005 and 2011 and between 2012 and 2017) seem to be performing quite well. The initial gap for the 2005-11 cohort was about 40 log points, but it was reduced by half within 10 years, while the 2012-17 cohort only had an initial gap of 24 log points. These encouraging findings are also confirmed in Figure 3.3, which shows the convergence in Panel A and the initial gap and convergence in Panel B while only considering Mexicans, Central Americans, and natives with a high school education or less. This figure reveals that when Mexicans and Central Americans are compared to similarly educated natives, the gap is significantly smaller.

Fig. 3.3 Age-Adjusted Convergence for the Relative Weekly Earnings of Mexican and Central American Cohorts, Only High School Educated or Less

A: Normalized Convergence **B:** Inital Gap and Convergence



Note: The wage differentials presented in this figure are calculated from regressions that are estimated separately for each cross section. The dependent variable in these regressions identifies the log weekly earnings of each individual, and the explanatory variables include a third-order polynomial for age and a set of fixed effects: one for each immigrant cohort. The omitted group is comprised of native-born workers such that the coefficients on the cohort fixed effects represent each cohort's log weekly earnings relative to native-born workers in a given survey year, which are used to construct the data points in the graphs. In Figure 3.3A, the relative log weekly earnings for each cohort is normalized to zero at the time of the entry. In Figures 3.3A and 3.3B, all individuals (both native and immigrant) have a high-school education or less.

Table 3.2 Age-Education-Adjusted Relative Weekly Earnings of Mexicans and Central American Immigrants: Initial Gap and Convergence After First 10 Years

	(1)	(2)
Panel A		
Relative Entry Wage		
1965-1969 arrivals	-0.523**	-0.449**
	(0.0439)	(0.0283)
1975-1979 arrivals	-0.626**	-0.437**
	(0.0418)	(0.0284)
1985-1989 arrivals	-0.670**	-0.445**
	(0.0463)	(0.0342)
1995-1999 arrivals	-0.674**	-0.423**
	(0.0225)	(0.0262)
2005-2011 arrivals	-0.732**	-0.427**
	(0.0159)	(0.0272)
2012-17 arrivals	-0.530**	-0.237**
	(0.00379)	(0.0260)
Panel B		
Relative Wage Growth in First 10 Years		
1965-1969 arrivals	0.081	0.221**
	[0.202]	[.000]
1975-1979 arrivals	0.088	0.216**
	[0.162]	[0.001]
1985-1989 arrivals	0.109	0.198**
	[0.102]	[0.002]
1995-1999 arrivals	0.099*	0.181**
	[0.031]	[0.000]
2005-2011 arrivals	0.189**	0.239**
	[0.000]	[0.000]
Basic Specification	X	–
Age-Educ-Year FE	–	X
N	9,669,594	9,669,594

Note: The wage differentials presented in Panel A are generated from regressions that are ran on the set of pooled cross sections from 1970, 1980, 1990, 2000, 2010, and 2017. The dependent variable identifies the log weekly earnings of each individual. The explanatory variables in column (1) include a third order polynomial for age, a third order polynomial for the number of years that immigrants have spent in the U.S., a set of cohort fixed effects, and a set of cohort fixed effects that are each interacted with a continuous variable identifying the number of years that immigrants have spent in the U.S. The explanatory variables for column (2) contain the same set of variables as in column (1) but additionally include a set of education-age-year fixed effects. The omitted group is comprised of native-born workers such that the coefficients in Panel A each represent a separate cohort's log weekly earnings relative to native born workers. The predicted relative wage growth in the first 10 years in Panel B assumes that all immigrants arrive in the country at the age of 25. Standard errors are in parentheses. P-values are in brackets. † $p < .1$, * $p < .05$, ** $p < .01$

3.5.2 Employment Rate Gaps and Convergence

Mexicans and Central Americans show a substantial initial earnings gap, albeit with convergence and no deterioration for the most recent cohorts. However, a different picture is painted by analyzing the employment rate gap. Panels C and D of Figure 3.2 show the convergence and initial gap for the employment probability. It is clear that the low earnings of Mexicans and Central Americans are not due to their lower probability of working or their marginal attachment to the labor market. This group of immigrants has a high employment rate, and after 10-20 years in the U.S., their employment rate consistently exceeds that of similarly aged natives (the graphs do not even correct for schooling). What is also true is that, in terms of the relative employment rate, the performance of recent cohorts has improved, with the last two cohorts surpassing natives within 10 years. This phenomenon is consistent with the idea that low skilled immigrants have taken a large number of manual and physically demanding jobs, whose job creation has been relatively strong in the recent decades (Basso et al. 2017). The flexible U.S. labor market has employed many immigrants, although at low wages. This situation is quite different from what has occurred with refugees in Europe, where employment rates have remained quite low due, in part, to the more generous support of governments but also because it is harder to gain access to labor markets (Fasani et al. 2018). While the U.S. labor market employs low-skilled immigrants at a high rate, the fact that they have a significant wage penalty, coupled with the fact that less educated natives' wages have also performed badly, imply that employment convergence, per se, is not sufficient to ensure the economic success of this group.

3.5.3 How Large is Return Migration?

The cohort method we adopt has been used as the main tool of analysis of immigrant assimilation, and the U.S. Census and ACS data have been the main source for this type of analysis. However, we need to emphasize two important caveats about these data. First, if

there is return migration to the source country, a cohort may change size and composition over time. If return migration is selective, part of the earnings convergence may be driven by poorly performing immigrants returning to their home country. Second, there may be some recall error in the arrival time, which would introduce measurement error in the size and composition of each cohort. Table 3.3 shows the population of each cohort used in our empirical analysis, which we can follow along the rows of the table. Notice that the cohorts we use for the labor market analysis include people 25 to 64 years old, not in group quarters, and includes all Mexicans and Central Americans who enter the U.S. at 18 or older. The change in size of the cohort in the first decade after arrival is always positive, and it is due to the people who arrived at age 18-24 and enter the considered age group. After that, notice that the cohort size shrinks, and this attrition is largely due to return migration and, to a lesser extent, to aging out of the group. However, given that the average age at arrival is rather young, the aging out is not significant until 3 or 4 decades after arrival. The reduction in cohort size 30 years after arrival can be substantial (comparing the number after 30 years with that after 10 years). Attrition seems differential across cohorts, and while we cannot do too much about it, it should be kept in mind as a possible source of selection of the remaining migrants.

Table 3.3 Population Estimates for Mexican and Central American Immigrant Cohorts

	Survey Year					
Cohort	1970	1980	1990	2000	2010	2017
1965-1969 arrivals	39,467	81,060	72,985	59,455	2,736	...
1975-1979 arrivals	...	147,640	240,400	267,721	149,135	38,555
1985-1989 arrivals	...	...	286,304	631,788	486,691	369,182
1995-1999 arrivals	...	...	...	640,099	768,334	653,910
2005-2011 arrivals	...	...	...	...	595,641	682,617
Natives	34,734,070	40,998,200	47,947,840	53,784,860	57,155,860	61,335,820

Note: These figures estimate the population of native-born and Mexican and Central American immigrant males between the ages of 25 and 64 who had between 1 and 40 years of potential work experience, were not in school or living in group quarters, and (for immigrants) entered the U.S. at the age of 18 or older.

3.6 The Role of Sector and Location

3.6.1 Convergence by Sector of Employment

It is difficult to produce causal evidence that identifies which economic conditions or policies promote rapid earnings convergence for Mexican and Central American immigrants. We can, however, identify some features of the labor market and location choices that are associated with different rates of earnings growth. In particular, by focusing on the sectors and locations where Mexicans and Central Americans are highly concentrated, we can determine whether the choice of sector or location is correlated with better outcomes relative to similar natives. Specifically, we analyze whether being located in an urban area or in a state with a large share of Mexicans and Central Americans (enclaves) is associated with an earnings advantage or disadvantage. Different sectors may provide different opportunities for upward mobility, and some specific urban locations are associated with more rapid earnings growth and stronger inter-generational mobility of natives (Chetty and Hendren 2018; Moretti 2013). As a result, certain locations may possibly generate benefits for immigrants, too.

Table 3.4 shows the proportion of Mexicans and Central Americans in four sectors of the economy, the proportion residing in urban and rural locations, and the proportion residing in enclave and non-enclave states. We define enclave states as the states with the largest percentage of Mexicans and Central Americans in the population calculated over the period 1970 to 2017. These states include California, Texas, Arizona, New Mexico, Nevada, and Illinois. In each of the sectors chosen, the immigrant group is over-represented relative to its average presence in the labor force. In particular, in the agriculture (respectively construction) sector, 23.8% (respectively 15.1%) of the labor force was Mexican or Central American in 2017. These statistics imply a very high degree of over-representation because Mexicans and Central Americans were only 5.4% of the overall labor force during that year. The other two sectors we consider, manufacturing

Table 3.4 Percent of Workforce Comprised of Mexican and Central American Immigrants by Sector and Location

	Survey Year					
	1970	1980	1990	2000	2010	2017
	Panel A: *By Sector*					
Agriculture and Farming	1.5	4.7	10.5	19.6	27.8	23.8
Construction	0.4	1.1	2.8	7.5	13.1	15.1
Manufacturing	0.8	2.7	4.4	8.5	9.7	7.5
Personal and Household Services	0.7	2.1	4.3	7.5	9.3	7.2
All Sectors	0.4	1.1	2.1	4.4	6.0	5.4
	Panel B: *By Location*					
Rural	0.2	0.4	0.6	1.8	2.8	2.5
Urban	0.5	1.5	2.8	5.2	6.9	6.1
Enclave	1.6	4.2	6.9	11.7	13.3	11.5
Non-Enclave	0.2	0.2	0.5	1.8	3.3	3.1

Note: These figures only include U.S.-born, Mexican, and Central American males between the age of 25 and 64 who had between 1 and 40 years of potential work experience, were not in school or living in group quarters, had positive earnings, worked at least one week during the survey year, and (for immigrants) entered the U.S. at the age of 18 or older. The enclave states used here are based on the share of Mexican and Central American immigrants calculated over the time period 1970-2017. They include California, Texas, Arizona, New Mexico, Nevada, and Illinois.

and personal and household services, include a larger than average share of Mexicans and Central Americans but not by much. The growth of the share of Mexicans and Central Americans in the workforce, especially in agriculture (respectively construction), has also been substantial, starting at 1.5% (respectively 0.4%) in 1970 and increasing to 23.8% (respectively 15.1%) in 2017. Panel B of the table reveals that Mexicans and Central Americans are more concentrated in urban areas and enclave states (by definition).

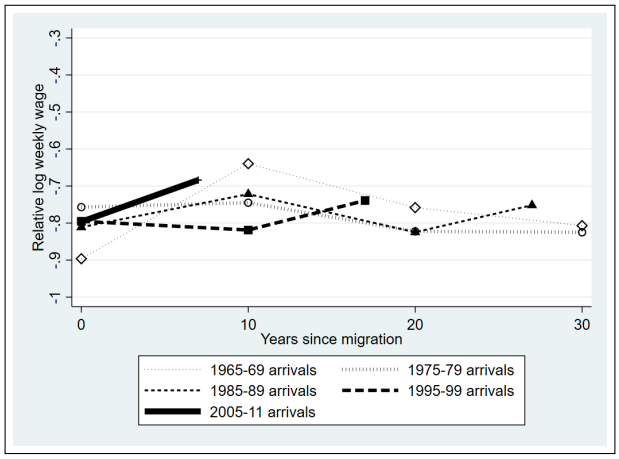
The earnings gap and convergence of Mexicans and Central Americans employed in different sectors relative to natives of the same age is shown in Figure 3.4, where each panel includes only Mexicans and Central Americans working in one specific sector and natives in all sectors. These graphs compare the average earnings of Mexicans and Central Americans in a sector to the average American of the same age. The sectors we consider represent the ones with the largest Mexican and Central American presence. Each panel of the figure shows the initial earnings gap and 30-year convergence for each cohort starting with the one that arrived between 1965 and 1969 and ending with the one that arrived

between 2005 and 2011.

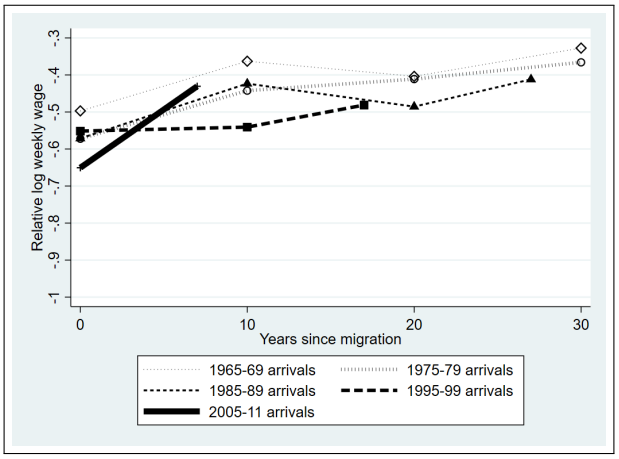
Our analysis reveals three main findings. First, within a sector, both the initial earnings gap and convergence rates for each cohort have not changed systematically over time. It is also important to note that this is a "within sector" convergence, so part of the aggregate convergence may result from Mexicans and Central Americans changing sector of work (say from agriculture to construction), which our analysis does not resolve. Second, agriculture is the sector with the largest initial earnings gap (about 80 log points, corresponding to a stunning 55 percent gap) and least rapid (almost insignificant) convergence. These results are not surprising because the agricultural sector has a negative wage differential with most other sectors, and agricultural workers do not experience much earnings growth over a career. Furthermore, to the extent that immigrant workers are more concentrated in the lower paying non-supervisory occupations, the lack of assimilation in the agricultural sector may be partially driven by earnings growth for native supervisors and farmers. Mexicans and Central Americans in the personal and household services sector do not perform much better than those in the agricultural sector. Third, Mexicans and Central Americans in the construction sector show a smaller initial earnings gap and a more rapid and persistent convergence rate over thirty years, much better any other sector. An initial gap of 60 log points is reduced to around 30 log points after 30 years. It is important to keep in mind that these are gaps relative to the average U.S. native of a similar age. If we compare Mexicans and Central Americans to similarly aged and educated natives, as we do in Table 3.5, the results become even more striking.

Fig. 3.4 Age-Adjusted Convergence for the Relative Weekly Earnings of Mexican and Central American Cohorts by Sector: Initial Gap and Convergence

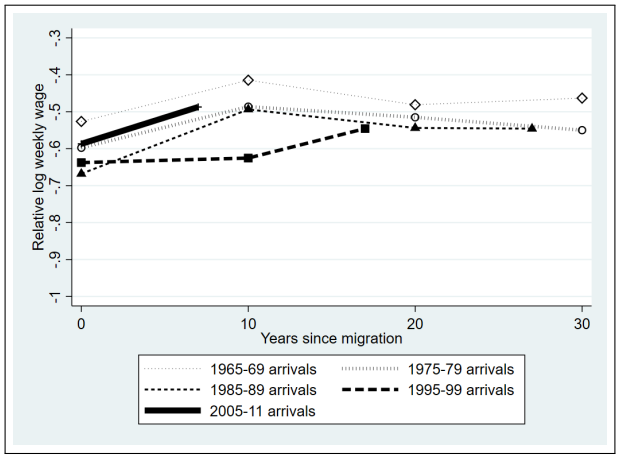
A: Agriculture and Farming



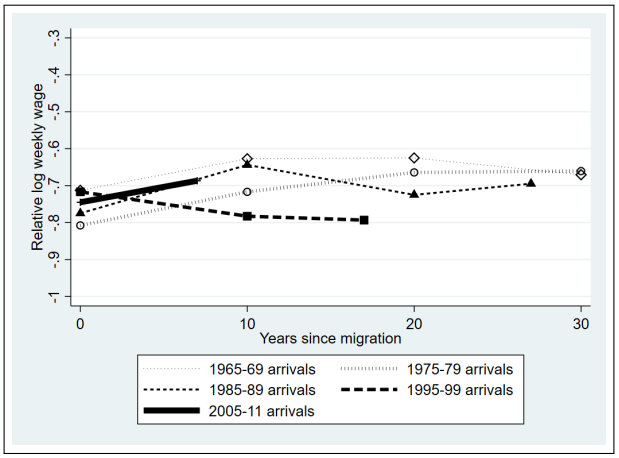
B: Construction



C: Manufacturing



D: Personal and Household Services



Note: The wage differentials presented this figure are calculated from regressions that are estimated separately for each cross section using data that only includes individuals employed in the sector identified in the panel being considered. The dependent variable in these regressions identifies the log weekly earnings of each individual, and the explanatory variables include a third-order polynomial for age and a set of fixed effects: one for each immigrant cohort. The omitted group is comprised of native-born workers such that the coefficients on the cohort fixed effects represent each cohort's log weekly earnings relative to native-born workers in a given survey year, which are used to construct the data points in the graphs.

Table 3.5 Age-Education-Adjusted Relative Weekly Earnings of Mexicans and Central American Immigrants by Sector: Initial Gap and Convergence After First 10 Years

	Agriculture		Construction		Manufacturing		Personal and Household Services	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A				
				Relative Entry Wage				
1965-1969 arrivals	-0.753**	-0.606**	-0.449**	-0.339**	-0.478**	-0.409**	-0.654**	-0.570**
	(0.033)	(0.020)	(0.047)	(0.036)	(0.032)	(0.025)	(0.037)	(0.025)
1975-1979 arrivals	-0.733**	-0.500**	-0.534**	-0.325**	-0.592**	-0.393**	-0.816**	-0.612**
	(0.034)	(0.021)	(0.041)	(0.034)	(0.031)	(0.024)	(0.036)	(0.025)
1985-1989 arrivals	-0.803**	-0.490**	-0.581**	-0.337**	-0.686**	-0.448**	-0.779**	-0.536**
	(0.036)	(0.022)	(0.041)	(0.031)	(0.038)	(0.030)	(0.036)	(0.024)
1995-1999 arrivals	-0.833**	-0.489**	-0.605**	-0.317**	-0.702**	-0.423**	-0.746**	-0.481**
	(0.020)	(0.013)	(0.020)	(0.017)	(0.017)	(0.013)	(0.018)	(0.012)
2005-2011 arrivals	-0.872**	-0.461**	-0.746**	-0.368**	-0.641**	-0.289**	-0.817**	-0.459**
	(0.014)	(0.013)	(0.014)	(0.017)	(0.015)	(0.011)	(0.013)	(0.012)
2012-17 arrivals	-0.860**	-0.424**	-0.562**	-0.149**	-0.506**	-0.209**	-0.609**	-0.262**
	(0.002)	(0.008)	(0.003)	(0.013)	(0.002)	(0.003)	(0.003)	(0.007)
				Panel B				
				Relative Wage Growth in First 10 Years				
1965-1969 arrivals	0.010	0.147**	0.079	0.207**	0.072	0.243**	0.041	0.184**
	[0.854]	[0.000]	[0.161]	[0.000]	[0.162]	[0.000]	[0.457]	[0.000]
1975-1979 arrivals	-0.026	0.114**	0.082	0.209**	0.086†	0.237**	0.094†	0.224**
	[0.618]	[0.002]	[0.122]	[0.000]	[0.093]	[0.000]	[0.093]	[0.000]
1985-1989 arrivals	0.022	0.117**	0.105†	0.197**	0.134*	0.253**	0.076	0.169**
	[0.679]	[0.003]	[0.060]	[0.000]	[0.021]	[0.000]	[0.166]	[0.000]
1995-1999 arrivals	0.046	0.129**	0.094*	0.176**	0.118**	0.227**	0.006	0.099**
	[0.245]	[0.000]	[0.016]	[0.000]	[0.003]	[0.000]	[0.862]	[0.000]
2005-2011 arrivals	0.161**	0.201**	0.261**	0.313**	0.107**	0.165**	0.099**	0.120**
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
N	9,425,202	9,425,202	9,423,810	9,423,810	9,423,649	9,423,649	9,426,230	9,426,230
Basic Specification	X	–	X	–	X	–	X	–
Age-Educ-Year FE	–	X	–	X	–	X	–	X

Note: The wage differentials presented in Panel A are generated from regressions that are ran on the set of pooled cross sections from 1970, 1980, 1990, 2000, 2010, and 2017. The dependent variable identifies the log weekly earnings of each individual. The explanatory variables in columns (1), (3), (5), and (7) include a third order polynomial for age, a third order polynomial for the number of years that immigrants have spent in the U.S., a set of cohort fixed effects, and a set of cohort fixed effects that are each interacted with a continuous variable identifying the number of years that immigrants have spent in the U.S. The explanatory variables for columns (2), (4), (6), and (8) contain the same set of variables as in column (1) but additionally include a set of education-age-year fixed effects. The omitted group is comprised of native-born workers such that the coefficients in Panel A each represent a separate cohort's log weekly earnings relative to native born workers. The predicted relative wage growth in the first 10 years in Panel B assumes that all immigrants arrive in the country at the age of 25. Standard errors are in parentheses. P-values are in brackets. † $p < .1$, * $p < .05$, ** $p < .01$

Table 3.5 shows a comparison of the initial log earnings gap (Panel A) and relative earnings growth (Panel B), by sector, when comparing Mexicans and Central Americans to similarly aged natives (columns 1,3,5 and 7) and when adjusting for education (columns 2,4,6 and 8). The results in Table 3.5 confirm the advantage of Mexicans and Central Americans in the construction sector (lagging only 32-33 log points at arrival and catching up by 20 log points in the first decade) and their disadvantage in the agricultural sector (lagging 50 log points at arrival and catching up only by 11-12 log points in the first decade). After controlling for education, Mexicans and Central Americans in manufacturing also seem to perform relatively well, with a high relative growth rate in the first decade (20-25 log points). The performance of Mexican and Central American agricultural workers is significantly improved when they are compared to similarly educated natives, which is a sign that the negative selection of workers in agriculture and the slow earnings growth for the low educated, which is true for natives too, explains a large part of the under-performance of workers in that sector. Still, the agricultural sector is the one with largest initial earnings gap and the slowest convergence even relative to similarly aged and educated natives.

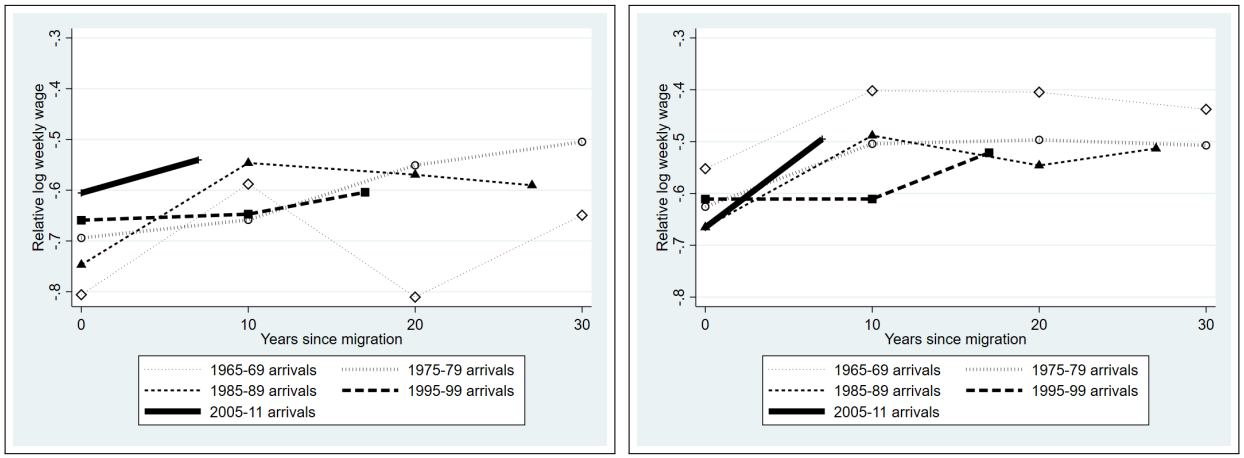
3.6.2 Convergence in Rural and Urban Areas

Despite their heavy presence in agricultural jobs, the concentration of Mexicans and Central Americans is larger in urban areas because most jobs are non-agricultural. Therefore, it is useful to see if urban location is associated with better wage performance relative to natives. Figure 3.5 shows the initial earnings gap and convergence relative to similarly aged natives for urban and rural areas, separately. The figure reveals that the initial earnings gap is smaller for those living in urban areas, but the convergence does not seem significantly different between the two graphs in the figure. Except for the 1965-69 rural cohort, which was small and shows a rather noisy estimate of convergence, the other cohorts in both rural and urban areas seem to perform similarly over time in terms of relative earnings growth.

Table 3.6 shows the initial earnings gap and convergence for rural and urban Mexicans and Central Americans when we compare them to similarly aged and then to similarly aged and educated natives. The results confirm a smaller initial earnings gap for those living in urban areas but a similar rate of earnings growth. Urban location may provide some initial advantage, but it is not so clear that it produces a sustained advantage over time. It would be interesting to separate urban locations between fast growing and declining ones as the wage dynamics may be very different among them (as noted by Moretti 2013) to see if the "divergence" between those two types of urban areas is also reflected in the economic convergence of Mexican and Central American immigrants.

Fig. 3.5 Age-Adjusted Convergence for the Relative Weekly Earnings of Mexican and Central American Cohorts by Location: Initial Gap and Convergence

A: Rural **B:** Urban



Note: The wage differentials presented this figure are calculated from regressions that are estimated separately for each cross section using data that only includes individuals employed in the region identified in the panel being considered. The dependent variable in these regressions identifies the log weekly earnings of each individual, and the explanatory variables include a third-order polynomial for age and a set of fixed effects: one for each immigrant cohort. The omitted group is comprised of native-born workers such that the coefficients on the cohort fixed effects represent each cohort's log weekly earnings relative to native-born workers in a given survey year, which are used to construct the data points in the graphs.

Table 3.6 Age-Education-Adjusted Relative Weekly Earnings of Mexicans and Central American Immigrants by Location: Initial Gap and Convergence After First 10 Years

	Rural		Urban		Enclave		Non-Enclave	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A				
				Relative Entry Wage				
1965-1969 arrivals	-0.720**	-0.625**	-0.481**	-0.413**	-0.536**	-0.448**	-0.494**	-0.452**
	(0.0433)	(0.0298)	(0.0449)	(0.0291)	(0.054)	(0.035)	(0.024)	(0.017)
1975-1979 arrivals	-0.722**	-0.521**	-0.612**	-0.425**	-0.650**	-0.450**	-0.480**	-0.362**
	(0.0451)	(0.0306)	(0.0422)	(0.0289)	(0.043)	(0.029)	(0.040)	(0.028)
1985-1989 arrivals	-0.714**	-0.447**	-0.663**	-0.441**	-0.698**	-0.461**	-0.581**	-0.379**
	(0.0514)	(0.0357)	(0.0466)	(0.0344)	(0.048)	(0.035)	(0.041)	(0.029)
1995-1999 arrivals	-0.704**	-0.406**	-0.668**	-0.419**	-0.705**	-0.445**	-0.629**	-0.369**
	(0.0229)	(0.0169)	(0.0228)	(0.0249)	(0.025)	(0.023)	(0.017)	(0.015)
2005-2011 arrivals	-0.647**	-0.320**	-0.743**	-0.434**	-0.751**	-0.434**	-0.711**	-0.393**
	(0.0152)	(0.0124)	(0.0163)	(0.0255)	(0.017)	(0.019)	(0.012)	(0.019)
2012-17 arrivals	-0.512**	-0.147**	-0.531**	-0.243**	-0.570**	-0.270**	-0.484**	-0.176**
	(0.00284)	(0.00513)	(0.00363)	(0.0236)	(0.003)	(0.015)	(0.002)	(0.018)
N	9,112,492	9,112,492	9,331,676	9,331,676 *N*	9,588,212	9,588,212	9,465,090	9,465,090
				Panel B				
				Relative Wage Growth in First 10 Years				
1965-1969 arrivals	0.047	0.207**	0.073	0.213**	0.086	0.226**	0.096†	0.208**
	[0.476]	[0.000]	[0.257]	[0.000]	[0.224]	[0.000]	[0.057]	[0.000]
1975-1979 arrivals	0.093	0.232**	0.084	0.212**	0.103	0.228**	0.024	0.167**
	[0.188]	[0.000]	[0.183]	[0.000]	[0.124]	[0.000]	[0.653]	[0.000]
1985-1989 arrivals	0.097	0.193**	0.107	0.197**	0.126†	0.211**	0.056	0.161**
	[0.184]	[0.001]	[0.110]	[0.000]	[0.077]	[0.000]	[0.304]	[0.001]
1995-1999 arrivals	0.079†	0.171**	0.098*	0.180**	0.119*	0.202**	0.062†	0.147**
	[0.092]	[0.000]	[0.034]	[0.000]	[0.019]	[0.000]	[0.075]	[0.000]
2005-2011 arrivals	0.067**	0.170**	0.204**	0.249**	0.195**	0.240**	0.182**	0.242**
	[0.001]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Basic Specification	X	–	X	–	X	–	X	–
Age-Educ-Year FE	–	X	–	X	–	X	–	X

Note: The wage differentials presented in Panel A are generated from regressions that are ran on the set of pooled cross sections from 1970, 1980, 1990, 2000, 2010, and 2017. The dependent variable identifies the log weekly earnings of each individual. The explanatory variables in columns (1), (3), (5), and (7) include a third order polynomial for age, a third order polynomial for the number of years that immigrants have spent in the U.S., a set of cohort fixed effects, and a set of cohort fixed effects that are each interacted with a continuous variable identifying the number of years that immigrants have spent in the U.S. The explanatory variables for columns (2), (4), (6), and (8) contain the same set of variables as in column (1) but additionally include a set of education-age-year fixed effects. The omitted group is comprised of native-born workers such that the coefficients in Panel A each represent a separate cohort's log weekly earnings relative to native born workers. The predicted relative wage growth in the first 10 years in Panel B assumes that all immigrants arrive in the country at the age of 25. The enclave states are the states with the largest percentage of Mexican and Central Americans in the population over the period 1970-2017. They include California, Texas, Arizona, New Mexico, Nevada, and Illinois. Standard errors are in parentheses. P-values are in brackets. † $p < .1$, * $p < .05$, ** $p < .01$

3.6.3 Convergence in Enclave vs. Non-Enclave States

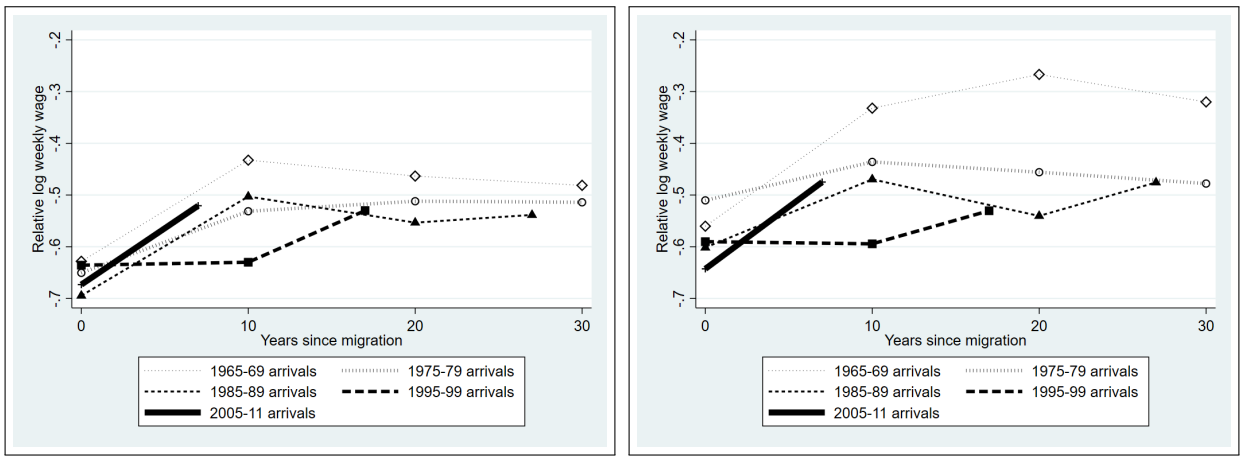
Some studies identify the local crowding of immigrants as a reason for slower integration (e.g., Borjas 2015). If immigrants live in an enclave that has a large share of co-ethnics, they may be less inclined to learn English, socially integrate, or they may remain marginal to some job and career opportunities. However, other studies use a more careful causal identification strategy and find that living in an enclave significantly increases earnings

(Piil Damm 2009). In recent work on German refugees, Battisti et al. (2016) find that living in enclaves may provide an initial employment advantage for new immigrants, but it may reduce their investment in human capital and diminish their earnings potential in the long run.

In order to test whether there is some association between living in an enclave and the initial earnings gap and convergence, we separate Mexicans and Central Americans living in the 6 largest enclave states from those not living in those states. This classification of states provides a rough categorization as one would like to check enclaves in smaller geographical units, such as counties or metropolitan areas. However, this state-level categorization will provide some preliminary evidence. As usual, we show the representation of convergence to similarly aged natives in Figure 3.6, and we show the initial earnings gap and relative earnings growth in the first ten years for similarly aged and for similarly aged and educated natives in Table 3.6.

Fig. 3.6 Age-Adjusted Convergence for the Relative Weekly Earnings of Mexican and Central American Cohorts by Enclave Region: Initial Gap and Convergence

A: Enclave **B:** Non-Enclave



Note: The wage differentials presented this figure are calculated from regressions that are estimated separately for each cross section using data that only includes individuals employed in the region identified in the panel being considered. The dependent variable in these regressions identifies the log weekly earnings of each individual, and the explanatory variables include a third-order polynomial for age and a set of fixed effects: one for each immigrant cohort. The omitted group is comprised of native-born workers such that the coefficients on the cohort fixed effects represent each cohort's log weekly earnings relative to native-born workers in a given survey year, which are used to construct the data points in the graphs. The enclave states are the states with the largest percentage of Mexican and Central Americans in the population over the period 1970-2017. They include California, Texas, Arizona, New Mexico, Nevada, and Illinois.

Both Figure 3.6 and the education adjusted coefficients in Table 3.6 do not show a large difference in the initial earnings gap between immigrants located in enclave or non-enclave states. Possibly, a worsening of the initial gap is visible in enclave states, which may reveal some crowding of Mexicans in some jobs, especially in the more recent decades. Several studies emphasize how the strongest labor market competition for new immigrants comes from other immigrants, and this finding may be partially consistent with that observation.

Summarizing this analysis of Mexicans and Central Americans in the U.S. over the last five decades we highlight four main findings: First, recent arrival cohorts did not do worse than previous ones in terms of the initial earnings gap or in terms of relative earnings growth. Second, there is a significant earnings gap relative to similar natives (on the order of 40 log points at arrival), which is reduced by about one third to one half but is not eliminated within 20-30 years. Third, there is a small initial employment gap, but after 20 years in the country, Mexicans and Central Americans are employed at a rate higher than similarly aged natives. Last, Mexicans and Central Americans in the construction sector and in urban areas do better than others in terms of the initial earnings gap, and those in the construction and manufacturing sectors do better in terms of relative earnings growth during the first 10 years.

The picture revealed by this analysis is one of a group coming to the U.S. to work in manual/physical intensive jobs and assimilating rapidly, in terms of being employed, but lagging behind in terms of earnings. Jobs in a sector like construction, which has a significant upward potential and usually is located in urban areas may be well suited to boost the economic success of immigrants. This finding could be an important consideration when discussing the potential for new job-related visas for less educated immigrants in terms of how to distribute them across occupations and sectors.

3.6.4 Composition and Language Skills

The location and sector of employment may be important factors in improving the initial earnings gap and convergence. Here we document potential factors affecting the performance of the two most recent cohorts analyzed in our study. Were those cohorts better positioned in terms of schooling or knowledge of English upon arrival? Were there differences in the composition of the two groups? Could Central Americans be at a further disadvantage coming from poorer countries relative to Mexicans? Table 3.7 shows some characteristics of each cohort upon entry between 1965 and 2011, revealing evidence of potential trends, which may affect skills and earnings differentials even after controlling for age and education.

Table 3.7 Summary Statistics for
Mexican and Central American Immigrant Cohorts Upon Arrival

Cohort	1965-69	1975-79	1985-89	1995-99	2005-11
Age	34.24	33.11	32.85	32.98	33.73
Years of Schooling	6.79	6.79	7.29	7.79	8.29
Share Central American	0.18	0.14	0.26	0.15	0.29
Share Speaking Some English	No Data	0.66	0.69	0.65	0.67
Share Speaking Good English	No Data	0.28	0.31	0.29	0.27
Share Nonwhite	0.06	0.06	0.56	0.55	0.42

First of all, in terms of education and age it appears that, upon arrival to the U.S., more recent cohorts are slightly better educated but are about the same age as the cohort that arrived in the 70's. These changes are small and controlled for in the convergence equation. The share of Central Americans, while varying by cohort, does not seem to have a clear trend nor does the share of those speaking English (at all or proficiently) at arrival. One variable that has increased substantially since the 1975-79 cohort arrived is the share of nonwhite, but the changed nature of the census question, which allowed people to indicate more than one ethnicity after 1980, may have affected the numbers. The share of nonwhite appears to decrease in the most recent cohort (2005-11) relative to the one arrived between 1995 and 1999. Overall, the most recent cohort, whose performance seems better

than the previous ones, does not seem much different in terms of language, but it includes a larger share of Central Americans and has a slightly higher level of education. Overall, these variables do not suggest that the skill composition of new arrivals has changed much over the past 20 years, yet more recent cohorts have performed better relative to similar natives.

3.7 The Other Largest Groups of Immigrants: Chinese and Indians

One important and novel finding of this paper is that the more recent cohorts of Mexican and Central American immigrants performed better than earlier ones in terms of earnings and employment rate gaps. Certainly, they have been migrating into an economy where the wages of less educated Americans have been deteriorating relative to the wages of high skilled Americans. This wage evolution has hurt immigrants in absolute terms, but it has not penalized them more than natives. This finding is interesting because several studies have pointed to a deterioration in the economic assimilation of immigrants that arrived in the 80's and early 90's (Borjas 2015). However, we show that a comparison of a more homogeneous group of immigrants to similar natives contradicts previous findings.

Do these findings also hold true for other large groups of immigrants? Are more recent cohorts of immigrants from other countries doing better than previous cohorts from the same countries? To answer these questions, we consider Chinese and Indian immigrants, separately, the two largest groups after Mexicans and Central Americans. Their immigration flows have become larger than that of Mexicans and Central Americans over the last decade. The Chinese and Indians have had a much larger share of highly educated individuals migrating to the U.S., both relative to Mexicans and Central Americans and relative to the U.S. population. Table 3.8 shows the share of people with a high school education or less, with some college, and with a college degree for the three groups of immigrants (Mexicans and Central Americans, Indians, and Chinese).

Table 3.8 Percent of Immigrants with High-School and College Education

	Survey Year					
	1970	1980	1990	2000	2010	2017
	Panel A					
	With a High School Diploma or Less					
Mexicans and Central Americans	89.4	89.9	87.3	86.4	84.5	81.6
Chinese	49.4	40.6	36.5	31.8	32.1	28.0
Indians	10.7	12.3	17.5	16.9	13.1	12.6
Natives	69.2	55.2	44.2	39.1	35.4	33.2
	Panel B					
	With At Least Some College					
Mexicans and Central Americans	10.6	10.1	12.7	13.6	15.5	18.4
Chinese	50.6	59.4	63.5	68.2	67.9	72.0
Indians	89.3	87.7	82.5	83.1	86.9	87.4
Natives	30.8	44.8	55.8	60.9	64.6	66.8
	Panel C					
	With a Bachelor's Degree or Higher					
Mexicans and Central Americans	4.3	3.8	4.2	4.6	5.6	7.2
Chinese	41.4	49.6	50.2	58.0	59.2	62.9
Indians	83.5	78.6	72.0	73.5	78.7	81.0
Natives	17.0	24.8	27.9	30.9	34.2	36.6

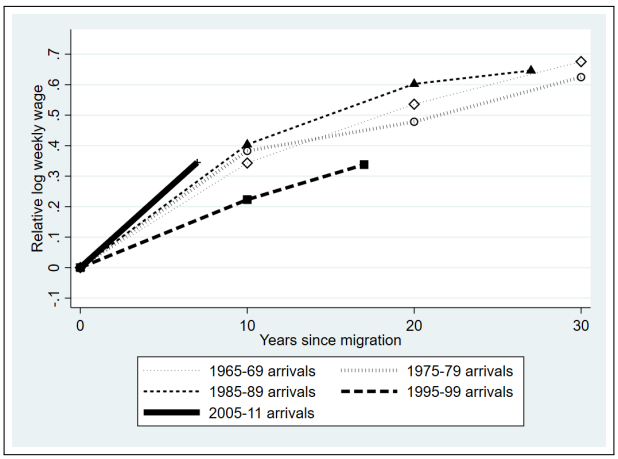
Note: These figures only include males between the age of 25 and 64 who had between 1 and 40 years of potential work experience, were not in school or living in group quarters, had positive earnings, worked at least one week during the survey year, and (for immigrants) entered the U.S. at the age of 18 or older.

It appears clearly from the table that Chinese and, even more so, Indian immigrants have been selected among highly educated individuals. These groups of immigrants have consistently had high levels of education since 1970. This very strong selection makes them more educated than natives on average. Of course Chinese and Indian immigrants have largely been employed in different sectors than Mexicans and Central Americans, with a large concentration in high tech, engineering, science, and professional occupations. Still, it is very interesting to see how subsequent cohorts of these immigrants compare to similar natives. Figures 3.7.B and 3.7.D show the initial gap and convergence of earnings and employment rates, respectively, for Chinese immigrants relative to similarly aged natives. In terms of earnings and employment, this group of immigrants consistently enters the U.S. with a relatively small gap and outperforms natives within 20 years of stay

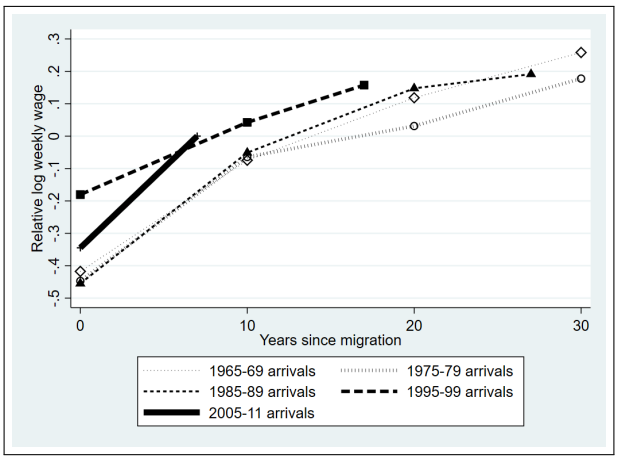
Fig. 3.7 Age-Adjusted Convergence for the Relative Weekly Earnings and Employment Rate of Chinese Cohorts

Earnings

A: Normalized Convergence

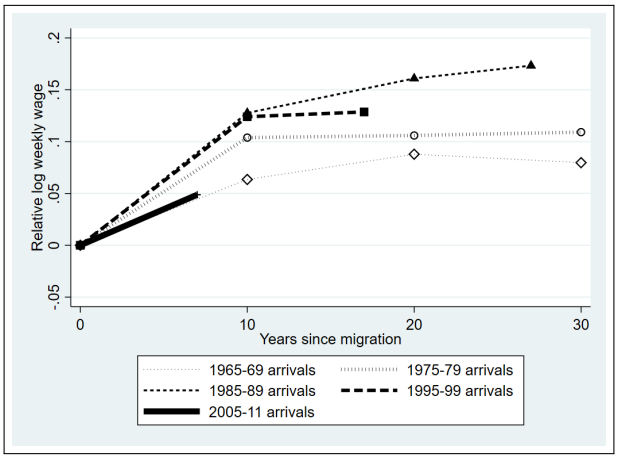


B: Initial Gap and Convergence

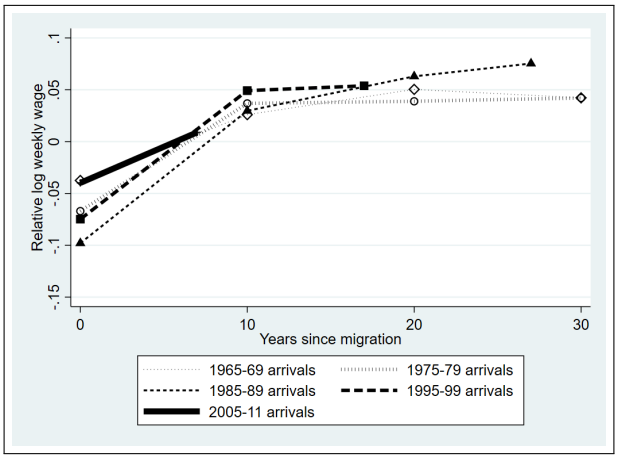


Employment

C: Normalized Convergence



D: Initial Gap and Convergence

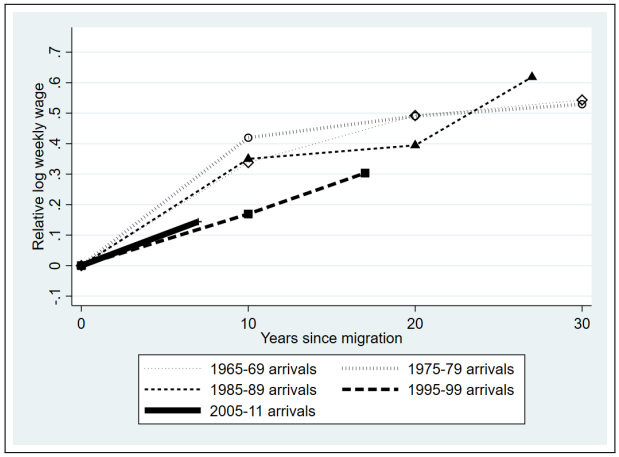


Note: The log weekly earnings and employment rate differentials presented this figure are calculated from regressions that are estimated separately for each cross section. The dependent variable in the earnings regressions identifies the log weekly earnings of each individual. The dependent variable in the employment regressions identifies whether each individual was employed for at least one week during the previous year. The explanatory variables for both the earnings and employment regressions include a third-order polynomial for age and a set of fixed effects: one for each immigrant cohort. The omitted group is comprised of native-born workers such that the coefficients on the cohort fixed effects represent each cohort's log weekly earnings or employment rate relative to native-born workers in a given survey year, which are used to construct the data points in the graphs. In Figure 3.7A, the relative log weekly earnings is normalized to zero at the time of entry. In Figure 3.7A, the relative employment rate for each cohort is normalized to zero at the time of the entry.

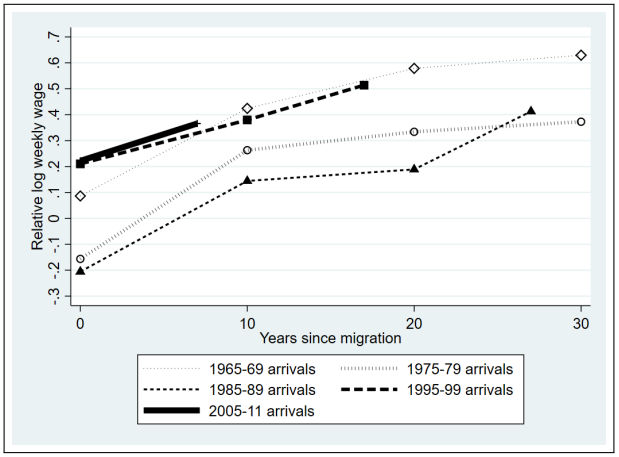
Fig. 3.8 Age-Adjusted Convergence for the Relative Weekly Earnings and Employment Rate of Indian Cohorts

Earnings

A: Normalized Convergence

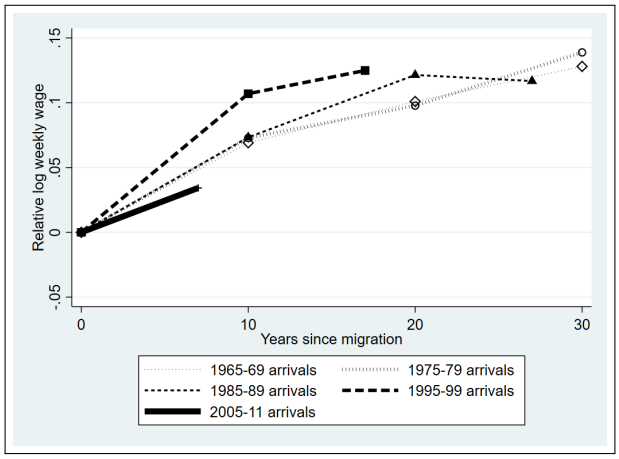


B: Initial Gap and Convergence

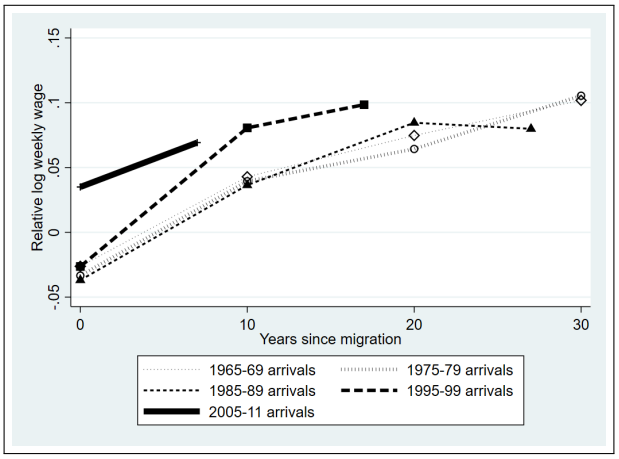


Employment

C: Normalized Convergence



D: Initial Gap and Convergence



Note: The log weekly earnings and employment rate differentials presented this figure are calculated from regressions that are estimated separately for each cross section. The dependent variable in the earnings regressions identifies the log weekly earnings of each individual. The dependent variable in the employment regressions identifies whether each individual was employed for at least one week during the previous year. The explanatory variables for both the earnings and employment regressions include a third-order polynomial for age and a set of fixed effects: one for each immigrant cohort. The omitted group is comprised of native-born workers such that the coefficients on the cohort fixed effects represent each cohort's log weekly earnings or employment rate relative to native-born workers in a given survey year, which are used to construct the data points in the graphs. In Figure 3.8A, the relative log weekly earnings is normalized to zero at the time of entry. In Figure 3.8C, the relative employment rate for each cohort is normalized to zero at the time of the entry.

in the U.S. The graphs also show that the performance of the two most recent cohorts are the best overall.

Figure 3.8 shows the same graphs for Indians, and in this case we see an even smaller initial gap and more rapid convergence. Indian immigrants consistently outperform natives in terms of earnings and employment rates within 10 years of arrival. For this group of immigrants, the two most recent cohorts already start at par or with an advantage in earnings and probability of working and continue to improve their relative performance over time.

The analyses of Chinese and Indian immigrants confirm that the labor market performance of immigrants who arrived in the last two decades has been remarkably strong. This evidence suggests that: (i) the quality of recent immigrants, in terms of labor market skills, is not worse than that of previous immigrants when we analyze specific countries of origin and (ii) the decline in the inflow of new immigrants in recent years may have stimulated rapid convergence. Moreover, the extremely high employment rates, when compared to similar natives, confirm that immigrants come to the U.S. to work and that the U.S. labor market continues to demand these workers. The exceptional relative performance of the Indians who arrived during or after 1995, many of whom entered the U.S. on an H1-B visa, suggests that the stories of underpaid H1-B visa workers parked in jobs with little upward mobility may represent the experience of some recent arrivals, but it is not representative of the whole group, whose salary and employment prospects are better than those of natives within 10 years of arrival to the U.S.

3.8 Conclusion

The economic assimilation of low skilled immigrants is a very important issue often dominating the debate about immigration. Several receiving countries claim that immigrants are, and remain, a burden to the receiving country because they do not have skills that can be integrated in the labor market, hence their employment rate is low and their earnings

lag behind those of similar natives. In the U.S., there is anecdotal and empirical evidence showing that immigrants who arrived in the 1980's and 1990's have had a harder time assimilating into the labor market.

In this paper we analyze whether such a characterization is true when extending the analysis to cohorts of entry in the 1990's and 2000's while focusing on Mexican and Central American immigrants, traditionally a group of low educated immigrants earning low wages. This is also a very large group of immigrants, comprising almost 6% of the U.S. labor force, hence their success is very important to the U.S. economy and society as a whole. While we do find an initial earnings gap and only incomplete convergence after 30 years of stay, we also find that recent cohorts of Mexicans and Central Americans (i.e., those that arrived during or after 1995) have not performed worse than the earlier ones that arrived in the 70's and 80's. Moreover, we find that, in terms of employment probability, Mexicans and Central Americans outperform similarly aged natives after 20-30 years in the country. In particular, when focusing on the cohorts that arrived between 1995 and 1999 and between 2005 and 2011, they seem to perform particularly well.

The findings from this analysis suggest that the appearance of a worsening quality of recent cohorts arises from the grouping of all immigrants together. Once we focus on immigrants from certain countries of origin and compare them to similar natives, we find that recent cohorts have performed rather well relative to earlier ones.

Finally, we also show that those employed in the construction sector and those living in urban areas start out with a smaller earnings gap, and those in the construction and manufacturing sectors have stronger relative earnings growth. On the other hand, those employed in the agricultural sector start out with larger initial gap and have the weakest convergence.

Our analysis suggests that there is no basis to claim that new immigrants are of lower labor-market quality relative to earlier ones. A consideration of immigrants from specific countries of origin reveals that subsequent cohorts have actually performed similarly or better in the U.S. Moreover, despite many hurdles, the U.S. labor market has done an

exceptional job of offering employment opportunities to immigrants, at least up until 2017 (the last year of our analysis). However, the poor earnings performance of low skilled workers, in general, has had a disproportionate impact on Mexicans and Central Americans, who are heavily represented in this group. Given the high demand for labor in the construction sector and the opportunities that it affords immigrants for upward mobility, one could think of a special visas linked to jobs in that sector.

Bibliography

Abramitsky, R., Boustan, L. P., and Eriksson, K. (2014). A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration. *Journal of Political Economy*, 122(3):467–717.

Alesina, A., Miano, A., and Stantcheva, S. (2018). Immigration and Redistribution. *National Bureau of Economic Research (NBER)*. Working Paper No. 24733.

Altonji, J. G. and Card, D. (1991). The Effects of Immigration on the Labor Market Outcomes of Less-Skilled Natives. In Abowd, J. M. and Freeman, R. B., editors, *Immigration, Trade, and the Labor Market*, pages 201–234. University of Chicago Press.

Basso, G. and Peri, G. (2015a). The Association Between Immigration and Labor Market Outcomes in the United States. *IZA Discussion Paper*, (9436).

Basso, G. and Peri, G. (2015b). The Association Between Immigration and Labor Market Outcomes in the United States. IZA Discussion Paper No. 9436, Bonn, Germany.

Basso, G., Peri, G., and Rahman, A. (2017). Computerization and Immigration: Theory and Evidence from the United States. *National Bureau of Economic Research (NBER)*. Working Paper No. 23935.

Battisti, M., Felbermayr, G., Peri, G., and Poutvaara, P. (2018). Immigration, Search and Redistribution: A Quantitative Assessment of Native Welfare. *Journal of the European Economic Association*, 16(4):1137–1188.

Battisti, M., Peri, G., and Romiti, A. (2016). Dynamic Effects of Co-Ethnic Networks on Immigrants' Economic Success. *National Bureau of Economic Research (NBER)*. Working Paper No. 22389.

BLS (2018). Quarterly Census of Employment and Wages, Average Annual and Quarterly Employment. Retrieved November 15th, 2018 from California Employment Development Department's QCEW Data Search Tool:`https://www.labormarketinfo.edd.ca.gov/cgi/dataanalysis/areaselection.asp?tablename=industry`.

Borjas, G. (1985). Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants. *Journal of Labor Economics*, 3(4):463–89.

Borjas, G. J. (2001). Does Immigration Grease the Wheels of the Labor Market? *Brookings Papers on Economic Activity*, 2001(1):69–119.

Borjas, G. J. (2003). The Labor Demand Curve *is* Downward Sloping: Reexamining the Impact of Immigration on the Labor Market. *The Quarterly Journal of Economics*, 118(4):1335–1374.

Borjas, G. J. (2014). *Immigration Economics.* Harvard University Press.

Borjas, G. J. (2015). The Slowdown in the Economic Assimilation of Immigrants: Aging and Cohort Effects Revisited Again. *Journal of Human Capital*, 9(4):483–517.

Borjas, G. J., Freeman, R. B., and Katz, L. F. (1997). How Much Do Immigration and Trade Affect Labor Market Outcomes? *Brookings Papers on Economic Activity*, 1997(1):1–90.

Boucher, S. R., Smith, A., and Taylor, J. E. (2007). Impacts of Policy Reforms on the Supply of Mexican Labor to U.S. Farms: New Evidence from Mexico. *American Journal of Agricultural Economics*, 72(3):567–573.

Brady, M. P., Gallardo, R. K., Badruddozza, S., and Jiang, X. (2016). Regional Equilibrium Wage Rate for Hired Farm Workers in the Tree Fruit Industry. *Western Economics Forum*, Fall 2016.

Bratsberg, B. and Raaum, O. (2012). Immigration and Wages: Evidence from Construction. *The Economic Journal*, 122(565):1177–1205.

Bratsberg, B., Raaum, O., and Røed, K. (2017). Immigrant Labor Market Integration Across Admission Classes. *Centre for Research and Analysis of Migration (CReAM)*, Discussion Paper Series 1702.

Brooks, N. R. (1991). Agriculture: Despite Heavy Damages from Drought and Cold, 1990 State Farm Revenues are Expected to Show an Increase. *Los Angeles Times*, retrieved March 30, 2020 at: `https://www.latimes.com/archives/la-xpm-1991-01-06-fi-10871-story.html`.

Cadena, B. C. and Kovak, B. K. (2016). Immigrants Equilibrate Local Labor Markets: Evidence from the Great Recession. *American Economic Journal: Applied Economics*, 8(1):257–90.

California Agricultural Commissioners and Sealers Association (2020). About. [website]. Retrieved from: `https://cacasa.org/about/#:~:text=County%20Agricultural%20Commissioner,-During%20California's%20period&text=State%20law%20now%20requires%20every,by%20the%20Board%20of%20Supervisors.`

Cameron, A. C. and Miller, D. L. (2015). A Practitioner's Guide to Cluster-Robust Inference. *Journal of Human Resources*, 50(2):317–372.

Card, D. (1990). The Impact of the Mariel Boatlift on the Miami Labor Market. *ILR Review*, 43(2):245–257.

Card, D. (2001). Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration. *Journal of Labor Economics*, 19(1):22–64.

Card, D. (2009). Immigration and Inequality. *American Economic Review: Papers & Proceedings*, 99(2):1–21.

Card, D. and Peri, G. (2016). Immigration Economics by George J. Borjas: A Review Essay. *Journal of Economic Literature*, 54(4):1333–49.

Cassey, A. J., Lee, K., Sage, J., and Tozer, P. R. (2018). Assessing Post Harvest Labor Shortages, Wages, and Welfare. *Agricultural and Food Economics*, 6(17).

CDFA (2018). California Agricultural Production Statistics. Retrieved February 10th, 2019 from: https://www.cdfa.ca.gov/statistics/.

CFBF and UC Davis (2019). Still Searching for Solutions: Adapting to Farm Worker Scarcity Survey 2019. *CFBF*.

Charlton, D. and Taylor, J. E. (2016). A Declining Farm Workforce: Analysis of Panel Data from Rural Mexico. *American Journal of Agriculutral Economics*, 98(4):1158–1180.

Charlton, D., Taylor, J. E., Vougioukas, S., and Rutledge, Z. (2019a). Can Wages Rise Quickly Enough to Keep Workers in the Fields? *Choices Magazine (forthcoming)*.

Charlton, D., Taylor, J. E., Vougioukas, S., and Rutledge, Z. (2019b). Innovations for a Shrinking Agricultural Workforce. *Choices Magazine (forthcoming)*.

Chetty, R. and Hendren, N. (2018). The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates. *The Quarterly Journal of Economics*, 133(3):1163–1228.

Chiswick, B. (1978). The Effect of Americanization on the Earnings of Foreign-born Men. *Journal of Political Economy*, 86(5):897–921.

Clark, K. and Lindley, J. (2006). Immigrant Labour Market Assimilation and Arrival Effects: Evidence from the UK Labor Force Survey. *IZA−Institute of Labor Economics*. Discussion Paper No. 2228.

Clemens, M., Montenegro, C., and Pritchett, L. (2016). The Effect of Americanization on the Earnings of Foreign-born Men. *Center for Global Development*. Working Paper No.428.

Cumby, R. E. and Huizinga, J. (1990). Testing the Autocorrelation Structure of Disturbances in Ordinary Least Squares and Instrumental Variables Regressions. *NBER Technical Working Paper*, No. 10.

Davis, G. C. and Espinoza, M. C. (1998). A Unified Approach to Sensitivity Analysis in Equilibrium Displacement Models. *Southern Journal of Agricultural Economics*, 80:868–879.

De Hoyos, R. and Sarafidis, V. (2006). Testing for Cross-Sectional Dependence in Panel-Data Models. *The Stata Journal*, 6(4).

della Cava, M. and Lopez, R. (2019). Could California Produce Soon Cost You More? Farms Face Labor Shortages, Immigration Woes. *USA Today*, Retrieved Feb 23 from: https://www.usatoday.com/story/news/2019/01/27/could-california-produce-soon-cost-you-more/2594964002/?fbclid=IwAR0clU-zge5QHNCgz0oDPa6yplQrrILS6lpf1HrKS_pmVe_3CTJfOk7xlbA.

Dickerson, C. and Medina, J. (2017). California Farmers Backed Trump, but Now Fear Losing Field Workers. *New York Times*, Retrieved February 23, 2019 from: https://www.nytimes.com/2017/02/09/us/california-farmers-backed-trump-but-now-fear-losing-field-workers.html.

Driscoll, J. C. and Kraay, A. C. (1998). Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data. *The Review of Economics and Statisticss*, 80(4):549–560.

Duffield, J. A. (1990). Estimating Farm Labor Elasticities to Analyze The Effects Of Immigration Reform. *USDA, Economic Research Service, Agriculture and Rural Economy Division*. Staff Report No. AGES 9013.

Duffy, P. A. and Wohlgenant, M. K. (1991). Effects of an Export Subsidy on the U.S. Cotton Industry. *Southern Journal of Agricultural Economics*, 3(3):1–7.

Dustmann, C., Fabbri, F., and Preston, I. (2005). The Impact of Immigration on the British Labour Market. *The Economic Journal*, 115(507):F324–F341.

Dustmann, C., Schönberg, U., and Stuhler, J. (2016). The Impact of Immigration: Why Do Studies Reach Such Different Results? *The Journal of Economic Perspectives*, 30(4):31–56.

Dustmann, C., Schönberg, U., and Stuhler, J. (2017). Labor Supply Shocks, Native Wages, and the Adjustment of Local Employment. *The Quarterly Journal of Economics*, 132(1):435–483.

Fan, M., Gabbard, S., Pena, A. A., and Perloff, J. M. (2015). Why Do Fewer Agricultural Workers Migrate Now? *American Journal of Agriculutral Economics*, 97(3):665–679.

Fasani, F., Frattini, T., and Minale, L. (2018). (The Struggle for) Refugee Integration into the Labour Market: Evidence from Europe. *IZA−Institute of Labor Economics*. Discussion Paper No. 11333.

Fishman, S. (2013). The Underground Economy of Unreported Income. NOLO.com. Accessed on September 13, 2018 at `https://www.nolo.com/legal-encyclopedia/the-underground-economy-unreported-income.html`.

Frees, E. (1995). Assessing Cross-Sectional Correlation in Panel Data. *Journal of Econometrics*, 69:393–414.

Friedberg, R. M. and Hunt, J. (1995). The Impact of Immigrants on Host Country Wages, Employment and Growth. *The Journal of Economic Perspectives*, 9(2):23–44.

Frisch, R. and Waugh, F. V. (1933). Partial Time Regressions as Compared with Individual Trends. *Econometrica*, 1(4):387–401.

Gardner, B. L. (1975). The Farm-Retail Price Spread in a Competitive Food Industry. *American Journal of Agricultural Economics*, 57:399–409.

Giles, D. E. (1984). Instrumental Variables Regressions Involving Seasonal Data. *Economics Letters*, 14(4):339–343.

Glaister, D. (2006). US Crops Left to Rot as Mexicans Leave the Fields for Better-Paid Jobs. *The Guardian*. February 3, 2006. Retrieved December 30, 2018 from: https://www.theguardian.com/world/2006/feb/04/usa.uk.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2019). Bartik Instruments: What, When, Why, and How. Technical report, National Bureau of Economic Research. NBER Working Paper No. 24408.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik Instruments: What, When, Why, and How. *American Economic Review*, 110(8):2586–2624.

Good, K. (2017). Trump Administration Immigration Policies Could Hamper Some Ag Sectors. *Farm Policy News*, Retrieved Feb 23 from: https://farmpolicynews.illinois.edu/2017/02/trump-administration-immigration-policies-hamper-ag-sectors/.

Grossman, J. B. (1982). The Substitutability of Natives and Immigrants in Production. *The Review of Economics and Statistics*, 64(4):596–603.

Gunter, L. F., Jarett, J. C., and Duffield, J. A. (1992). Effect of U.S. Immigration Reform on Labor-Intensive Agricultural Commodities. *American Journal of Agricultural Economics*, 74(4):897–906.

Hagerty, N. (2019). The Scope for Climate Adaptation: Evidence from Irrigated Agriculture in California. *AAEA Conference Paper*.

Hainmueller, J., Hangartnerand, D., and Lawrence, D. (2016). When Lives are Put on Hold: Lengthy Asylum Processes Decrease Employment Among Refugees. *Social Advances*, 2(8).

Hertz, T. and Zahniser, S. (2012). Is There a Farm Labor Shortage? *American Journal of Agricgultural Economics*, 95(2):476–481.

Ho, G. and Turk-Ariss, R. (2018). The Labor Market Integration of Migrants in Europe: New Evidence from Micro Data. *IMF−International Monetary Fund*. Working Papers 18/32.

Hoechle, D. (2007). Robust Standard Errors for Panel Regressions with Cross-Sectional Dependence. *The Stata Journal*, 7(3):281–312.

Ifft, J. and Jodlowski, M. (2016). Is ICE Freezing US Agriculture? Impacts of Local Immigration Enforcement on US Farm Profitability and Structure. *AAEA Conference Paper*.

Jaeger, D. A., Ruist, J., and Stuhler, J. (2018a). Shift-Share Instruments and the Impact of Immigration. NBER Working Paper 24285. Accessed on April 5 2018 at http://www.nber.org/papers/w24285.

Jaeger, D. A., Ruist, J., and Stuhler, J. (2018b). Shift-Share Instruments and the Impact of Immigration. NBER Working Paper 24285. Accessed on April 5 2018 at `http://www.nber.org/papers/w24285`.

Jones, B. (2020). Prevention-by-Deterrence Policies Have Counterintuitive Relationship to Migrant Death Crisis. *U.C. Davis Global Migration Center: Immigration Fact*. Retrieved on April 15, 2020 from: `https://globalmigration.ucdavis.edu/prevention-deterrence-policies-have-counterintuitive-relationship-migrant-death-crisis`.

Khan, A., Martin, P., and Haridman, P. (2004). Expanded Production of Labor-Intensive Crops Increases Agricultural Employment. *California Agriculture*, 58(1).

Kostandini, G., Mykerezi, E., and Escalante, C. (2013). The Impact of Immigration Enforcement on te U.S. Farming Sector. *American Journal of Agricultural Economics*, 96(1):172–192.

Llull, J. (2017). The Effect of Immigration on Wages: Exploiting Exogenous Variation at the National Level. *Journal of Human Resources*, 53(3):608–662.

Lochmann, A., Rapoport, H., and Speciale, B. (2019). The Effect of Language Training on Immigrants' Economic Integration: Empirical Evidence from France. *European Economic Review*, 113:265–296.

López, G. and Bialik, K. (2017). Key Findings About U.S. Immigrants. Pew Research Center. Accessed at `http://www.pewresearch.org/fact-tank/2017/05/03/key-findings-about-u-s-immigrants/` on May 2, 2018.

Lovell, M. C. (1963). Seasonal Adjustment of Economic Time Series and Multiple Regression Analysis. *Journal of the American Statistical Association*, 58(304):993–1010.

Manacorda, M., Manning, A., and Wadsworth, J. (2012). The Impact of Immigration on the Structure of Wages: Theory and Evidence from Britain. *Journal of the European Economic Association*, 10(1):120–151.

Martin, P. (2006). Braceros: History, Compensation. *Rural Migration News*, 12(2).

Martin, P. (2019). Farm Labor Shortages. *Rural Migration News*. Retrieved from: `https://migration.ucdavis.edu/rmn/blog/post/?id=2264`.

Martin, P., Hooker, B., Akhtar, M., and Stockton, M. (2016). How Many Workers are Employed in California Agriculture? *UCANR, California Agriculture.* Retrieved December 18, 2018 from: http://calag.ucanr.edu/Archive/?article=ca.2016a0011.

Martin, W. E. (1966). Alien Workers in United States Agriculture: Impacts on Production. *Journal of Farm Economics*, 48(5):1137–1145.

Massey, D. S., Durand, J., and Malone, N. J. (2002). *Beyond Smoke and Mirrors: Mexican Immigration in an Era of Free Trade.* New York: Russel Sage Foundation.

Mérel, P. and Rutledge, Z. (2020). The Short-Run Impacts of Immigration on Native Workers: A Sectoral Approach. *The World Enclyclopaedia of Migration*. https://www.zachrutledge.com/uploads/1/2/5/6/125679559/versionbook.pdf.

Mincer, J. (1958). Investment in Human Capital and Personal Income Distribution. *The Journal of Political Economy*, 66(4):281–302.

Moretti, E. (2013). Real Wage Inequality. *American Economic Journal: Applied Economics*, 5(1):65–103.

Muth, R. F. (1964). The Derived Demand Curve for a Productive Factor and the Industry Supply Curve. *Oxford Economic Papers*, New Series 16(2):221–234.

NASS (2018). California County Ag Commissioners' Reports. Retrieved November 15th, 2018 from: `https://www.nass.usda.gov/Statistics_by_State/California/Publications/AgComm/index.php`.

NASS (2019). [website]. Retrieved March 12 from NASS Quickstats at: `https://quickstats.nass.usda.gov/`.

NASS (2020). Farm Labor Survey. Retrieved April 22nd, 2020 from NASS Quickstats at: `https://quickstats.nass.usda.gov/`.

Nevo, A. and Rosen, A. M. (2012). Identification with Imperfect Instruments. *The Review of Economics and Statistics*, 94(3):659–671.

NOAA (2019). Climate Data Online [website]. Retrieved at: `https://www.ncdc.noaa.gov/cdo-web/`.

Oatman, M. (2018). California's Vineyard Workers Already Faced Long Hours, Low Pay, and Harsh Conditions. Then Came Trump's Immigration Crackdown. *Mother Jones*. Retrieved Feb 23 from: https://www.motherjones.com/food/2018/07/californias-vineyard-workers-wine-labor-shortage-guestworker-visas-immigration-crackdown/.

Orrenius, P. (2004). The Effect of U.S. Border Enforcement on the Crossing Behavior of Mexican Migrants. Chapter 14 in *Crossing the Border: Research from the Mexican Migration Project*. Russel Sage Foundation.

Ottaviano, G. I. and Peri, G. (2012). Rethinking the Effect of Immigration on Wages. *Journal of the European Economic Association*, 10(1):152–197.

Passel, J. S., Cohn, D., and Gonzalez-Barrera, A. (2012). Net Migration from Mexico Falls to Zero - and Perhaps Less. *Pew Research Center - Hispanic Trends*, April 23.

Peri, G. and Sparber, C. (2009). Task Specialization, Immigration, and Wages. *American Economic Journal: Applied Economics*, 1(3):135–169.

Piggott, R. R., Piggott, N. E., and Wright, V. E. (1995). Approximating Farm-Level Returns to Incremental Advertising Expenditure: Methods and an Application to the Australian Meat Industry. *American Journal of Agricultural Economics*, 77(3):497–511.

Piil Damm, A. (2009). Ethnic Enclaves and Immigrant Labor Market Outcomes: Quasi-Experimental Evidence. *Journal of Labor Economics*, 27(2):281–314.

Plummer, B. (2013). We're Running Out of Farm Workers. Immigration Reform Won't Help. *Washington Post*, Retrieved Feb 23 from: https://www.washingtonpost.com/news/wonk/wp/2013/01/29/the-u-s-is-running-out-of-farm-workers-immigration-reform-may-not-help/?utm_term==.85a4c6595cb2.

Reed, D. and Danziger, S. (2007). The Effects of Recent Immigration on Racial/Ethnic Labor Market Differentials. *The American Economic Review*, 97(2):373–377.

Richards, T. J. (2018). Immigration Reform and Farm Labor Markets. *American Journal of Agricultural Economics*, 100(4):1050–1071.

Richards, T. J. and Patterson, P. M. (1998). Hysteresis and the Shortage of Agricultural Labor. *American Journal of Agricultural Economics*, 80:683–695.

Riosmena, F. (2004). Return versus Settlement Among Undocumented Mexican Immigrants. Chapter 13 in *Crossing the Border: Research from the Mexican Migration Project*. Russel Sage Foundation.

Rogers, W. (1993). Regression Standard Errors in Clustered Samples. *Stata Technical Bulletin*, 13.

Ruggles, S., Genedek, K., Goeken, R., Grover, J., and Sobek, M. (2017). Integrated Pulbic Use Microdata Series: Version 7.0 [dataset]. Minneapolis: University of Minnesota. https://doi.org/10.18128/D010.V7.0.

Ruggles, S., Genedek, K., Goeken, R., Grover, J., and Sobek, M. (2019). Integrated Pulbic Use Microdata Series: Version 7.0 [dataset]. Minneapolis: University of Minnesota. https://doi.org/10.18128/D010.V7.0.

Rutledge, Z. and Taylor, J. E. (2019a). California Farmers Change Production Practices as the Farm Labor Supply Declines. *ARE Update*, 22(6).

Rutledge, Z. J. and Taylor, J. E. (2019b). Farmworkers and Non-Farm Work: Evidence from the NAWS (Working Paper). Retrieved from author's personal website: `https://www.zachrutledge.com/uploads/1/2/5/6/125679559/rutledge-taylor_naws_paper.pdf`.

Ryssdal, K. (2017). As Farmers Grow Scarce, Wages Are on the Rise. *Marketplace.org*. Retrieved from: `https://www.marketplace.org/2017/03/23/farmworkers-grow-scarce-farm-wages-go/`.

Sarvimäki, M. and Hämäläinen, K. (2016). Integrating Immigrants: The Impact of Restructuring ALMP. *Journal of Labor Economics*, 34(2):479–508.

Shultz, M. K. (1965). The Bracero & California Agriculture. *Industrial and Labor Relations Forum*, 2(2):143–160.

Smith, A. (2010). Farm Workers: Take Our Jobs, Please! *CNN Money*, Retrieved Feb 24 from: https://money.cnn.com/2010/07/07/news/economy/farm_worker_jobs/index.htm.

Taylor, J. E., Charlton, D., and Yúnez-Naude, A. (2012). The End of Farm Labor Abundance. *Applied Economic Perspectives and Policy*, 34(4):587–598.

Thilmany, D. (1996). FLC Usage among California Growers under IRCS: An Empirical Analysis of Farm Labor Market Risk Management. *American Journal of Agricultural Economics*, 78(4).

Thilmany, D. and Blank, S. C. (1996). FLCs: An Analysis of Labor Management Transfers among Agricultural Producers. *Agribusiness: An International Journal*, 12(1).

U.S. Department of Labor, B. o. L. S. (2018). Foreign-Born Workers: Labor Force Characteristics—2017. News release USDL-18-0786, May 17, 2018.

Villareal, A. and Tamborini, C. (2018). Immigrants' Economic Assimilation: Evidence from Longitudinal Earnings Records. *American Sociological Review*, 83(4):686–715. DOI: 10.1177/0003122418780366.

White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica*, 48:817–838.

Wohlgenant, M. K. (1993). Distribution of Gains from Research and Promotion in Multi-Stage Production Systmes: The Case of the U.S. Beef and Pork Industries. *Southern Journal of Agricultural Economics*, 75:642–651.

Zahniser, S., Hertz, T., Dixon, P., and Rimmer, M. (2011). U.S. Agriculture and the Market for Hired Farm Labor: A Simulation Analysis. *American Journal of Agricultural Economics*, 94(2):477–482.

Zahniser, S., Taylor, J. E., Hertz, T., and Charlton, D. (2018). Farm Labor Markets in the United States and Mexico Pose Challenges for U.S. Agriculture. *USDA-ERS Economic Research Bulletin*, No. 201.

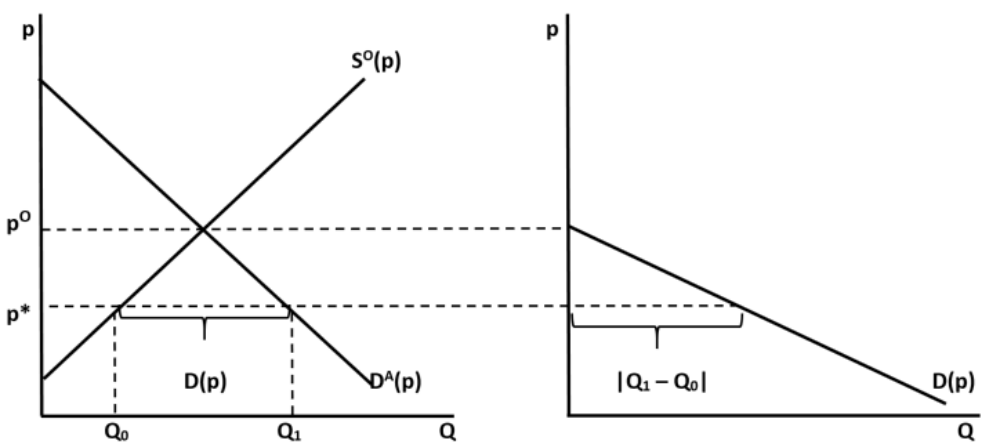
Zahniser, S., Taylor, J. E., Hertz, T., and Charlton, D. (2019). Rising Wages Point to a Tighter Farm Labor Market in the United States. *Amber Waves*. Retrieved August 20, 2020 from: `https://www.ers.usda.gov/amber-waves/2019/february/rising-wages-point-to-a-tighter-farm-labor-market-in-the-united-states`.

Appendices

1.A Derivation of County-Level Output Demand Elasticity

Let p denote the output price, $D^A(p)$ denote the aggregate demand for output, $S^O(p)$ denote the output supply of all other regions (i.e., aggregate supply minus the supply of county C), and $D(p)$ denote the output demand for county C. In Figure 1.A.1, the equilibrium price is denoted by p^*. At the equilibrium price, consumers demand Q_1 tons of fruits and vegetables, but suppliers in all the other regions only supply Q_0 at that price. At price p^*, the demand facing county C is equal to the amount $|Q_1 - Q_0|$.

Fig. 1.A.1 Aggregate and County-Level Demand



Mathematically, the county-level demand can be expressed as follows:

$$D(p) = D^A(p) - S^O(p). \tag{1.A.1}$$

By taking the derivative of (1.A.1), one can derive the following equation:

$$\frac{dD(p)}{dp} = \frac{dD^A(p)}{dp} - \frac{dS^O(p)}{dp}. \tag{1.A.2}$$

By multiplying both sides of (1.A.2) by p/Q, where Q is the equilibrium quantity in county C, one can derive the following equation:

$$\eta = \eta^A \frac{Q^A}{Q} - e_Q^O \frac{Q^O}{Q},$$

where $\eta \leq 0$ is the output demand elasticity for county C, $\eta^A \leq 0$ is the aggregate output demand elasticity, $e_Q^O \geq 0$ is the output supply elasticity for the other regions, $Q^A > Q$ is the aggregate output, and $Q^O > Q$ is the quantity supplied by all other regions. Because $\eta^A \leq 0$, $e_Q^O \geq 0$, $\frac{Q^A}{Q} > 1$ and $\frac{Q^O}{Q} > 1$, it follows that $\eta \leq \eta^A$. Therefore, the demand facing the a county is more elastic than the aggregate demand.

Assuming that the aggregate output demand and output supply of all other regions are not perfectly elastic or inelastic (i.e., $\eta^A < 0$, $\eta^A > -\infty$, $e_Q^O > 0$, and $e_Q^O < \infty$), the extent to which the demand facing the county is more elastic than the aggregate demand depends on the inverse of the share of output produced by the county ($\frac{Q^A}{Q}$), the output supply elasticity of the other regions (e_Q^O), and the ratio of the output in the other counties relative to the output in the county of interest ($\frac{Q^O}{Q}$). Even if a county produces 5% of the total fruit and vegetable output (i.e., if $\frac{Q^A}{Q} = 20$), the elasticity of demand facing that county will be at least 20 times larger than the aggregate demand elasticity. Moreover, if the output supply from the other regions is not perfectly inelastic (i.e., if $e_Q^O > 0$), the county-level output demand elasticity can potentially be much larger.

1.B Proof that $\Lambda \geq 0$

Let Λ be defined as[10]

$$
\begin{aligned}
\Lambda \quad &\equiv \left[\frac{\xi_3\rho_1 - \xi_1\rho_3}{\rho_1}\right] = \\
&= \left[\frac{\sigma^2(1 + k_A e_A + k_B e_B) + \sigma(k_B e_A + k_A e_B + e_A e_B) + \sigma(1 + k_A e_A + k_B e_B)k_A e_B}{D(\sigma + k_A e_B)}\right] + \\
&+ \left[\frac{k_A k_B e_A e_B + k_A^2 e_B^2 + k_A e_A e_B^2 - \sigma^2 k_A e_A - \sigma k_A e_A e_B - \sigma k_A e_A e_B - k_A e_A e_B^2}{D(\sigma + k_A e_B)}\right] \qquad (1.B.1) \\
&= \left[\frac{e_A e_B k_B(k_A + \sigma k_B) + k_A e_B^2(k_A + \sigma k_B) + \sigma e_B(k_A + \sigma k_B) + \sigma D}{D(\sigma + k_A e_B)}\right] \qquad (1.B.2) \\
&= \left[\frac{(k_A + \sigma k_B)(e_A e_B k_B + k_A e_B^2 + \sigma e_B) + \sigma D}{D(\sigma + k_A e_B)}\right] \\
&= \left[\frac{(k_A + \sigma k_B)e_B D + \sigma D}{D(\sigma + k_A e_B)}\right] \\
&= \left[\frac{D[\sigma + e_B(k_A + \sigma k_B)]}{D(\sigma + k_A e_B)}\right] \\
&= \left[\frac{\sigma + e_B(k_A + \sigma k_B)}{(\sigma + k_A e_B)}\right] \geq 0.
\end{aligned}
$$

Therefore, $\Lambda \geq 0$.

1.C Proof that $\Upsilon \geq 0$

Let Υ be defined as

$$
\begin{aligned}
\Upsilon &\equiv \left[\frac{\xi_2\rho_1 - \xi_1\rho_2}{\rho_1}\right] = \\
&= \left[\frac{k_B(\sigma + e_A)(\sigma + k_A e_B) - k_A k_B e_A(\sigma + e_B)}{D(\sigma + k_A e_B)}\right].
\end{aligned}
$$

In order to prove that $\Upsilon \geq 0$, it is sufficient to show that

$$k_B(\sigma + e_A)(\sigma + k_A e_B) \geq k_A k_B e_A(\sigma + e_B).$$

[10] To derive equation (1.B.2) from (1.B.1), I make use of the identity $k_B^2 = (1 - k_A)^2$ such that $\sigma e_A e_B - 2\sigma k_A e_A e_B + \sigma k_A^2 e_A e_B = \sigma e_A e_B k_B^2$.

Dividing both sides of (1.C.1) by $k_B e_A$ delivers the following inequality:

$$\underbrace{\left[\frac{(\sigma + e_A)}{e_A}\right]}_{G}(\sigma + k_A e_B) \geq k_A \sigma + k_A e_B,$$

which holds because $G \geq 1 \geq k_A$ and thus $G\sigma \geq k_A \sigma$ and $G k_A e_B \geq k_A e_B$. Therefore, $\Upsilon \geq 0$.

1.D Simulations for the Relative Magnitudes of θ_1, θ_2, and θ_3

In order to determine how the magnitudes of each source of bias compare to each other, I assign each structural parameter and variance term a value from a random uniform distribution and simulate estimates for the bias terms 100,000 times. All parameter values are taken from the $\mathrm{U}(0, 9999)$ distribution except for k_A, which is taken from the $\mathrm{U}(0, 1)$ distribution where $k_B = 1 - k_A$.

1.D.1 Relative Magnitude of θ_1 and θ_2

Let θ_1 and θ_2 be defined as follows:

$$\theta_1 = \frac{k_A k_B e_A(\sigma + e_B)}{(\sigma + k_A e_B)(\sigma + k_B e_A + k_A e_B)}$$

$$\theta_2 = \frac{\Lambda}{\rho_3}\left[\frac{\rho_3^2 \mathrm{var}(\ln(\delta'))}{\rho_1^2 \mathrm{var}(\ln(\beta')) + \rho_3^2 \mathrm{var}(\ln(\delta'))}\right].$$

If the value determined by the ratio θ_1/θ_2 is always greater than the value one, then that would provide evidence that θ_1 is always greater than θ_2. However, if the simulations produce values that are both greater than one and less than one, then the relative magnitude of the two sources of bias cannot be determined without additional information. As can be seen in row one of Table 1.D.1, the values range from very close to zero to very large, indicating that the relative magnitude of θ_1 to θ_2 is uncertain.

Table 1.D.1 Summary Statistics from Simulations for Bias Magnitudes

	Mean	Median	Min	Max
Value of θ_1/θ_2	3.21	0.45	3.69×10^{-9}	2,169.70
Value of θ_1/θ_3	3.21	0.45	3.69×10^{-9}	2,169.70
Value of θ_1/θ_{3A}	3.21	0.45	3.69×10^{-9}	2,169.70
Value of θ_1/θ_{3B}	4.31×10^{9}	3.53×10^{7}	0.10	3.23×10^{13}

1.D.2 Relative Magnitude of θ_1 and θ_3

Let θ_1 be defined as in Appendix 1.D.1 and θ_3 be defined as follows:

$$\theta_3 = \underbrace{\frac{\Lambda}{\rho_3}\left[\frac{\rho_3^2 \text{var}(\ln(\delta'))}{E}\right]}_{\theta_{3A}} + \underbrace{\frac{\Upsilon}{\rho_2}\left[\frac{\rho_2^2 \text{var}(\ln(\delta'))}{E}\right]}_{\theta_{3B}}, \tag{1.D.1}$$

where

$$E = \rho_1^2 \text{var}(\ln(\beta')) + \rho_2^2 \text{var}(\ln(\gamma')) + \rho_3^2 \text{var}(\ln(\delta')).$$

In order to determine if $\theta_1 < \theta_3$ (or vice versa), I simulate values for each source of bias and calculate the ratio θ_1/θ_3. The results can be seen in the second row of Table 1.D.1. As the table reveals, the values, when rounded to two decimal points, are identical to the results in row 1. Therefore, the relative magnitude of θ_1 to θ_3 cannot be determined without additional information. This result also applies to each source of omitted variables bias, defined as θ_{3A} and θ_{3B} in equation (1.D.1), considered independently of each other. As can be seen in rows 3 and 4 of Table 1.D.1, the values of θ_1/θ_{3A} and θ_1/θ_{3B} range from close to zero to much greater than one.

1.E Derivation of 2SLS Bias

Suppose I want to estimate Γ using equation (1.11) but there are two omitted variables, $\ln(\gamma')$ and $\ln(\delta')$. Assuming there is an instrumental variable (Z) that satisfies the following exclusion restriction: $E[\nu|Z] = \text{cov}(Z, \ln(\gamma')) = \text{cov}(Z, \ln(\delta')) = \text{cov}(Z, e) = 0$, Γ could be estimated using the following 2SLS model:

$$\ln(Q) = c + \Gamma \ln(A) + \underbrace{\Upsilon \ln(\gamma') + \Lambda \ln(\delta') + e}_{\nu} \tag{1.E.1}$$

$$\ln(A) = d + \Pi Z + u$$

where

$$\widehat{\ln(A)} = d + \Pi Z. \tag{1.E.2}$$

Note that $\Pi = \frac{\text{cov}(Z, \ln(A))}{\text{var}(Z)}$ and $\text{var}(\widehat{\ln(A)}) = \Pi^2 \text{var}(Z)$. The probability limit of the 2SLS coefficient on the variable $\ln(A)$ is

$$\Gamma_{2SLS} = \frac{\text{cov}(\widehat{\ln(A)}, \ln(Q))}{\text{var}(\widehat{\ln(A)})}. \tag{1.E.3}$$

By substituting (1.E.1), (1.E.2), and the formula for $\text{var}(\widehat{\ln(A)})$ into (1.E.3), one can derive the following formula for Γ_{2SLS}:

$$\begin{aligned}\Gamma_{2SLS} &= \frac{\Gamma}{\Pi}\overbrace{\left[\frac{\text{cov}(Z,\ln(A))}{\text{var}(Z)}\right]}^{=\Pi} + \\ &\quad + \frac{\Upsilon}{\Pi}\overbrace{\left[\frac{\text{cov}(Z,\ln(\gamma'))}{\text{var}(Z)}\right]}^{=0} + \frac{\Lambda}{\Pi}\overbrace{\left[\frac{\text{cov}(Z,\ln(\delta'))}{\text{var}(Z)}\right]}^{=0} + \frac{1}{\Pi}\overbrace{\left[\frac{\text{cov}(Z,e)}{\text{var}(Z)}\right]}^{=0} \\ &= \Gamma.\end{aligned}$$

But Γ is the OLS coefficient from equation (1.11) under the case where there are no omitted variables, and it was shown in section 1.3.2.1 that

$$\Gamma = \frac{\xi_1}{\rho_1} = \xi_1 + \theta_1.$$

Therefore, using 2SLS with an instrument that satisfies the exclusion restriction will alleviate the omitted variables bias but not the employment-labor supply mismatch bias (i.e., θ_1 will remain).

1.F List of Labor-Intensive Commodities Used in Analysis

Table 1.F.1 List of Hand-Harvested Fruit and Vegetable Crops

ANISE (FENNEL)	LETTUCE LEAF
APPLES ALL	LETTUCE ROMAINE
APRICOTS ALL	LIMES ALL
ARTICHOKES	MELONS CANTALOUPE
ASPARAGUS FRESH MARKET	MELONS HONEYDEW
AVOCADOS ALL	MELONS UNSPECIFIED
BEANS FRESH UNSPECIFIED	MELONS WATERMELON
BEANS SNAP FRESH MARKET	MUSHROOMS
BERRIES BUSHBERRIES UNSPECIFIED	NECTARINES
BERRIES RASPBERRIES	OKRA
BERRIES STRAWBERRIES FRESH MARKET	OLIVES
BERRIES STRAWBERRIES PROCESSING	ONIONS GREEN & SHALLOT
BROCCOLI FOOD SERVICE	ORANGES NAVEL
BROCCOLI FRESH MARKET	ORANGES VALENCIA
BROCCOLI PROCESSING	PARSLEY
CABBAGE CHINESE & SPECIALTY	PEACHES CLINGSTONE
CABBAGE HEAD	PEACHES FREESTONE
CAULIFLOWER FOOD SERVICE	PEACHES UNSPECIFIED
CAULIFLOWER FRESH MARKET	PEARS ASIAN
CELERY FOOD SERVICE	PEARS UNSPECIFIED
CELERY FRESH MARKET	PEAS EDIBLE POD (SNOW)
CHERRIES SWEET	PEPPERS BELL
CITRUS UNSPECIFIED	PEPPERS CHILI HOT
CORN SWEET ALL	PERSIMMONS
CUCUMBERS	PLUMS
CUCUMBERS GREENHOUSE	POMEGRANATES
DATES	PUMPKINS
EGGPLANT ALL	QUINCE
ENDIVE ALL	RADICCHIO
ESCAROLE ALL	RAPPINI
GRAPEFRUIT ALL	SALAD GREENS MISC.
GRAPES TABLE	SQUASH
KALE	SWISS CHARD
KIWIFRUIT	TANGELOS
KUMQUATS	TANGERINES & MANDARINS
LEMONS ALL	TOMATOES CHERRY
LETTUCE HEAD	TOMATOES FRESH MARKET

2.A City Comparisons and the Short-Run Wage Effects of Immigration

This section shows that in the presence of trade in capital between cities, the spatial correlation approach tends to underestimate the overall impact of immigration on wages, even if there is no trade in goods across cities.

Consider two cities, A and B. In the short run, capital is mobile between cities, but fixed in the aggregate at $\bar{K}$. Labor L_i, $i \in \{A, B\}$, is immobile. For the sake of the argument, here we assume that immigrant and native labor are perfectly substitutable and that labor is supplied perfectly inelastically. Each city uses the same constant-returns-to-scale technology to produce a homogeneous good Q_i: $Q_i = f(L_i, K_i)$. The production function satisfies the law of diminishing marginal returns. The associated unit cost function is denoted $c(w, r)$, with w the wage rate and r the rental on capital. The labor endowment of city B is assumed to be fixed at $\bar{L}_B$, while city A experiences an increase in its labor endowment due to immigration, $\Delta L_A > 0$. For simplicity, we assume that demand in city A is unaffected by immigration, and we write the demand functions as $Q_i = D_i(p_i)$, with $D_i' < 0$ and p_i the local price of the good.

We are interested in the comparative statics $\frac{\partial w_i}{\partial L_A}$, for $i = A, B$, and also in the difference between them, which is what would be identified by exploiting city comparisons in a spatial correlation approach.

2.A.1 Scenario 1: Traded Good

If the good is traded between cities, then in equilibrium we have $p_A = p_B$. Under constant returns to scale, we also have $p_i = c(w_i, r)$. Therefore, we must have $w_A = w_B$ (the cost function is monotonically increasing in input prices), and as a result the wage is equalized between cities. Intercity comparisons will reveal an absence of a wage effect.

Nonetheless, the wage decreases in both cities. To see why, note that in equilibrium the wage-to-output-price ratio must be equal to the marginal product of labor in each city, i.e., $\frac{w_i}{p_i} = \frac{\partial f}{\partial L}\left(\frac{L_i}{K_i}, 1\right)$, where we have used the fact that the marginal product of labor is homogeneous of degree zero. Since total labor increases in the aggregate due to $\Delta L_A > 0$, while total capital is fixed, the ratio $\frac{L_i}{K_i}$ increases in each city. Because the marginal product of labor decreases in the labor argument, the ratio $\frac{w_i}{p_i}$ declines. Since demand slopes down in each city and the additional labor results in more output in each city, output prices must decline. As a result, wages w_i decline as well.

In this scenario with traded good and traded capital between cities, intercity comparisons of wages would thus reveal *none* of the short-run wage effects of immigration. Note that if the good was traded internationally rather than just between cities, the same conclusion would obtain as long as capital is fixed in the aggregate. If capital and the good were traded internationally, then there would be no wage effect of immigration.

2.A.2 Scenario 2: Non-Traded Good

If the good is not traded between cities, then the equilibrium can be described by the following set of equations:

$$
\begin{aligned}
D_A(p_A) &= f(L_A, K_A) \\
D_B(p_B) &= f(\bar{L}_B, K_B) \\
\frac{w_A}{p_A} &= \frac{\partial f}{\partial L}\left(\frac{L_A}{K_A}, 1\right) \qquad (2.A.1) \\
\frac{w_B}{p_B} &= \frac{\partial f}{\partial L}\left(\frac{\bar{L}_B}{K_B}, 1\right) \\
p_A \frac{\partial f}{\partial K}\left(1, \frac{K_A}{L_A}\right) &= p_B \frac{\partial f}{\partial K}\left(1, \frac{K_B}{\bar{L}_B},\right) \qquad (2.A.2) \\
K_A + K_B &= \bar{K}
\end{aligned}
$$

which constitute a system of 6 equations in 6 unknowns: p_A, p_B, K_A, K_B, w_A, and w_B. We are interested in the effect of a change in L_A, $\Delta L_A > 0$, on these equilibrium variables, specifically the wages w_i.

Case 1: Gross Complements

First assume that in each city, labor and capital are *gross complements*: that is, an increase in the labor endowment results in an increase in the derived demand for capital. As shown, for instance, in Muth (1964), labor and capital are gross complements whenever the substitution elasticity in production is lower than the (absolute) output demand elasticity. This happens if capital and labor are not too substitutable and output demand is not too inelastic.

Under the assumption of gross complements, the demand for capital rises in city A, which leads to a transfer of capital from city B to city A: $\Delta K_A = -\Delta K_B > 0$. Because labor and capital are gross complements, the outflow of capital from city B results in a reduction in the derived demand for labor, and therefore a reduction in the wage w_B. Output declines in city B, and thus output price increases: $\Delta p_B > 0$. But then, condition (2.A.2) together with the fact that the marginal product of capital decreases in the capital-to-labor ratio implies that either p_A increases or the ratio $\frac{K_A}{L_A}$ decreases or both. Since $\Delta L_A > 0$ and $\Delta K_A > 0$, output increases in city A and therefore $\Delta p_A < 0$. Therefore, $\Delta\left(\frac{K_A}{L_A}\right) < 0$, and condition (2.A.1) implies that the wage-to-output-price ratio declines in city A. Since $\Delta p_A < 0$, $\Delta w_A < 0$.

Summarizing, the wage w_i declines in both cities. If we relax the assumption that the inflow of labor into city A does not change the output demand, the conclusion that w_B declines still holds because capital still flows to city A due to the combined effects of labor-capital gross complementarity and the increase in output demand. The conclusion that w_A declines holds as long as it is still the case that $\Delta p_A < 0$, that is, the increase in output demand is not so high as to result in an output price increase. This will hold if the immigrant inflow makes local goods cheaper.

Case 2: Gross Substitutes

Now assume that labor and capital are gross substitutes in both cities. If the immigrant inflow does not shift the output demand in city A (or not too much), then the derived demand for capital decreases and capital flows towards city B. In city B, the derived demand for labor declines due to gross substitutability, hence the wage rate decreases. Output increases and output price decreases. Condition (2.A.2) then implies that p_A decreases. Condition (2.A.1) implies that $\frac{w_A}{p_A}$ decreases, and therefore w_A decreases as well.

Because the spatial correlation approach identifies the effect of immigration from comparing wage changes between city A (the treatment city) and city B (the control city), and the wage declines in both cities, this approach underestimates the total effect and might even predict a positive wage effect if the wage decline in city A is less than in city B.

2.B Borjas' "Relevant Wage Elasticity"

In his book *Immigration Economics*, as well as in earlier work (Borjas 2003), George Borjas defines the "relevant wage elasticity" as the percentage change in native wages associated with a percent change in labor supply attributable to immigration (past and present). Denote by w the native wage, by $m = \frac{M}{N}$ the ratio of the immigrant to native workforce, and by $p = \frac{M}{M+N}$ the share of immigrants in the workforce. Borjas' relevant elasticity is then $\eta = \frac{\partial \ln w}{\partial m}$, while the elasticity given by the coefficient on the immigrant share in a regression of the log wage is $\beta = \frac{\partial \ln w}{\partial p}$. Because $p = \frac{m}{1+m}$, it follows that $\eta = \frac{\beta}{(1+m)^2} = \beta\left(1-p\right)^2$. Therefore, Borjas' "relevant wage elasticity" is directly deducible from the regression of log wage on the immigrant share.

Card and Peri (2016) (and other authors) choose to regress the first-difference of the log wage, $\Delta \ln w$, on the regressor $\frac{\Delta M}{M_{-1}+N_{-1}}$, where $\Delta M = M - M_{-1}$ is the (net) immigrant inflow between the prior and current periods and N_{-1} is the number of native workers in the prior period. Although this specification is sometimes referred to as a first-difference model in the literature, it cannot be obtained by first-differencing any underlying data generating process (DGP) for the determination of wages. Rather, it is a *sui generis* DGP that specifies wage growth as a function of the relative inflow of immigrants. The wage elasticity in Card and Peri (2016)'s model is $\epsilon = \frac{\partial \Delta \ln w}{\partial\left(\frac{\Delta M}{M_{-1}+N_{-1}}\right)}$. Only if $N_{-1} = N$ and $M_{-1} = 0$ can ϵ be related to η. In that case, $\frac{\Delta M}{M_{-1}+N_{-1}} = \frac{M}{N} = \frac{M}{N} - \frac{M_{-1}}{N_{-1}} = \Delta\left(\frac{M}{N}\right)$ and Card and Peri (2016)'s regression becomes the first-differenced version of a regression with m as the regressor, which implies $\epsilon = \eta$.

2.C Residualized Dependent Variables

As explained in Section 2.4.1, the dependent variables used in our analysis are generated by running sector-specific regressions with individual-level outcomes on a full set of MSA-

year fixed effects as well as a set of individual observables.[11] Following Reed and Danziger (2007), we use the MSA-year effects to construct "residualized" dependent variables that are used in our final analysis, the difference being that we construct MSA-year effects separately for each economic sector considered. The regressions we use to residualize our dependent variables are commonly referred to as Mincer models, which originates from the work of Mincer (1958) who is accredited with pioneering the use of factors other than school, such as work experience, to explain differences in individual labor market outcomes. Our model controls for educational attainment, race, potential work experience, gender, and marital status, as follows (to alleviate notation, there is no explicit index to denote the sector):

$$\begin{aligned} \mathrm{O}_{kit} &= \gamma_0 + y_{it} + \gamma_1 \mathrm{HS}_{kit} + \gamma_2 \mathrm{AA}_{kit} + \gamma_3 \mathrm{Black}_{kit} + \gamma_4 \mathrm{Other}_{kit} \\ &\quad + \gamma_5 \mathrm{Exp}_{kit} + \gamma_6 \mathrm{Exp}^2_{kit} + \gamma_7 \mathrm{Fem}_{kit} + \gamma_8 \mathrm{Mar}_{kit} + \psi_{kit} \end{aligned}$$

where O_{kit} is the outcome for individual k in MSA i in survey year t, HS_{kit} is a dummy variable that identifies individuals who have at least a high school education but not an Associate's (or higher) degree, AA_{kit} is a dummy variable for having at least an Associate's degree, Black_{kit} is a dummy variable that identifies black individuals, Other_{kit} is a dummy variable that identifies individuals who are neither white or black, Exp_{kit} is an individual's potential work experience (assumed to be non-negative), which is defined as the individual's age minus their years of schooling minus six (the typical age for starting school), Exp^2_{kit} is potential work experience squared, Fem_{kit} is a dummy variable for being female, Mar_{kit} is a dummy variable for being married, and ψ_{kit} is the error term. The MSA-year fixed effects from these regressions y_{it} are then used as the dependent variables in our main sectoral analysis. For a given sector, y_{it} effectively captures the average outcome for each MSA in each year after controlling for a set of individual-level observables.

[11] This is not a panel regression. Although the sample contains multiple time periods, we are not able to identify individuals over time. For each sector, we pool the individual observations from each cross-section into a single sample and run the regression on the sample of pooled cross-sections.

2.D Annual Earnings Results for the Manufacturing and Higher-Skilled Sectors

Table 2.D.1 Effect of Immigration on the Annual Earnings of Native-born Workers

			(1) OLS	(2) IIV-10	(3) IIV-All	(4) IIV-All but	(5) IIV-All but Pop
Computers	All occ.		0.137*	0.122	0.108	0.108	0.103
			(0.077)	(0.078)	(0.077)	(0.076)	(0.079)
		N	1,386	1,386	1,386	1,386	1,386
	Exposed occ.		0.181*	0.149	0.137	0.139	0.139
			(0.102)	(0.095)	(0.097)	(0.096)	(0.099)
		N	1,386	1,386	1,386	1,386	1,386
Engineering	All occ.		0.056	0.029	0.022	0.023	0.028
			(0.053)	(0.054)	(0.057)	(0.057)	(0.054)
		N	1,385	1,385	1,385	1,385	1,385
	Exposed occ.		-0.046	-0.105	-0.114	-0.112	-0.096
			(0.105)	(0.113)	(0.116)	(0.116)	(0.110)
		N	1,377	1,377	1,377	1,377	1,377
Sciences	All occ.		-0.060	-0.085	-0.080	-0.080	-0.085
			(0.090)	(0.094)	(0.091)	(0.091)	(0.092)
		N	1,382	1,382	1,382	1,382	1,382
	Exposed occ.		-0.081	-0.134	-0.121	-0.120	-0.130
			(0.174)	(0.180)	(0.181)	(0.180)	(0.179)
		N	1,363	1,363	1,363	1,363	1,363
Manufacturing	All occ.		0.057	0.053	0.026	0.049	0.003
			(0.083)	(0.136)	(0.149)	(0.195)	(0.168)
		N	1,387	1,387	1,387	1,387	1,387
	Exposed occ.		0.040	-0.060	-0.101	-0.127	-0.127
			(0.113)	(0.171)	(0.193)	(0.272)	(0.220)
		N	1,387	1,387	1,387	1,387	1,387

Note: All regressions include MSA and year fixed effects. Standard errors are clustered at the MSA level. Column (1) reports the OLS estimate of β in Equation (2.17). Column (2) reports the IV estimate obtained by using the immigrant share in the 10 sectors with the highest immigrant shares. Column (3) reports the IV estimate obtained by using the immigrant share across all sectors. Column (4) reports the IV estimate obtained by using the immigrant share across all other sectors. Columns (2) - (4) use instruments constructed from individuals who were between the ages of 18-64, were not in school, were not living in group quarters, and who were jointly in the labor force at the time of the survey and worked a positive number of weeks during the previous year. Column (5) reports the IV estimate obtained by using the immigrant share of the entire population across all other sectors of the economy. All occ. refers to the analysis conducted on all occupations within the sector. Exposed occ. refers to the analysis conducted on the immigrant-exposed occupations within the sector. * (resp. **, resp. ***) denotes statistical significance at the 10% (resp. 5%, resp. 1%) level.

2.E Employment Results for the Higher-Skilled Sectors

Table 2.E.1 Effect of Immigration on the Employment Rate of Native-born Workers

			(1) OLS	(2) IIV-10	(3) IIV-All	(4) IIV-All but	(5) IIV-All but Pop
Computers	All occ.		0.022	0.002	0.015	0.016	-0.030
			(0.025)	(0.067)	(0.069)	(0.088)	(0.087)
		N	1,386	1,386	1,386	1,386	1,386
	Exposed occ.		0.012	0.010	0.009	0.009	-0.025
			(0.026)	(0.064)	(0.027)	(0.027)	(0.081)
		N	1,386	1,386	1,386	1,386	1,386
Engineering	All occ.		-0.018	-0.042	-0.092	-0.119	-0.084
			(0.020)	(0.064)	(0.072)	(0.096)	(0.093)
		N	1,385	1,385	1,385	1,385	1,386
	Exposed occ.		-0.021	-0.092	-0.168*	-0.215	-0.165
			(0.027)	(0.093)	(0.100)	(0.136)	(0.119)
		N	1,377	1,377	1,377	1,377	1,377
Sciences	All occ.		-0.019	-0.020	-0.040	-0.049	-0.204
			(0.023)	(0.023)	(0.103)	(0.142)	(0.137)
		N	1,382	1,382	1,382	1,382	1,382
	Exposed occ.		-0.021	-0.022	-0.103	-0.136	-0.042
			(0.043)	(0.041)	(0.162)	(0.219)	(0.235)
		N	1,363	1,363	1,363	1,363	1,363

Note: All regressions include MSA and year fixed effects. Standard errors are clustered at the MSA level. Column (1) reports the OLS estimate of β in Equation (2.17). Column (2) reports the IV estimate obtained by using the immigrant share in the 10 sectors with the highest immigrant shares. Column (3) reports the IV estimate obtained by using the immigrant share across all sectors. Column (4) reports the IV estimate obtained by using the immigrant share across all other sectors. Columns (2) - (4) use instruments constructed from individuals who were between the ages of 18-64, were not in school, were not living in group quarters, and who were jointly in the labor force at the time of the survey and worked a positive number of weeks during the previous year. Column (5) reports the IV estimate obtained by using the immigrant share of the entire population across all other sectors of the economy. All occ. refers to the analysis conducted on all occupations within the sector. Exposed occ. refers to the analysis conducted on the immigrant-exposed occupations within the sector. * (resp. **, resp. ***) denotes statistical significance at the 10% (resp. 5%, resp. 1%) level.

2.F Relative Endogeneity of Imperfect Instruments

In this section, we derive a testable condition that, when met, can be used to determine the relative endogeneity of each imperfect instrument. We use this test to provide evidence that our preferred instrument, the IIV-All but (p_{it}^{S-s}) variable, is plausibly the least endogenous. Let $\tilde{x}$ denote the residuals from a regression of p_{it}^{s} on a set of year and MSA fixed effects. Let ρ_{vw} (resp. σ_{vw}) denote the correlation coefficient (resp. covariance) between two random variables v and w. Let σ_v denote the standard deviation of a random variable v. Define $z_1 \equiv p_{it}^{10}$, $z_2 \equiv p_{it}^{S}$, and $z_3 \equiv p_{it}^{S-s}$. Suppose that $\beta_{z_1}^{IV} > \beta_{z_2}^{IV} > \beta_{z_3}^{IV}$, as is the case for the majority of the effects we estimate. By equation (2.20) in the main text,

$$\begin{aligned} \beta_{z_1}^{IV} > \beta_{z_2}^{IV} > \beta_{z_3}^{IV} &\iff \\ \frac{\sigma_{z_1 u}}{\sigma_{\tilde{x} z_1}} > \frac{\sigma_{z_2 u}}{\sigma_{\tilde{x} z_2}} > \frac{\sigma_{z_3 u}}{\sigma_{\tilde{x} z_3}} &\iff \\ \rho_{z_1 u}\frac{\sigma_{z_1}}{\sigma_{\tilde{x} z_1}} > \rho_{z_2 u}\frac{\sigma_{z_2}}{\sigma_{\tilde{x} z_2}} > \rho_{z_3 u}\frac{\sigma_{z_3}}{\sigma_{\tilde{x} z_3}}. & \end{aligned} \tag{2.F.1}$$

Therefore, assuming that the instruments are indeed "imperfect instruments" in the sense that $\rho_{z_j u} \geq 0 \ \forall j$, a sufficient condition for $\rho_{z_1 u} > \rho_{z_2 u} > \rho_{z_3 u}$, that is, the instruments become progressively "less endogenous," is that

$$\frac{\sigma_{z_1}}{\sigma_{\tilde{x} z_1}} < \frac{\sigma_{z_2}}{\sigma_{\tilde{x} z_2}} < \frac{\sigma_{z_3}}{\sigma_{\tilde{x} z_3}}. \tag{2.F.2}$$

In our case, condition (2.F.1) generally holds, while (2.F.2) always holds (see Table 2.F.1). As a result, when considering these three imperfect instruments, it is plausible that z_1 (p_{it}^{10}) is the most endogenous, z_2 (p_{it}^{S}) falls in the middle, and z_3 (p_{it}^{S-s}) is the least endogenous.

Table 2.F.1 Sample Values for Condition (2.F.2), by Sector

	(1) $\frac{\hat{\sigma}_{z_1}}{\hat{\sigma}_{\tilde{x} z_1}}$	(2) $\frac{\hat{\sigma}_{z_2}}{\hat{\sigma}_{\tilde{x} z_2}}$	(3) $\frac{\hat{\sigma}_{z_3}}{\hat{\sigma}_{\tilde{x} z_3}}$	(4) $\frac{\hat{\sigma}_{z_4}}{\hat{\sigma}_{\tilde{x} z_4}}$
Food Service	223.2	300.8	432.0	460.7
Maintenance	160.7	196.9	246.4	279.8
Personal Service	298.8	335.3	417.4	529.2
Construction	166.4	211.2	288.1	331.3
Manufacturing	232.0	292.9	371.3	414.6
Transportation	246.4	272.2	327.4	386.1
Computers	380.0	412.3	534.7	622.2
Engineering	471.4	526.1	675.1	732.3
Science	553.5	754.2	1084.7	1251.5

Defining $z_4 \equiv p_{it}^{S-sPop}$, we also show in Table 2.F.1 that $\frac{\sigma_{z_3}}{\sigma_{\tilde{x} z_3}} < \frac{\sigma_{z_4}}{\sigma_{\tilde{x} z_4}}$ in all nine sectors. However, the use of z_4 improves upon z_3 in less than one half of the cases (that is, in the

majority of cases $\beta^{IV}_{z_3} < \beta^{IV}_{z_4}$). Because $\beta^{IV}_{z_3} < \beta^{IV}_{z_4} \iff \rho_{z_3 u} \frac{\sigma_{z_3}}{\sigma_{\tilde{x} z_3}} < \rho_{z_4 u} \frac{\sigma_{z_4}}{\sigma_{\tilde{x} z_4}}$, when $\frac{\sigma_{z_3}}{\sigma_{\tilde{x} z_3}} < \frac{\sigma_{z_4}}{\sigma_{\tilde{x} z_4}}$ the relative endogeneity between z_3 and z_4 cannot be determined. Therefore, we rely on the "IIV-All but" (p^{S-s}_{it}) instrument as our preferred instrument.