

Immigrant Earnings Assimilation in the United States: A Panel Analysis

Deborah Rho

Department of Economics
University of St. Thomas
deborah.rho@stthomas.edu

Seth Sanders

Department of Economics and
Sanford School of Public Policy
Duke University
seth.sanders@duke.edu

December 14, 2017 (current version)

August 20, 2018 (version pending census review)

Abstract

We construct the first long-term comparison of cross-sectional and panel estimates of immigrant earnings assimilation in the U.S. from a single data source. Unlike Lubotsky (2007), we find that selective outmigration of higher earning immigrants biases downwards cross-sectional estimates for all education groups. Cross-sectional estimates dramatically understate earnings growth for high-skilled foreign-born workers. The bias stems from both selective outmigration and selective employment; among high-skilled immigrants, low earners find employment with a substantial delay while high earners work immediately upon arrival. We present suggestive evidence that the H-1B visa program may play a role in estimated immigrant earnings dynamics.

We thank Peter Arcidiacono, Joseph Hotz, Kristin McCue, Marjorie McElroy, and participants at Duke Labor Lunch for helpful comments. Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. This research uses data from the Census Bureau's Longitudinal Employer-Household Dynamics Program, which was partially supported by the following NSF Grants SES-9978093, SES-0339191 and

ITR-0427889; NIA Grant AG018854; and grants from the Alfred P. Sloan Foundation. All errors and omissions are our own.

1. Introduction

The inflow of new immigrants into the United States has been a hotly debated policy issue for many decades. There is a great deal of work investigating the earnings differences between immigrant and native-born workers.¹ Most of these studies use U.S. Decennial Census data which has large cross-sectional samples of immigrants with a measure of year of entry into the U.S. and earnings.² Some inference can be made by following immigrants who migrated in a specific time period (e.g. 1975-1979) across census years. However, analyses relying on synthetic cohorts are limited by return migration changing the composition of a synthetic cohort over time. Additionally, cross-sectional studies on wages or earnings of immigrants rely on calculating statistics on the working population. What has been difficult to address, due to data limitations, is that selective labor force participation and nonrandom employment over time of natives and immigrants who remain in the U.S. can also bias cross-sectional estimates of earnings assimilation.

The nature of these biases is likely to be changing over time and the relative earnings patterns of more recent immigrant cohorts may differ from those found in previous cohorts for a number of reasons.³ First, the composition of immigrants has changed both in terms of countries of origin and skill level. For example, among immigrants entering the U.S. between 1965 and 1970, 28% had a college or advanced degree.⁴ Forty years later, the fraction of college-educated among new immigrants has grown to 38%. This reflects both the increase in the fraction of immigrants from countries that have traditionally sent highly educated workers and the increasing education levels within countries of origin. The rise in the share of college-educated immigrants is likely in part due to policies that have favored high-skilled immigration in more recent decades such as the L1 and H1-B visa programs. Next, these visa programs have affected selective employment of immigrants. The L1 and H1-B visa programs require employers to petition for visas for specific employees, meaning visa recipients arrive to the U.S. with a job in hand. This will lead to bias in cross-sectional estimates of

¹ Borjas (1994) and LaLonde and Topel (1997) survey the economic literature on immigration.

² Chiswick (1978) and Borjas (1985) are seminal papers that use these data.

³ See Borjas (2015) for a cross-sectional study of recent immigrant cohorts.

⁴ Following the literature, we refer to immigrants entering within 5 years of a census as “recent” immigrants.

earnings assimilation if lower earning immigrants entering through traditional channels take time to obtain employment. Finally, as immigrant sending countries have developed, the patterns of return migration may differ from earlier entry cohorts, potentially changing the direction of the bias due to selective outmigration in cross-sectional studies.

This paper examines the earnings dynamics of recent immigrants, specifically immigrants who arrived in the U.S. in 1995-1999. We examine three aspects of earnings dynamics of this recent cohort — (1) selective outmigration, (2) selective delay in labor force participation upon arrival and (3) differences across educational groups. We extend the insight of Lubotsky (2007) that longitudinal earnings data on immigrants can be assembled by matching administrative earnings data to standard demographic surveys. Lubotsky matches a 1990/91 cross-section from the Survey of Income and Program Participation (SIPP) and a 1994 cross-section of the Current Population Survey (CPS) to Social Security Earnings records from 1951 to three years after the survey year. These cross-sections record country of birth which identifies which matched administrative earnings records belong to the immigrants in the cross-sectional samples. There are two limits to this path breaking work. First, longitudinal analysis is possible only for immigrants who remain in the U.S. long enough to be sampled in the SIPP or CPS; second, immigrants identified through the SIPP or CPS cross-sections yield sample sizes sufficient for aggregate analysis but not for detailed subgroup analysis.⁵

Our data construction differs from Lubotsky's in two important ways. First, our cross-sectional dataset is very large, the 1-in-6 long form of the 2000 Decennial Census. While only a 5% sample from that file is publicly available, we use the Census Bureau's internal version that contains the full 16.67% of the population. This provides a large analysis file which is capable of analysis on important subpopulations. Second, we match this to population-level administrative earnings records contained in the Longitudinal Employer-Household Dynamics (LEHD) database. We observe immigrant earnings from the time of their arrival until 2008, the last year earnings data is available in our version of the LEHD data. We then further match these earnings histories to the full 2010

⁵ The 1994 CPS has contains 21, 296 natives and 1,643 immigrants of which only 203 were recent immigrants migrating within 5 years of the survey. Lubotsky pools SIPP and CPS data in order to obtain sufficient sample size for analysis by year of entry cohort.

population census to determine whether the immigrant remains in the U.S. in 2010. The result is that we are able to study both immigrants who remain and those who likely have left the U.S. in order to understand selection issues.⁶

To understand contemporary earnings dynamics of immigrants we argue that subgroup analysis is important. We find that cross-sectional earnings assimilation patterns are starkly different across education groups and this may rationalize changes in the average earnings patterns among recent immigrant cohorts and the differences across countries in assimilation patterns.⁷ Our ability to examine earnings dynamics by level of education is directly tied to our ability to construct very large samples of immigrants.

Consistent with studies of earlier entry cohorts, our cross-sectional results indicate that immigrants with less than a college education start at an earnings disadvantage but converge towards natives with time in the U.S. On the other hand, the cross-sectional analyses suggest that immigrants whose highest level of education is a bachelor's degree who arrive in the U.S. early in their labor market careers start at an earnings advantage, but experience a relative earnings decline with time. Unlike Lubotsky (2007), we find evidence of selective outmigration of higher earning immigrants for all education groups. While outmigration leads to a small downward bias in cross-sectional estimates of assimilation for low-skilled immigrants (those with less than a college degree), it results in substantial downward bias in estimates for high-skilled immigrants (those with a bachelor's degree or more). Controlling for selective outmigration, the observed cross-sectional decline in relative earnings goes to zero for immigrants whose highest level of education is a bachelor's degree.

While selective outmigration may account for the anomalous finding of "negative earnings assimilation" for high-skilled immigrants, the finding of no assimilation remains inconsistent with most notions of assimilation based on human capital theory. Another less studied source of bias in measuring assimilation is delayed labor market entry. Among college-educated immigrants, the most skilled immigrants may have been

⁶ Our earnings data include only individuals who have SSNs, which implies that our study is limited to legal immigrants. In the appendix, we present results based on reweighting the sample to reflect the characteristics of immigrants more generally as a robustness check.

⁷ Borjas (2015) finds that more recent immigrant cohorts experience much lower rates of wage convergence with natives than earlier cohorts. Chiswick and Miller (2011) find that immigrants from English-speaking developed countries have starkly different earnings patterns than other immigrants.

recruited to the U.S. with the promise of work and work immediately upon arrival; less skilled college-educated workers may enter in the more traditional way, searching for work once they arrive. For immigrants with exactly a bachelor's degree, we find that the increase in employment with time in the U.S. among lower earning immigrants is what makes it appear that immigrants experience zero relative earnings growth. The inclusion of an individual fixed effect shows an increase in relative earnings growth for immigrants in this education group. This is in part due to college-educated immigrants who work in industries that are highly represented in the H-1B visa program; they work immediately and continually upon arrival while other immigrants delay entry into the labor force. These results indicate that cross-sectional analyses of earnings assimilation may result in biased estimates not only due to selective outmigration but also because of selective participation in the labor market. While there is less evidence of bias for other education groups, we find evidence of substantial bias among individuals whose highest level of education is a bachelor's degree.

The paper proceeds as follows. Section 2 discusses the literature on immigrant assimilation. Section 3 describes the data and presents descriptive statistics on the sample. Section 4 discusses the empirical framework to measure earnings assimilation and Section 5 presents the results. Section 6 discusses auxiliary analyses that address how the results relate to findings using alternative data sources. Section 7 concludes.

2. Background

Many studies have established that immigrants earn less than native workers upon arrival but converge toward the native-born with time in the U.S. Most earlier work uses U.S. Decennial Census data which affords large cross-sectional samples of immigrants with consistent measurement of year of immigration and self-reported earnings. A second advantage is that, like most household surveys, the Census long form samples include workers whether or not they are documented, the latter group estimated to be about 30% of the foreign-born population in 2006.⁸ The typical approach is to define an entry cohort, for example immigrants entering between 1975 and 1979. The 1980 level of earnings is interpreted as the initial earnings in the U.S., the 1990 level of earnings is

⁸ Passel (2006)

interpreted as the level of earnings after 11-15 years, and the 2000 earnings as the level of earnings after 21-25 years in the country. Most studies find that initial earnings of immigrants are lower than similar native-born workers but the earnings gap diminishes with time (Borjas 1985; LaLonde and Topel 1992; Borjas 1995).

A recent exception to this finding is Borjas (2015) which finds that while earlier immigrant entry cohorts narrow the earnings gap by around 15 percentage points during the first two decades in the U.S., immigrants who entered the country after the 1980s have a negligible rate of wage convergence. Additionally, Chiswick and Miller (2011) finds evidence of “negative assimilation” for immigrants to the U.S. from English Speaking Developed Countries (ESDC). Chiswick and Miller find that immigrants from ESDC countries start at an earnings advantage relative to natives which diminishes with time in the U.S. This finding supports their model of international migration among countries in which immigrants’ skills are highly transferable. Immigration occurs when individuals experience a favorable wage draw in the potential destination relative to the wage available in the country of origin. A high wage offer that attracts immigrants need not persist indefinitely and a “regression to the mean” occurs without a deterioration of skills. Negative assimilation is found for ESDC immigrants in the same cross-sectional Census data in which relative earnings increases are found for immigrants born in other countries.

It is widely recognized that the synthetic cohort approach can be affected by both back-and-forth migration and permanent outmigration which change the composition of immigrants remaining to be measured at any census.⁹ For instance, rising relative earnings of foreign-born workers could reflect lower earning immigrants leaving the U.S., potentially biasing upward immigrant earnings growth. Additionally, bias may arise from differences in employment rates between immigrants and natives that becomes selective over time. Newly arrived foreign-born workers may not be accustomed to local job search practices but may learn about the labor market with time, increasing their chances of employment. Immigrants may also develop better social networks with

⁹ Borjas and Bratsberg (1996) estimate that 18 to 22 percent of legal immigrants who arrived in the U.S. in 1970-1980 left the country by 1980.

duration in the new country which can increase job prospects.¹⁰ Using the 2000 Decennial Census, Duncan and Trejo (2012) shows that while only about 73% of native high school dropouts are employed, 79% of similarly educated recent immigrants are working. On the other hand, recent college-educated immigrants are less likely to be employed than their native counterparts. In this single cross-section, immigrant employment rates increase with duration in the U.S. for all education levels. Studies using synthetic cohorts also find evidence of a relative increase in employment of immigrants with time.¹¹ We are aware of one panel study that examines employment of immigrants relative to natives in the U.S. context. Kaushal et al. (2015) uses the Survey of Income and Program Participation (SIPP) panel data that start in 1996, 2001, and 2004 and last 3-4 years, a relatively short panel, and finds that immigrants with a high school or lower education experience net employment growth over similar native workers with time in the U.S. Higher educated immigrants have the same employment trajectory as similar natives, suggesting that selection into the labor market varies by skill level. The employment of immigrants relative to natives can bias estimates of earnings assimilation in cross-sectional studies. It is difficult to know a priori the sign of the bias since it depends both on changes in the relative employment rates of immigrants and natives as they age and on how selection into and out of the U.S. labor force takes place for each group.

While the availability of appropriate data has been limiting, a number of studies have used longitudinal samples to consider earnings assimilation.¹² Lubotsky (2007), a groundbreaking study, uses earnings panel data from the Social Security Administration to investigate the nature of selective outmigration by comparing estimates of earnings convergence between workers who remained in the U.S. long enough to be measured in the 1990/91 SIPP and the 1994 Current Population Survey (CPS) to estimates from Decennial Census cross-sections. His panel estimates imply that during the first 20 years in the U.S., immigrants close the earnings gap by 10-15 percent which is about half as

¹⁰ Goel and Lang (2010); Munshi (2003)

¹¹ See Chiswick, Cohen, and Zach (1997), Funkhouser (2000), and Antecol, Kuhn, and Trejo (2006).

¹² Borjas (1989), Duleep and Dowhan (2002), and Hu (2000) are early studies using U.S. longitudinal data. Li (2003), Banerjee (2009), and Picot and Piraino (2013) use Canadian panel data to estimate immigrant earnings assimilation. Barth et al. (2012) and de Matos (2013) are examples of European panel studies.

fast as what is implied by synthetic cohort estimates, suggesting that lower earning foreign-born workers selectively leave the country.

Our analysis of selective outmigration and how it biases cross-sectional estimates is similar to Lubotsky's approach in that we examine assimilation estimates with and without conditioning on being in the U.S. in a particular year. Unlike Lubotsky (2007), our study compares cross-sectional and panel results from the same data source, which has the advantage of eliminating differences in estimates that stem from discrepancies in data collection methods. Additionally, immigrant status in Lubotsky's panel could only be determined if a worker was matched to the SIPP or CPS and therefore only if the worker was in the country long enough to be included in those surveys, which for some entry cohorts in the sample is decades. Our data allow us to directly examine earnings differences between immigrants with short and long stays. Immigrants in our sample, who arrived in 1995-1999, must have stayed in the U.S. for 1-5 years such that they are found in the 2000 Census. The data allow us to examine the earnings of immigrants who leave the country anytime after 2000, including the earnings of those with relatively short stays.

This paper also builds upon Lubotsky's in that we consider the labor market entry of immigrants in addition to outmigration. We investigate whether high earning immigrants are more likely to obtain employment upon arrival than their lower earning counterparts and consider how selective labor market entry biases estimates of assimilation by adding individual fixed effects to the panel model of earnings assimilation. We limit the analysis of labor market entry to workers who are still in the U.S. in 2010 to separately consider how selective employment affects estimates of earnings assimilation.

Kim (2012) and Kaushal et al. (2015) also include individual fixed effects in panel analyses of earnings assimilation and find that cross-sectional estimates overstate the relative earnings growth of foreign-born workers. While these studies observe individuals for two to four consecutive years, our quarterly panel spans a fourteen year period. Additionally, unlike most previous panel studies, we examine the relative earnings patterns of immigrants by education and find differences in assimilation across

skill levels.¹³ In addition to investigating the impact of selective outmigration on measures of assimilation, the focus of previous panel studies, as mentioned above, we also examine how selective employment affects these estimates.

3. Longitudinal Employer-Household Dynamics Data and Descriptive Statistics

3.1 Longitudinal Employer-Household Dynamics Data

The data comes from the 2000 Decennial Census 1-in-6 long form data which we combine with earnings panel data from the Longitudinal Employer-Household Dynamics (LEHD) database. The version of the LEHD we use contains quarterly earnings data from complete sets of unemployment insurance (UI) records for 47 states from the early nineties to 2008.¹⁴ We limit our analysis to the 27 states with earnings records that are available starting in 1995 or earlier.¹⁵ This allows us to observe earnings of immigrants who entered the U.S. in 1995-1999 from the start of their U.S. careers. About 81.6% of immigrants reside in one of these 27 states in the 2000 Decennial Census.¹⁶ Because the data is from UI records, earnings information is not available for the self-employed and for those who work in the informal sector. We exclude earnings observations in agriculture and public administration since coverage is incomplete.¹⁷ While the LEHD contains basic demographic information such as sex, age, and place of birth for all workers as well as the year of application for a Social Security number (SSN), education information is not available and we must rely on the 2000 Census for this measure. The data source of the variables used are reported in Appendix 1 Table A1.

Our sample consists of individuals in the 2000 Census who lived in one of the 27 states considered in 1995 and who are linked to LEHD earnings records for at least 2 quarters. For natives, we use the Decennial Census question “Where did you live 5 years ago?” The answer to this question indicates the state of residence of the individual in

¹³ Kaushal (2011) is an exception in that it focuses on the assimilation of college-educated immigrants with a science or engineering degree. Kaushal et al. (2015) considers those with a high school or lower education and those with more than a high school education separately.

¹⁴ Data from Connecticut, New Hampshire, Massachusetts, and Washington D.C. are not available.

¹⁵ The states in our analysis include: AK, AZ, CA, CO, FL, GA, HI, ID, IL, IN, KS, LA, MD, MN, MO, MT, NC, NM, NY, OR, PA, RI, SD, TX, WA, WY, and WI.

¹⁶ The sample includes male immigrants age 25 or older who are not residing in group quarters.

¹⁷ See Abowd et al. (2009) for a full description of the database.

1995 which we use to exclude those who lived outside of the 27 states considered. Since most of the immigrants in our sample were not in the U.S. at the start of the panel, we use the state in which they applied for an SSN as their initial state of residence. The proportion of immigrants in the Decennial Census who are successfully linked to LEHD earnings records is substantially lower than the proportion of natives. Undocumented immigrants are unlikely to be matched, as are those who have visas that do not authorize them to work, and those working in jobs not covered by UI. Andersson et al. (2014) find that a similar matched sample is reasonably representative of UI-covered employment.¹⁸ We discuss how this affects the composition of our immigrant sample below.

For this analysis we consider male immigrants who entered the U.S. in 1995-1999 and a random sample of native-born men of the same age who can be linked to LEHD earnings data.¹⁹ The sample is limited to individuals who are 25 years or older at the time of the 2000 Census and 65 years or younger at the end of our sample period. We only observe education in the Decennial Census and individuals younger than 25 are less likely to have completed their education.

The year of application for an SSN is taken as the year of immigration rather than the self-reported year of entry in the Decennial Census. The Decennial Census asks “When did this person come to live in the United States?” For those who have had multiple trips to the U.S., it is unclear how they would interpret this question. Table 1 gives the mean of the absolute value of the difference between the self-reported year of arrival and the year of applying for an SSN for immigrants in our sample. The mean absolute difference for immigrants taken as a whole is 2.72 years. For more than half of all immigrants, the self-reported year of immigration corresponds to the year in which they applied for an SSN. The discrepancy is primarily driven by Mexican immigrants and for about half of this group the difference is greater than 6.53 years. This may be due to back-and-forth migration. Ellis and Wright (1998) compare immigrants’ responses in the 1980 and 1990 Censuses to the question “When did this person arrive in the U.S. to

¹⁸ Starting from a sample drawn from LEHD earnings records in 2000, they find that the subset that also matches to 2000 Decennial data differs only modestly from the full sample (though their sample is limited to 31 metropolitan areas in 11 states and their immigrant sample is not restricted to individuals arriving between 1995 and 1999).

¹⁹ We take a 10% random sample of all native men who have less than a graduate degree and a 20% random sample of graduate degree holders. We only consider males since the labor force participation of women is more selective.

stay?” and “Where did this person live five years ago?” Over 27 percent of immigrants who reported that they arrived between 1985 and 1990 also reported that they lived in the U.S. on April 1, 1985 and this discrepancy is most common among Mexican immigrants. The Survey of Income and Program Participation (SIPP) and the Current Population Survey (CPS) also ask “When did this person come to the United States to stay?” Lubotsky (2007) matches administrative Social Security earnings records and the 1990/1991 SIPP and the 1994 March Supplement to the CPS and finds that fourteen percent of immigrants in the longitudinal data have earnings prior to the year of their self-reported arrival.

Table 1: Self-Reported Year of Arrival vs. SSN Application Year
Absolute Value of Difference

	n	Mean	SD	25 th percentile	Median	75 th percentile
All immigrants	80100	2.72	4.99	0.00	0.00	3.47
Mexican Immigrants	14800	7.09	5.55	3.00	6.53	10.00
Non-Mexican Immigrants	65300	1.80	4.33	0.00	0.00	1.00

Note. — SD=standard deviation. This table consists of all immigrants in the 2000 Decennial Census who fulfill our sample criteria and are matched to LEHD earnings in one of the 27 states considered. Sample sizes are rounded to the nearest hundred and the mean of observations falling within one percentile of the percentiles given are reported instead of the value at the percentile due to confidentiality requirements.

The earnings panel begins in 1995 and earnings of an individual who moves across states are included if they work in a UI covered job in one of the 27 states considered. Quarterly earnings are the total earnings from all firms in which the individual works during that quarter.²⁰ We observe labor force status in the 2000 Census but not in the LEHD. Although the LEHD has accurate earnings information for those working in UI covered work, when an individual does not work, we only observe a lack of earnings and have no information on labor force status or place of residence.

While our earnings panel ends in 2008, our measure of remaining in the U.S. occurs in 2010 – whether or not we can link an immigrant to the 2010 Decennial Census. About 88% of natives and 76% of immigrants are linked. Matching to the 2010 Census

²⁰ Since we are not considering all states and do not know the start date of a particular job, we exclude the first quarter of earnings we observed for each individual as this is less likely to be a full quarter of earnings than subsequent quarters.

indicates that an individual was in the U.S. but not matching may occur for a number of reasons. Less than complete Census coverage, mortality between when we last observed an individual and 2010, and emigration would all lead to our matched 2000 Census sample not matching to the 2010 Census. We suspect that issues of Census coverage play a limited role; while coverage rates for immigrants with a Social Security number are unreported, coverage rates by race and ethnicity are available. The Census Bureau reported that whites were overcounted by 0.8%, Asians were neither under nor overcounted, Hispanics were undercounted by 1.5% and Black Americans by 2.1%.²¹ As we show below, the nonmatch rates of immigrants are an order of magnitude larger than what would be expected by non-coverage alone if immigrant coverage rates are at all similar to the rates for all Asians and Hispanics.

Mortality is a factor in nonmatch rates. While there is controversy about whether healthier individuals migrate, whether sicker migrants return to their country of origins to die, or whether lifestyle choices are protective, the evidence that immigrant mortality is lower than native mortality despite lower socioeconomic status is voluminous.²² Figure 1 presents nonmatch rates by the final year a worker is observed in the LEHD earnings records. For natives, the probability of nonmatch diminishes only slightly with final year observed from 2000 to 2007.²³ This is what we would expect, assuming that mortality is relatively stable from year to year for the age group considered. Immigrants are much less likely than natives to be matched to the 2010 Decennial if they were not observed working in later years. For instance, 74% of immigrants with a graduate degree who are last observed working in 2000 are not matched to the 2010 Census while only 14% of their native counterparts are not matched. However, there is a much smaller difference in nonmatch rates between immigrants and natives who are seen working in 2008. While no strong correlation exists between final year of earnings and matching to the 2010 Census for natives, the nonmatch rates for immigrants decrease substantially as their final year of earnings approaches 2008. These patterns are consistent with immigrants leaving the U.S. before 2010 and we interpret the excess rate at which immigrant fail to match to the 2010 Census as a proxy for outmigration. We view natives who do not match to the

²¹ See https://www.census.gov/newsroom/releases/archives/2010_census/cb12-95.html

²² See Turra and Elo (2008), Turra and Goldman (2007), and Blue and Fenelon (2011).

²³ Note that 2008 is the final year of our earnings panel and therefore the last possible year to be observed.

2010 Census as largely made up of those who died prior to the Census; we view immigrants as a mixture of individuals who died or who migrated before 2010. If immigrant mortality rates are similar to native rates, then Figure 1 suggests that for immigrants not seen since the early 2000s, the vast majority likely left the U.S. By comparing statistics of natives vs. immigrants who do not match to the 2010, we can get a sense of the characteristics of selective outmigration as long as the characteristics of immigrants and natives who have died are similar.

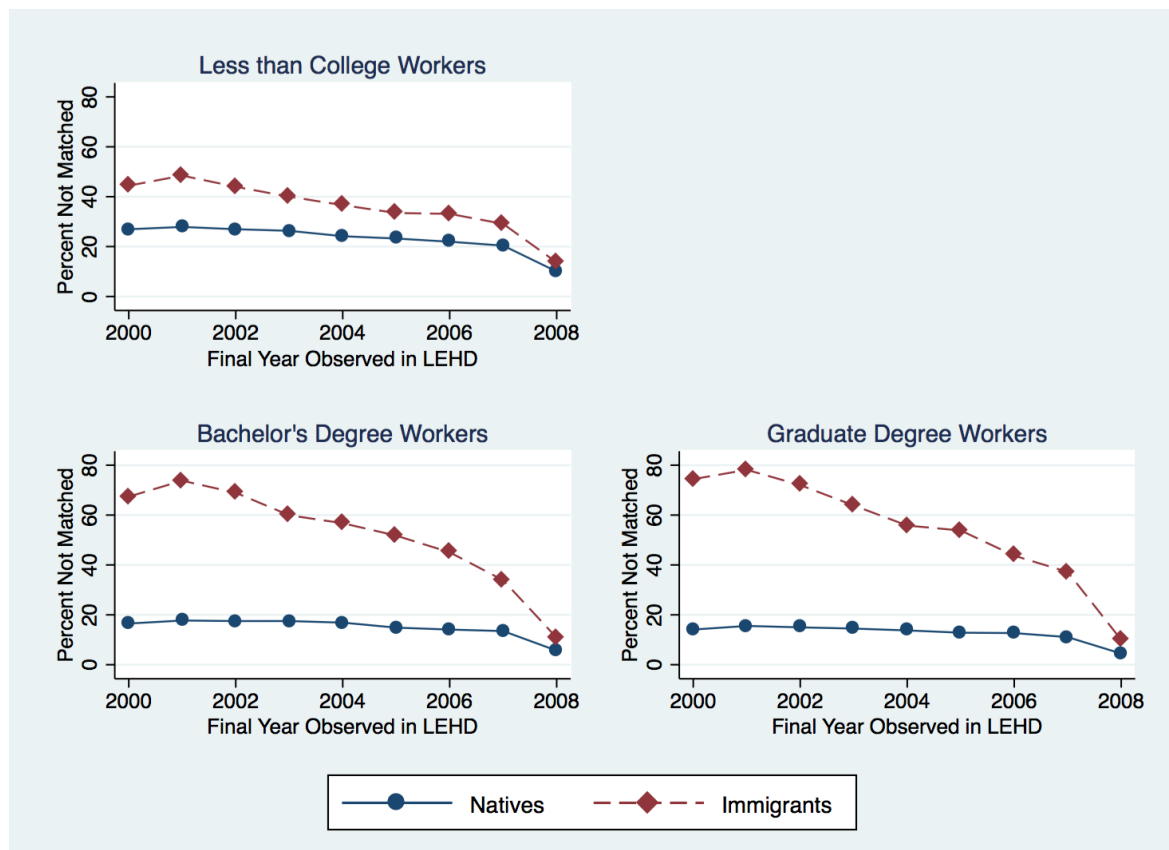


Figure 1: The percent of workers not matched to the 2010 Decennial Census by the final year observed in the LEHD earnings panel

In addition to considering the impact of outmigration, we examine employment in covered work for those who remain in the U.S. until 2010. As discussed above, there are differences in the employment rates of natives and immigrants and evidence of increasing employment for immigrants with time in the U.S. Given the limited evidence available on how assimilation affects immigrant-native differences in employment, we believe

these exercises provide valuable evidence on this topic, despite the fact that our data does not cover all types of jobs.

The earnings of a worker who moves across states are included in the panel if he works in one of the 27 states considered. Bartel (1989) finds that immigrants are more likely to move within the U.S. than similarly educated natives and those with higher education are more likely to move. If this is the case in our sample, we may be missing more immigrant earnings than native earnings due to their higher interstate mobility. However, we are likely capturing a majority of interstate moves as we are only missing moves to out-of-sample states. Of those who are matched to the 2010 Census, approximately 5% of natives and 7% of immigrants are observed in a state other than the 27 for which we have earnings data in 2010. Another reason for missing earnings is self-employment. Immigrants are more likely to be self-employed than natives and in the cross-section their likelihood for self-employment increases with time in the U.S. (Lofstrom 2002).

3.2 Descriptive Statistics

Columns 1 and 2 of Table 2 displays the education distribution of natives and immigrants in our sample as well as the place of birth distribution of immigrants from the most represented countries of origin.²⁴ Immigrants are much more likely than natives to be high school dropouts as well as graduate degree holders. Since immigrants in the LEHD are a subset of all U.S. immigrants, we compare the LEHD sample to 2000 Decennial Census data to assess how selection into UI covered jobs affects the sample of foreign-born workers.

²⁴ Those with less than a bachelor's degree are classified as high school graduates.

Table 2: Sample Characteristics

	LEHD Sample			2000 Decennial Public Use Sample	
	Natives	Immigrants		Immigrants	
		All	Excluding Mexican- born	All	Excluding Mexican- born
	(1)	(2)	(3)	(4)	(5)
n	434200	80100	65300	61675	40567
Education:					
High School Dropout	10.23	25.27	17.74	35.66	19.47
High School Graduate	60.15	34.4	35.11	31.79	35.13
Bachelor's Degree	19.88	20.67	24.07	16.78	23.32
Graduate Degree	9.75	19.66	23.07	15.77	22.08
Place of Birth:					
Mexico		17.46	-	32.48	-
India		8.74	10.59	8.03	11.89
Cuba		5.36	6.49	3.21	4.75
China		4.94	5.98	5.19	7.68
Former USSR		4.9	5.94	3.26	4.83
Philippines		3.5	4.24	2.45	3.63
Canada		3.2	3.88	2.26	3.35
UK		2.96	3.59	2.13	3.16
Former Yugoslavia		2.67	3.23	1.54	2.28
Vietnam		2.47	2.99	1.89	2.79

Note.—The samples depicted in Columns 1-4 consist of workers in the 2000 Decennial Census who fulfill our sample criteria and are matched to LEHD earnings in one of the 27 states considered. Columns 4 and 5 consist of male immigrants in the 2000 Decennial Public Use data who are age 25-57, live in one of the 27 states that are included in the LEHD sample, and report arriving in the U.S. in 1995-1999. Columns 3 and 5 exclude Mexican immigrants from Columns 2 and 4, respectively. Sample sizes for LEHD samples are rounded to the nearest hundred due to confidentiality requirements.

Column 4 consists of immigrants in the 2000 Census who are the same ages and live in the same states as those in the LEHD sample and self-report that they arrived in the U.S. in 1995-1999. About one third of this sample consists of Mexican immigrants. While Mexico is the largest source country in the LEHD sample, the proportion (17.46%) is much smaller than that found in the 2000 Census, which reflects the nature of our sample selection. As mentioned previously, the LEHD underrepresents undocumented workers and those who have visas that do not authorize them to work. It also excludes those who work in sectors not covered by UI. The underrepresentation of Mexicans in our sample likely reflects the overrepresentation of Mexican immigrants in the undocumented immigrant population, in work not covered by UI, and in agricultural

industries.²⁵ Any differences in characteristics associated with different year-of-arrival measures will also disproportionately affect Mexican immigrants, as documented in Table 1. In Columns 3 and 5 of Table 2, we exclude Mexican-born workers from the samples. The samples are remarkably similar in terms of place of birth and education conditional on the under-representation of Mexican immigrants.²⁶

In Table 3 we compare measures of earnings in the LEHD with those reported in the 2000 Census for our sample.²⁷ The immigrant-native earnings gap is calculated by subtracting native log earnings from that of immigrants. Comparing Columns 1 and 3, we see that self-reported 1999 earnings in the 2000 Census are substantially higher than the earnings found in the LEHD in the same year for all education levels. However, because this is the case for both immigrants and natives, the immigrant-native earnings gaps are similar between the two definitions of earnings. In Column 2, we report the log hourly wage of those with positive earnings in the Census.²⁸ Previous studies often use wage in considering earnings assimilation.²⁹ Comparing Columns 1 and 2 we see that the immigrant-native wage gap is substantially smaller than the gap in annual earnings. However, because we do not observe hours worked in the LEHD, we cannot consider wage. In Column 4 we present average log quarterly earnings from 1999 LEHD earnings records. The immigrant-native earnings gaps using quarterly earnings are closer to those using hourly wage so we use quarterly earnings to examine earnings assimilation.

²⁵ Passel (2006)

²⁶ In the appendix we reweight our LEHD sample to be representative of immigrants in the 2000 Decennial Census. We discuss this below.

²⁷ For comparability between the LEHD and the Census, we exclude immigrants who arrived in 1999 because their self-reported earnings may include earnings from their country of origin. We also exclude those who report working in agriculture in the Census.

²⁸ The wage rate is defined as annual earnings divided by hours worked.

²⁹ Lubotsky (2007) is an exception. The Social Security earnings records used in Lubotsky's study do not report hours worked so he uses log annual earnings.

Table 3: 1999 Earnings in the 2000 Census and LEHD

	<u>2000 Decennial Census</u>		<u>LEHD 1999 Earnings</u>	
	Annual Log Earnings	Log Hourly Wage	Annual Log Earnings	Average Log Quarterly Earnings
	(1)	(2)	(3)	(4)
Less than College				
Natives	10.2583	2.6807	10.0179	8.7706
n	253100	253100	236300	236300
Immigrants	9.8437	2.3858	9.6189	8.4242
% Gap	-41.46	-29.49	-39.90	-34.64
n	35500	35500	32900	32900
Bachelor's Degree				
Natives	10.7518	3.0848	10.5894	9.3096
n	69300	69300	62800	62800
Immigrants	10.4856	2.9579	10.3594	9.1401
% Gap	-26.62	-12.69	-23.00	-16.95
n	11800	11800	10500	10500
Graduate Degree				
Natives	11.0045	3.3135	10.8792	9.6106
n	66900	66900	59300	59300
Immigrants	10.5168	3.0031	10.3949	9.2183
% Gap	-48.77	-31.04	-48.43	-39.23
n	11600	11600	9500	9500

Note.—Earnings gaps are computed as the difference in log earnings between immigrants and natives. Immigrants who arrived in 1999 and all individuals who self-reported that they worked in agriculture in the 2000 Census are also excluded. Only those with positive earnings are considered. Sample sizes are rounded to the nearest hundred due to confidentiality requirements.

Table 4 presents average log quarterly earnings of workers in years 2000, 2004, and 2008. Natives are older than immigrants and have higher earnings in the raw data. Immigrants without a college education are at the greatest disadvantage in 2000 with earnings gaps of about 33%. While immigrants with a graduate degree are similarly disadvantaged in 2000, in 2008, their earnings gap diminishes to about 3%. For all education groups, the raw earnings disadvantage diminishes with time but is not eliminated. In our regression analysis, we examine whether these patterns represent the immigrant experience when controlling for selective outmigration and selective employment.

Table 4: Log Quarterly Earnings

	High School Dropouts and Graduates			Bachelor's Degree			Graduate Degree		
	2000	2004	2008	2000	2004	2008	2000	2004	2008
Natives									
Age	40.21	44.13	47.62	40.23	44.01	47.66	43.68	47.38	51.02
Log Earnings	8.8738	8.9083	8.9568	9.4134	9.4675	9.5317	9.6989	9.7423	9.7916
n	858300	772100	684800	232200	214500	196100	218400	201400	181700
Immigrants									
Age	35.97	39.92	43.86	35.53	39.37	43.30	35.43	39.09	43.02
Log Earnings	8.5405	8.6638	8.7613	9.2786	9.3259	9.4754	9.3833	9.5476	9.7647
n	140000	128000	116700	48800	41100	37400	43900	37600	34900
% Gap	-33.33	-24.45	-19.55	-13.48	-14.16	-5.63	-31.56	-19.47	-2.69

Note.—Earnings gaps are computed as the difference in log earnings between natives and immigrants. Sample sizes are rounded to the nearest hundred due to confidentiality requirements.

Next, Table 5 displays characteristics for those who are matched to the 2010 Decennial Census and those who are not. Approximately 88% of natives in our sample are found in the 2010 Census while only 76% of immigrants are matched. Interestingly, low-skilled natives are less likely to match than their college-educated counterparts while the reverse is true for immigrants. Additionally, immigrants who are not matched reported higher 1999 earnings than those who remained in the U.S. until 2010, suggesting that higher earning immigrants may be selectively leaving the country. In this study, we examine the degree to which selective outmigration biases estimates of earnings assimilation.

Table 5: Sample Characteristics- Match to the Decennial 2010

	Natives		Immigrants	
	Matched (1)	Not Matched (2)	Matched (3)	Not Matched (4)
n	383700	50500	61400	19000
% Sample	87.6	12.40	76.17	23.83
Education:				
High School Dropout	9.18	17.64	26.13	22.53
High School Graduate	59.58	64.16	35.87	29.7
Bachelor's Degree	20.87	12.85	19.76	23.58
Graduate Degree	10.37	5.36	18.24	24.18
Age in 2000:				
25-34	28.38	30.80	52.41	55.04
35-44	36.32	34.89	31.89	30.84
45 and older	35.30	34.31	15.70	14.12
Other Variables:				
1999 Self-reported earnings	\$44,126	\$31,342	\$29,251	\$38,310
Married	66.79	52.61	73.89	65.00
Disabled	14.78	23.59	20.32	17.39
Student	6.03	5.80	13.70	14.40
Speaks English Well	NA	NA	70.24	75.52
Citizen	NA	NA	6.20	5.75
Place of Birth:				
Mexico			18.36	14.61
India			9.23	7.16
Cuba			6.04	3.17
China			5.26	3.92
Former USSR			5.45	3.12
Philippines			3.92	2.15
Canada			2.75	4.65
Former UK			2.67	3.90
Yugoslavia			3.03	1.52
Vietnam			2.88	1.17

Note.— Sample sizes are rounded to the nearest hundred due to confidentiality requirements.

In addition to selective outmigration, nonrandom employment in the covered sector can also bias estimates of relative earnings growth. For those who are still in the U.S. until 2010, we present the distribution of the percentage of total time worked in Figure 2.³⁰ Natives are more likely than immigrants to work for 90% or more of quarters considered. In our analysis of selective employment below, we examine assimilation

³⁰ For natives, the total possible quarters worked excludes the first quarter of 1995. For immigrants, it excludes the quarters during the year of migration and those before migration.

patterns for workers who are more attached to the labor market by conditioning on those who work for at least 70% or more quarters, which is about 63-68% of natives and 60-64% of immigrants.

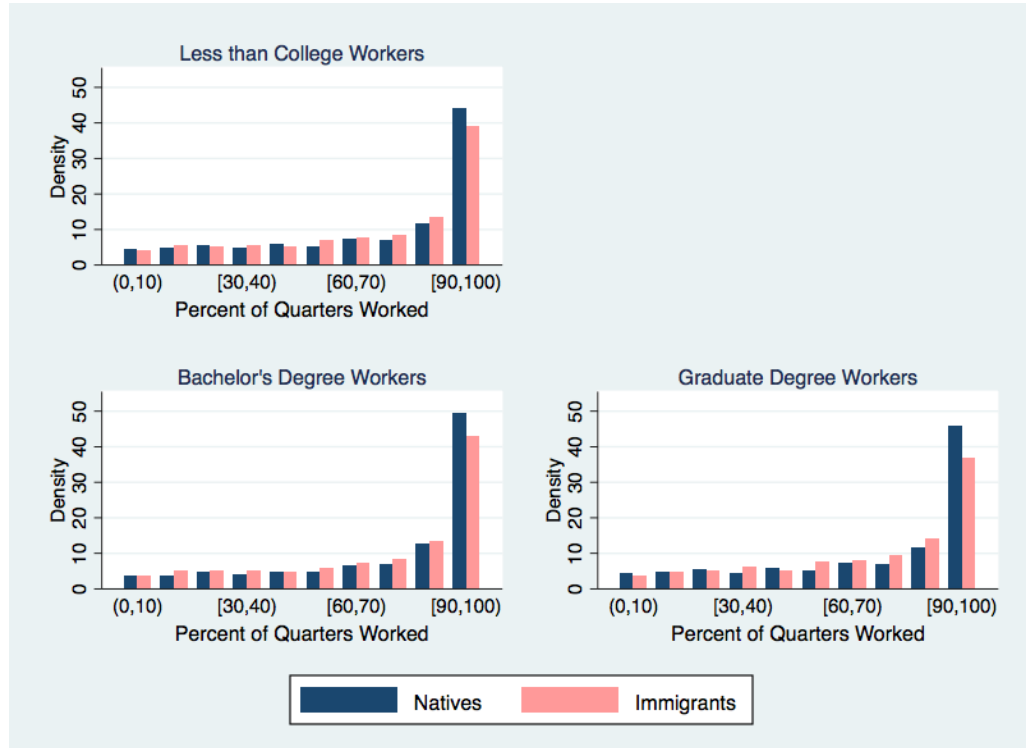


Figure 2: Distribution of Percent of Quarters Worked by Education Level and Immigrant Status

4. Earnings Assimilation Regressions

Before considering selection issues, we estimate earnings regressions in the spirit of those that are found in the immigrant earnings assimilation literature using synthetic cohorts. We then estimate the same regressions while limiting the sample to workers who are still in the U.S. in 2010 to examine the effect of selective outmigration on assimilation measures. To consider how selective employment biases estimates of immigrants' relative earnings growth, we further restrict the sample to individuals who are seen working in at least 70% of the possible quarters. Finally, we augment these regressions with person fixed effects to attain an estimate of earnings assimilation that is obtained from within person earnings growth.

For our cross-sectional analysis, log quarterly earnings of worker i in quarter t are given by:

$$w_{it} = \alpha_0 + \alpha_1 Age_{it} + \sum_{j=1}^6 \theta_j (Imm_i * AgeMig_i^j) + \sum_{k=1}^2 \beta_k (Imm_i * YSM_{it}^k) + \gamma_t + \epsilon_{it} \quad (1)$$

The earnings of natives are decomposed by Age_{it} , a quartic in age, and γ_t , year-quarter fixed effects. Immigrant earnings are further decomposed by the effect of age at migration and the effect of years since migration to the U.S. An indicator of immigrant status, Imm_i , is interacted with variables, $AgeMig_i^j$, where $j=1, 2, \dots, 6$, indicating whether an immigrant arrived at ages 20-25, 26-30, 31-35, 36-40, 41-45, or 46 and older.³¹ Natives are the omitted group. The variables YSM_{it}^k , where $k=1$ and 2 , are indicators for whether an immigrant has been in the U.S. for 6-10 or 11 or more years. The coefficient θ_j measures the initial immigrant-native earnings gap (in the first 5 years since migration) for the corresponding age at immigration cohort and β_k captures changes in the gap with years since migration.

Equation 1 is estimated using ordinary least squares separately for three groups — workers with less than a college degree (high school dropouts and high school graduates), those whose highest level of education is a bachelor's degree, and those with graduate degrees. We also include an indicator for a high school diploma in the regression for those with less than a college degree and an indicator for a professional/Ph.D. degree in the regression for the highest education group. Typically immigrant earnings assimilation equations such as equation 1 also include indicators for year of entry in the U.S. Coefficients on these indicators are taken as a measure of “cohort quality.” We are only considering those who immigrated in 1995-1999 and do not include such indicators. We estimate equation 1 for each education group on the full sample.

Then, to examine the role of selective outmigration, we estimate assimilation measures excluding those who are not matched to the 2010 Census. Differences in estimates from the full sample and those from workers who remain in the U.S., will

³¹ Most immigrant assimilation studies consider all immigrants who arrived in the U.S. at age 18 or older. Due to the age restrictions in the sample selection criteria all immigrants in our sample arrived at age 20 or older.

reveal the extent to which cross-sectional estimates of assimilation are biased by selective outmigration. Immigrants' relative earnings growth will be biased downward if lower earners selectively leave the country. On the other hand, if higher earners are more likely to return to their home countries, assimilation estimates will be biased upward.

Estimates of earnings assimilation may be biased even if we control for outmigration to the extent that participation in the covered sector in a given year is not random. In order to address concerns regarding selective employment, we further limit the sample of those who remain in the U.S. until 2010 to those who are also observed working for 70% or more of the total possible quarters. We estimate equation 1 for this sample. In only considering workers who are relatively attached to the labor market, we greatly reduce the role of changes in sample composition on estimates of assimilation.

Even with considering a more balanced panel of workers, who are relatively attached to the labor market, the timing of work may bias assimilation estimates. For instance, if higher earners are more likely to start working immediately while lower earners take longer to search for employment, it may appear that the relative earnings of immigrants is diminishing with years since migration despite improvements in relative earnings. To address this, we estimate a fixed effects model. With the inclusion of individual fixed effects, η_i , log quarterly earnings is given by:

$$w_{it} = \alpha_1^{FE} Age_{it} + \sum_{k=1}^2 \beta_k^{FE} (Imm_i * YSM_{it}^k) + \eta_i + \epsilon_{it}^{FE} \quad (2)$$

For those in the U.S. in 2010 who work for 70% or more of the sample period, we compare estimates of β_k in equation (1) to estimates of β_k^{FE} in equation (2) to consider whether there is a bias in estimates of assimilation that do not control for the timing of employment.³²

³² We will refer to estimates of equation (1) on the full sample as “cross-sectional estimates” and estimates of equation (1) on the subset of those who are matched to the 2010 Census as “panel estimates.” We will refer to estimates of equation (2) as “panel with fixed effects” estimates.

5 Results

5.1 Assimilation and Selective Outmigration

Columns 1-3 of Table 6 report regression results for the cross-sectional model. Immigrants without a college degree who arrived at ages 25 or younger start at an earnings advantage relative to natives. However, the initial relative earnings of immigrants decrease as age at migration increases and later entrants are at large earnings disadvantages compared to similar native workers. The cross-sectional results indicate that most low-skilled immigrants start at an earnings disadvantage which diminishes substantially with time in the U.S. Relative earnings improve by about 8.6 percentage points during the first 6-10 years in the U.S. for this group. This is a smaller earnings catchup than similar estimates found by Lubotsky (2007) which suggests a 13.2 percentage point catchup after 6-10 years in the country.

Table 6: Log Quarterly Earnings Regressions

	Cross-Sectional Estimates			Panel Estimates		
	High School Dropouts Graduates/ (1)	Bachelor's Degree (2)	Graduate Degree (3)	High School Dropouts Graduates/ (4)	Bachelor's Degree (5)	Graduate Degree (6)
Age at Migration:						
20-25	0.0745* (0.0021)	0.3052* (0.0040)	0.2815* (0.0047)	0.0611* (0.0022)	0.2647* (0.0043)	0.2842* (0.0049)
26-30	-0.0594* (0.0018)	0.0801* (0.0032)	-0.0255* (0.0035)	-0.0907* (0.0019)	0.0404* (0.0034)	-0.0417* (0.0037)
31-35	-0.2005* (0.0020)	-0.1433* (0.0038)	-0.2644* (0.0039)	-0.2331* (0.0021)	-0.2254* (0.0040)	-0.3029* (0.0042)
36-40	-0.2965* (0.0024)	-0.2499* (0.0048)	-0.4282* (0.0051)	-0.3295* (0.0025)	-0.3708* (0.0051)	-0.4928* (0.0054)
41-45	-0.3859* (0.0026)	-0.4235* (0.0058)	-0.5627* (0.0069)	-0.4284* (0.0028)	-0.5342* (0.0060)	-0.6449* (0.0073)
>=46	-0.4282* (0.0026)	-0.5146* (0.0059)	-0.5895* (0.0077)	-0.4717* (0.0027)	-0.6470* (0.0061)	-0.6951* (0.0080)
Year Since migration						
6-10 YSM	0.0861* (0.0017)	-0.0508* (0.0034)	0.0151* (0.0036)	0.1030* (0.0018)	0.0293* (0.0035)	0.0615* (0.0038)
11-13 YSM	0.0926* (0.0027)	-0.0949* (0.0055)	0.0282* (0.0058)	0.1104* (0.0028)	-0.0083 (0.0056)	0.0731* (0.0059)
Observations	12532300	3505200	3271100	11051900	3213500	3032900
R-squared	0.0618	0.0746	0.0733	0.0616	0.0775	0.0756

Note.— Robust standard errors in parentheses. Regressions control for year-quarter dummies and a quartic in age. Columns 3 and 6 also include an indicator for professional/Ph.D. degrees. Sample sizes are rounded to the nearest hundred due to confidentiality requirements. + $p < .10$ ** $p < .05$. * $p < .01$.

We find dramatic initial earnings advantages of high-skilled foreign-born workers in the cross-section. As shown in Columns 2 and 3 of Table 6, immigrants who entered the U.S. before the age of 25 whose highest level of education is a bachelor's (graduate) degree start with earnings advantages of about 30.5% (28.2%). As in the low-skilled sample, the initial relative earnings of immigrants decrease as age at migration increases. However, college-educated immigrants do not appear to experience a substantial increase in relative earnings throughout the first 10 years in the U.S. In fact, immigrants with only a bachelor's degree, appear to experience slower earnings growth leading to an *overall decrease in relative earnings* of about 9.5 percentage points after more than a decade since migration. While immigrants with graduate degrees also do not experience

substantial improvements, their relative earnings are about 3 percentage points higher after 11-13 years than when they first arrived.³³

The estimates of assimilation in Columns 1-3 of Table 6 do not control for selective outmigration. Columns 4-6 of Table 6 present assimilation estimates for the sample of workers who are still in the U.S. in 2010, our panel estimates. Comparing these estimates to those in Columns 1-3, we see that for all education groups, immigrants who stayed in the U.S. for 11-15 years have lower initial relative earnings than those who left the country prior to 2010. Conditioning on staying in the U.S. increases measures of assimilation for all education groups. These panel estimates of assimilation reveal slightly faster relative earnings growth of low-skilled immigrants. While the cross-sectional analysis suggests an earnings catch-up of about 9.3 percentage points, the panel results indicate an 11.0 percentage point increase in relative earnings.

The difference between the cross-sectional results and the panel estimates for highly-educated immigrants is more dramatic. For those whose highest level of education is a bachelor's degree, limiting the analysis to workers who are in the U.S. in 2010, changes the considerable decline in relative earnings of immigrants to no change in relative earnings. Compared to the cross-sectional results, the panel estimates suggest a larger relative earnings catch-up for immigrants with a graduate degree of about 7.3 percentage points.

These results are consistent with the outmigration of high earning immigrants. As discussed above, 12.4% of natives in our sample are not matched while 23.8% of immigrants are not found in the 2010 Census and the difference between the native and immigrant match rates are likely due to the outmigration of foreign-born workers. Here, we more closely examine earnings and the probability of outmigration. We consider the probability of not being observed in the 2010 Census and self-reported 1999 earnings in the 2000 Census. The probability of matching to the 2010 Census is estimated with a probit model in which the match probability is a function of indicators for the decile of

³³ As we have previously discussed, the LEHD is not representative of the 2000 Census immigrant population. A reasonable question is whether the results reported here would hold for a representative sample of immigrants. To address this, in Appendix 2 we reweight a modified LEHD sample to be representative of immigrants in the 2000 Census and the results are substantively the same.

self-reported earnings in the 2000 Census and birth year indicators. The probit model is estimated separately for natives and immigrants for each education group.³⁴

Figure 3 displays the predicted probability of not matching to the 2010 Census across deciles of self-reported 1999 earnings for the 1960 birth cohort. For all education groups, the probability of nonmatch decreases with decile for native workers. However, this is not the case for immigrants. For low-skilled immigrants, while the probability of nonmatch decreases with decile of earnings at low deciles, the probability of nonmatch increases starting with the seventh decile. For high-skilled immigrants, the probability of nonmatch generally increases with self-reported 1999 earnings. For immigrants with exactly a bachelor's degree, those in the highest decile of earnings are 2.4 times more likely to leave the country than those in the fifth decile. Selective outmigration of high earning immigrants leads to a downward bias in our cross-sectional estimates of earnings assimilation.³⁵

³⁴ Deciles are calculated using the entire education group, including both natives and immigrants.

³⁵ Similar patterns are found using 1999 LEHD earnings as shown in Figure A1 in Appendix 3.

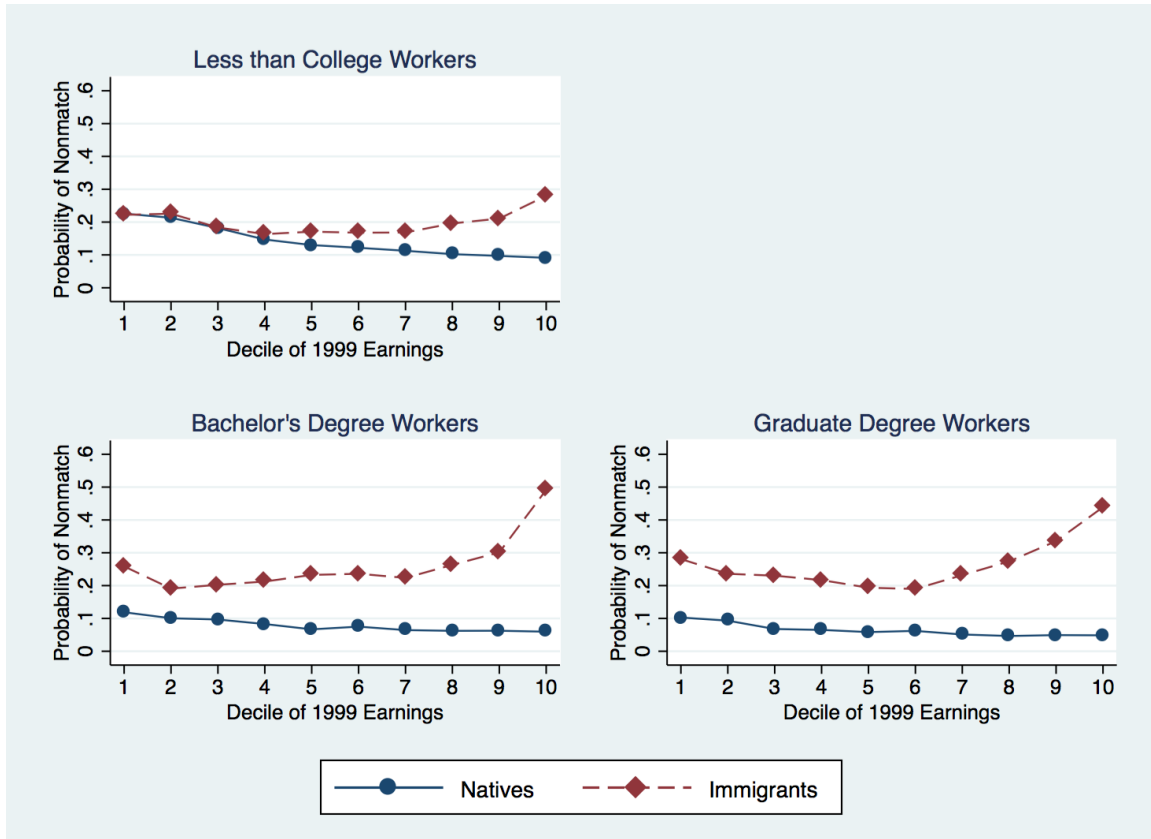


Figure 3: Predicted Probability of Not Matching to the 2010 Census by Decile of the Self-Reported 1999 Earnings Distribution
Note.—Estimates are for the 1960 birth cohort.

These results stand in contrast to the conclusion in Lubotsky (2007) that the cross-sectional analysis overestimates the earnings catchup of foreign-born workers. Lubotsky's panel estimates suggest that the earnings gap closes by half as fast as what is suggested by his cross-sectional analysis. We find the reverse for all education groups; the panel results suggest faster relative earnings growth than the cross-sectional estimates. The difference in our results from that of Lubotsky's study may be due to changes in immigrant flows in recent years.

For instance, the L-1 visa program was established in the 1970 amendments to the Immigration and Nationality Act.³⁶ The L-1A visa allows U.S. employers to transfer an executive or manager from one of its foreign offices to one of its offices in the U.S. and also enables foreign companies which do not have a U.S. office to send a manager to the

³⁶ See Implementation of L-1 Visa Regulations of Department of Homeland Security Office of Inspector General

U.S. with the purpose of establishing one.³⁷ A visa recipient is allowed an initial stay of 1 to 3 years and may request for extensions of stay until the employee has reached the maximum limit of seven years. The L-1B visa similarly, allows U.S. employers to transfer a professional employee with specialized knowledge from one of its foreign offices to one of its offices in the U.S. and also enables foreign companies which do not have a U.S. office to send a professional employee with specialized knowledge to the U.S. with the purpose of establishing one.³⁸ The number of L1 visas has increased in more recent years. In 1985 (1990), 65,349 (63,180) foreign-born workers were admitted with an L1 visa. In 1995, 112,124 workers were admitted under this visa and the number has increased every year for the rest of the decade, reaching 234,443 in 1999. If the outmigration patterns of foreign-born managers and other immigrants with specialized knowledge differ from other immigrant workers, the increase in these visas may lead to different immigrant flows in more recent entry cohorts.

While we do not observe visa status in our data, we consider those who are working as managers in our sample of college-educated immigrants. We observe that of the 13.1% of the sample that self-report their occupation as an executive or manager in the 2000 Census, 42.1% are not found in the 2010 Census while those who report other occupations in 2000 are not found at a rate of 26.1%. Immigrant managers on average, earn over twice as much as their non-manager counterparts and for both managers and non-managers, those who are not found in the 2010 Census have higher earnings than those who are definitively in the U.S.³⁹ The high propensity of foreign-born executives and managers to leave the U.S. may in part explain the change in selective outmigration patterns of more recent immigrants.

³⁷ See <https://www.uscis.gov/working-united-states/temporary-workers/l-1a-intracompany-transferee-executive-or-manager>

³⁸ See <https://www.uscis.gov/working-united-states/temporary-workers/l-1b-intracompany-transferee-specialized-knowledge>

³⁹ In contrast, of the 21.0% of college-educated natives who report their occupation as an executive or manager in the 2000 Census, 6.7% are not found in the 2010 Census while those who are not managers in 2000 are not found at a rate of 7.9%. Additionally, those who are not found in the 2010 Census have lower earnings than those who are matched.

5.2 Assimilation and Employment

In this section, we consider workers who are in the U.S. in 2010 and those who are working at least 70% of the quarters considered. Columns 1-3 of Table 7 display estimates of equation 1 for this sample. Comparing these estimates to those in Columns 4-6 of Table 6, we see that limiting the analysis to workers who are relatively attached to the labor force does not generally change the patterns of assimilation. For low-skilled workers, analysis of the more balanced panel leads to lower estimates of initial relative earnings of immigrants and higher estimates of relative earnings growth. There are no meaningful differences between estimates from the more balanced panel and the entire sample of those matched to the 2010 Census for college-educated workers.

Table 7: Log Quarterly Earnings Regressions-Matched to 2010 Census and Worked 70% of quarters

	Panel Estimates			Panel with Fixed Effects Estimates		
	High School Dropouts Graduates/ (1)	Bachelor's Degree (2)	Graduate Degree (3)	High School Dropouts Graduates/ (4)	Bachelor's Degree (5)	Graduate Degree (6)
Age at Migration:						
20-25	-0.0247* (0.0022)	0.2724* (0.0042)	0.2972* (0.0051)			
26-30	-0.1749* (0.0019)	0.0440* (0.0034)	-0.0388* (0.0038)			
31-35	-0.2946* (0.0021)	-0.2000* (0.0041)	-0.2926* (0.0043)			
36-40	-0.3836* (0.0025)	-0.3524* (0.0051)	-0.4677* (0.0053)			
41-45	-0.4872* (0.0028)	-0.5217* (0.0061)	-0.6205* (0.0074)			
≥46	-0.5519* (0.0027)	-0.6831* (0.0062)	-0.6888* (0.0082)			
Year Since Migration						
6-10 YSM	0.1160* (0.0018)	0.0388* (0.0035)	0.0627* (0.0039)	0.1006* (0.0029)	0.0239* (0.0051)	0.0432* (0.0061)
11-13 YSM	0.1308* (0.0028)	-0.0009 (0.0058)	0.0843* (0.0061)	0.1392* (0.0048)	0.0468* (0.0088)	0.0658* (0.0101)
Person fixed effects	No	No	No	Yes	Yes	Yes
Observations	8914900	2691900	2458900	8914900	2691900	2458900
R-squared	0.0702	0.091	0.0896	0.0463	0.1242	0.0938

Note.— Robust standard errors in parentheses. Regressions control for year-quarter dummies. Columns 1-3 also control for a quartic in age. Columns 3 also include an indicator for professional/Ph.D. degrees. Sample sizes are rounded to the nearest hundred due to confidentiality requirements.
+ p < .10. **p < .05. *p < .01.

Columns 4-6 of Table 7 present estimates from equation 2 which includes individual fixed effects. Controlling for individual heterogeneity does not substantially change estimates of assimilation for those without a college degree and those with graduate degrees. For immigrants whose highest level of education is a bachelor's degree, cross-sectional estimates suggest no relative earnings growth beyond 10 years in the U.S. However, including individual fixed effects reveals improvements in relative earnings of about 4.7 percentage points, suggesting that selective employment may bias estimates of assimilation for these workers. This finding stands in contrast to other studies that compare cross-sectional estimates of assimilation with estimates from fixed

effect models. Both Kim (2012) and Kaushal et al. (2015) also include individual fixed effects in panel analyses of earnings assimilation and find that cross-sectional estimates overstate the relative earnings growth of foreign-born workers.

Panel estimates of assimilation may be biased if not all immigrants work immediately and continuously upon arrival and if the timing of employment is correlated with earnings. To consider this, we categorize immigrants into three groups within education level. Group 1 is made up of immigrants who work immediately upon arrival and work in every quarter of the sample period after migration. Group 2 consists of those who are not in Group 1 but work for at least 85% of quarters after arrival and Group 3 includes all other workers.⁴⁰ For Groups 2 and 3, we consider whether an immigrant is more or less likely to work with time in the U.S. We estimate a probit model in which the dependent variable is an indicator for whether an immigrant is working in a given quarter and the independent variables are a constant and indicators for whether an immigrant has been in the U.S. for 6-10 or 11 or more years along with a quartic in age and year-quarter fixed effects. We report marginal effects in Table 8. The coefficients on the year since migration indicators convey how much more likely an immigrant is to work at later points in his U.S. career than in the first five years. For all education groups, immigrants are more likely to be working relative to similar natives with time in the U.S. This is consistent with the finding by Kaushal et al. (2015) for workers with a high school or lower education. However, Kaushal et al. finds that higher educated immigrants have the same employment trajectory as similar natives. They consider immigrants who arrived in the U.S. prior to 1970 to those who entered after 1990 while our study focuses on immigrants who arrived in 1995-1999.

⁴⁰ As mentioned above, we are only considering those who work for at least 70% of the sample period.

Table 8: Marginal effects from Probit Model

Dependent Variable: Indicator for Employed in Quarter

	High School Dropouts/Graduates		Bachelor's Degree		Graduate Degree	
	Percent of Total Quarters Worked		Percent of Total Quarters Worked		Percent of Total Quarters Worked	
	85% or more (1)	Less than 85% (2)	85% or more (3)	Less than 85% (4)	85% or more (5)	Less than 85% (6)
Year Since Migration						
6-10 YSM	0.0082* (0.0015)	0.0288* (0.0037)	0.0042* (0.0026)	0.0259* (0.0062)	0.0175* (0.0031)	0.0604* (0.0062)
11+ YSM	0.0129* (0.0021)	0.0528* (0.0049)	0.0125* (0.0036)	0.0447* (0.0086)	0.0336* (0.0038)	0.1164* (0.0078)
Observations	550300	247600	165300	75300	137700	78300
Pseudo R-squared	0.0640	0.0393	0.0754	0.0608	0.1181	0.1124

Note.— Regressions also include a constant, a quartic in age and year-quarter fixed effects. Samples used for Columns 1, 3, and 5 include immigrants who worked for at least 85% of quarters but not all quarters. The sample of natives remains constant within education group and does not depend on the percent of quarters worked. Sample sizes are rounded to the nearest hundred due to confidentiality requirements.

*p < .01.

The way in which delayed labor force participation of immigrants biases estimates of assimilation depends on differences in earnings levels and earnings growth between continuous workers and those who are less attached to the labor market among immigrants. For instance, if continuous workers have higher earnings than those who are less attached but both types of workers have similar earnings growth, cross-sectional estimates of earnings assimilation will be biased downward. The bias will be greater if less attached workers earn less and experience lower earnings growth than continuous workers. To understand the change in assimilation estimates between the regressions with and without fixed effects in Table 7, we examine panel estimates of assimilation by the groups we described above. We present these results in Table A5 in Appendix 4. For all education groups, immigrants who are the least attached to the labor market, those in Group 3, have the lowest initial earnings. Among workers with less than a college education, the differences in relative earnings growth across the three groups are not substantial. This corresponds to a small downward bias in the panel estimates of assimilation, indicated by the small increase in the coefficient on the indicator for 11 or

more years since migration going from the panel estimate to the panel with fixed effects estimate for this education group in Table 7. However, for workers whose highest level of education is a bachelor's degree, we see that those who are least attached to the labor market experience zero relative earnings growth while those who are more attached to the labor market have positive relative earnings growth with time in the U.S. This leads to the more substantial downward bias in estimates of assimilation that do not control for individual fixed effects. For those with a graduate degree, Table A5 in Appendix 4 shows that the least attached workers experience the greatest relative earnings growth. In this case, panel estimates of earnings assimilation may be biased either upward or downward. In our sample, there is a small upward bias in estimates without individual fixed effects.

To further examine the role of selective employment on estimates of earnings assimilation of college-educated immigrants, we consider the role of the H-1B visa program. First established in the Immigration and Nationality Act in 1965, the H-1B visa, allows firms to temporarily employ college-educated foreign workers in “specialty” occupations. Firms submit petitions on behalf of workers with the U.S. Department of Labor. Initially granted for up to three years, the visa can be renewed once for a total possible period of employment of six years and H-1B visa workers can apply for legal permanent status. In 2000, about 41% of H-1B beneficiaries had at least a Masters degree and 49% were from India. About 58% were working in computer related occupations and 12% were working in architecture and engineering.⁴¹ The Immigration Act of 1990 increased the cap of 65,000 H-1B visas per fiscal year. The cap was binding in 1997-2000 and was increased to 115,000 for fiscal years 1999 and 2000 and then to 195,000 for years 2001-2003. Borjas and Friedberg (2009) make a conservative estimate of 400,000 visas and state that given the 7.6 million people in the 2000 Census who report to have arrived in the U.S. in the previous five years, H-1B visa workers make up about 5% of the 1995-1999 immigrant entry cohort.

Similar to Borjas and Friedberg (2009), we approximate the set of H-1B visa recipients by considering immigrants whose first job is at a firm that is classified as providing Computer Related Services since the LEHD does not include information on visa status. We refer to these immigrants as H-1B visa recipients. Here we consider our

⁴¹ 2000 Statistical Yearbook of the Immigration and Naturalization Service

sample of workers who are seen in the 2010 Census and who work for at least 70% of the sample period. Classifying our sample based on this approximation of H-1B status, about 21% of immigrants whose highest level of education is a bachelor's degree and about 20% of immigrants with graduate degrees are H-1B recipients. Among immigrants who have only a bachelor's degree, about 27% of those without an H-1B visa are in Group 1 (i.e., work immediately upon arrival and work in every quarter of the sample period after migration) while 38% of H-1B holders work immediately and continuously.⁴² For this education group, 23% of immigrants without an H-1B visa are in Group 3 (i.e., work for less than 85% of all quarters after migration) while only 16% of H-1B holders are in this group. As discussed above, immigrants in Group 3 do not experience any earnings catch-up after 10 years in the U.S. This suggests that while H-1B recipients work upon arrival, other immigrants take time to find a first job and do not experience substantial relative earnings gains with U.S. work experience.

6. Auxiliary Analyses

For most of the education groups we consider, the cross-sectional earnings regression results resemble assimilation patterns in the literature. While past work has typically found faster earnings growth of immigrants, our cross-sectional results suggest a decline in relative earnings of foreign-born workers whose highest level of education is a bachelor's degree. Here, we investigate whether our unique cross-sectional findings are due to selection into the LEHD. We use the synthetic cohorts approach similar to that of Borjas (1995) using Decennial Census data and estimate assimilation rates by education.

We consider immigrant entry cohorts 1980-2011 and use the 1990 and 2000 Decennial Census 5% samples and the 2007-2011 American Community Survey. We restrict the analysis to males aged 25-64 who work in the civilian sector, are not self-employed, and do not reside in group quarters. An immigrant is defined as anyone who was born outside of the United States. We exclude those born in U.S. territories as well as those who immigrated before the age of 18. The wage rate is defined as annual

⁴² H-1B recipients work immediately but not necessarily for every quarter. For the full sample of immigrants whose highest level of education is a bachelor's degree (which does not condition on being in the U.S. in 2010 and on working for at least 70% of the sample period), about 88% of H-1B recipients work within the first year of arrival.

earnings divided by hours worked in the previous calendar year. Earnings are in 2000 dollars. As in Borjas (1995) workers who reported an hourly wage rate below \$1 and over \$250 are omitted from the analysis.

Although many previous studies have controlled for years of schooling in estimates of earnings assimilation, these papers have focused on the change in quality of immigrant entry cohorts rather than differences in the return to years since migration by education and have not allowed for the relative earnings growth of immigrants to differ by education level. We estimate log wage equations by education and separately consider those without a college degree (high school dropouts and high school graduates), those whose highest degree is a bachelor's, and those with a graduate degree.⁴³

Log wage for individual i is given by:

$$w_i = \alpha_0 + \alpha_1 Year_t + \alpha_2 Age_i + \alpha_3 (Age_i * Year_t) + \alpha_4 Imm_i + \alpha_5 (C_i * Imm_i) + \alpha_6 (YSM_i * Imm_i) + \alpha_7 (AgeMig_i * Imm_i) + \epsilon_i \quad (3)$$

where $Year_t$ is a vector of indicators for the year of the survey and Age_i is a cubic in the age of the worker at the time of the survey. Differences between natives and immigrants are captured by Imm_i , an indicator of foreign-born status, C_i , a vector of indicators for the entry cohort of immigrants, YSM_i , a cubic in the years since migration of immigrants, and $AgeMig_i$, age at migration for the foreign-born.⁴⁴

We report regression results in Table A6 in Appendix 5 and focus here on the age-earnings profile of immigrants relative to natives implied by the regression results. Figure 4 displays the predicted relative age-earnings profiles of immigrants who arrived at the age of 25. With the exception of workers whose highest degree is a bachelor's degree, foreign-born workers start at an earnings deficit and subsequently experience faster earnings growth than similarly educated natives. Immigrants with a bachelor's degree who entered the U.S. in the late 1990s on the other hand, have initial earnings that are higher than natives. Figure 4 suggests that this group of immigrants experience a decrease in relative earnings throughout the first 10 years in the U.S., much like the

⁴³ Those with less than four years of college are classified as high school graduates.

⁴⁴ C_i is a vector of indicators for arriving in years 1985-1989, 1990-1994, 1995-1999, 2000-2004, and 2005-2011. The 1980-1984 entry cohort is the omitted category.

cross-sectional results from the LEHD sample. While the differences between these synthetic cohort analyses and our LEHD panel study prohibit detailed comparisons, the decline in the relative earnings of immigrants whose highest level of education is a bachelor's degree in both cross-sectional analyses suggests that our seemingly unusual cross-sectional findings for this education group are not likely due to selection into the LEHD data.

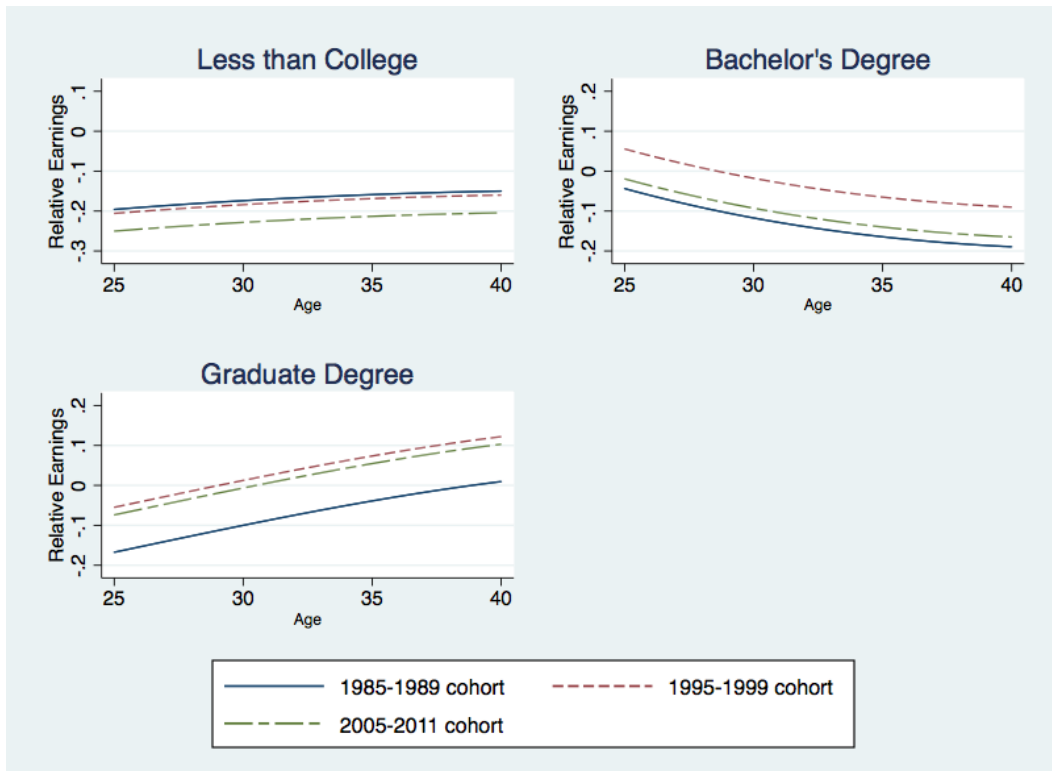


Figure 4: Public Use Data Analysis Results-Immigrant Relative Earnings Profiles.

Note- Predictions implied by results in Table A6 in Appendix 5

7. Conclusion

This paper is among the first to directly compare cross-sectional and panel analyses of immigrant earnings assimilation in the U.S. We consider recent male immigrants and find faster earnings growth in the cross-section for foreign-born workers with less than a college education. However, the cross-sectional analysis suggests that immigrants whose highest level of education is a bachelor's degree experience a relative earnings decline with time in the U.S. For all education groups we find evidence of

selective outmigration of higher earning foreign-born workers. This leads to a substantial downward bias in cross-sectional estimates of assimilation for high-skilled immigrants.

Despite the fact that previous works have emphasized the role of selective outmigration of immigrants, this paper suggests that the delayed labor market entry of immigrants could also result in biased estimates of relative earnings growth. This source of bias appears to especially pose a threat to accurate measures of assimilation of college-educated foreign-born workers. The relative earnings of immigrants whose highest level of education is a bachelor's degree diminish in the cross-section. However, the inclusion of individual fixed effects reverses the decline. This is not a result of selective attrition from the panel but due to lower earning immigrants becoming more likely to work with time in the U.S. Controlling for individual heterogeneity, we find that the earnings of college-educated immigrants also improve in relation to similar natives. The nature of selection into employment varies across education groups and results in different degrees of bias in our cross-sectional estimates of immigrant earnings assimilation, suggesting the importance of a panel study.

We find dramatic differences in estimates of relative earnings growth of immigrants between the cross-sectional and fixed effect analyses for college-educated immigrants who arrived in the late 1990s. Further research is necessary to examine whether delayed labor market entry and nonrandom employment also contribute to biased comparisons of earnings between other immigrant entry cohorts and their native counterparts. Such research would greatly contribute to a fuller understanding of the immigrant labor market experience in the U.S.

References

- Abowd, John M., Lars Vilhuber, Bryce E. Stephens, Fredrik Andersson, Kevin McKinney, Marc Roemer, and Simon Woodcock. 2009. The LEHD infrastructure files and the creation of the quarterly workforce indicators. In *Producer dynamics: New evidence from microdata*, eds. Timothy Dunne, J. Bradford Jensen, and Mark J. Roberts. University of Chicago Press, 149-230.
- Andersson, Fredrik, Monica Garcia-Perez, John C. Haltiwanger, Kristin McCue, and Seth Sanders. 2014. Workplace concentration of immigrants. *Demography* 51:2281-2306.
- Antecol, Heather, Peter Kuhn, and Stephen J. Trejo. 2006. Assimilation via prices or quantities? Sources of immigrant earnings grown in Australia, Canada, and the United States. *Journal of Human Resources* 41, no. 4:821-840.
- Banerjee, Rupa. 2009. Income growth of new immigrants in Canada: Evidence from the Survey of Labour and Income Dynamics. *Industrial Relations* 64, no. 3:466-488.
- Bartel, Ann P. 1989. Where do the new U.S. immigrants live? *Journal of Labor Economics* 7, no. 4:371-391.
- Barth, Erling, Bernt Bratsberg, and Oddbjørn Raaum. 2012. Immigrant wage profiles within and between establishments. *Labour Economics* 19, 541-556.
- Blue, Laura and Andrew Fenelon. 2011. Explaining low mortality among US immigrants relative to native-born Americans: the role of smoking. *International Journal of Epidemiology* 40, no. 3:786-793.
- Borjas, George J. 1985. Assimilation, changes in cohort quality, and the earnings of immigrants. *Journal of Labor Economics* 3, no. 4:463-489.
- Borjas, George J. 1989. Immigrant and emigrant earnings: A longitudinal study. *Economic Inquiry* 27, no. 1:21-37.
- Borjas, George J. 1994. The economics of immigration. *Journal of Economic Literature* 32, no. 4:1667-1717.
- Borjas, George J. 1995. Assimilation and the changes in cohort quality revisited: What happened to immigrant earnings in the 1980s? *Journal of Labor Economics* 13, no. 2:201-245.

- Borjas, George J. 2015. The slowdown in the economic assimilation of immigrants: Aging and cohort effects revisited again. *Journal of Human Capital* 9, no. 4:483-517.
- Borjas, George J. and Bernt Bratsberg. 1996. Who leaves? The outmigration of the foreign-born. *The Review of Economics and Statistics* 78, no. 1:165-76.
- Borjas, George J. and Rachel M. Friedberg. 2009. Recent trends in the earnings of new immigrants to the United States. Working Paper 15406, NBER.
- Chiswick, Barry R. 1978. The effect of americanization on the earnings of foreign-born Men. *Journal of Political Economy* 86, no. 5:897-921.
- Chiswick, Barry R., Yinon Cohen, and Tzippi Zach. 1997. The labor market status of immigrants: effects of the unemployment rate at arrival and duration of residence. *Industrial and Labor Relations Review* 50, no. 2:289-303.
- Chiswick, Barry R. and Paul W. Miller. 2011. The “negative” assimilation of immigrants: A special case. *Industrial and Labor Relations Review* 64, no. 3.
- de Matos, Ana D. 2013. The careers of immigrants. Unpublished manuscript, Department of Economic, London School of Economics.
- Duleep, Harriet Orcutt and Daniel J. Dowhan. 2002. Insights from longitudinal data on the earnings growth of U.S. foreign-born men. *Demography* 39, no. 3:485-506.
- Duncan, Brian and Stephen J. Trejo. 2012. The employment of low-skilled immigrant men in the United States. *American Economic Review: Papers & Proceedings* 102, no. 3:549-554.
- Ellis, Mark and Richard Wright. 1998. When immigrants are not migrants: Counting arrivals of the foreign born using the U.S. Census. *The International Migration Review* 32, no. 1:127-144.
- Funkhouser, Edward. 2000. *Issues in the economics of immigration*, chap. Convergence in Employment Rates of Immigrants. University of Chicago Press, 143-184.
- Goel, Deepti and Kevin Lang. 2010. Social Ties and the Job Search of Recent Immigrants. Working Paper 148, Collegio Carlo Alberto.
- Hu, Wei-Yin. 2000. Immigrant earnings assimilation: Estimates from longitudinal data. *A.E.R. Papers and Proceedings* 90, no. 2:368-372.

- Kaushal, Neeraj. 2011. Earnings trajectories of highly educated immigrants: Does place of education matter? *Industrial and Labor Relations Review* 64, no. 2:323-340.
- Kaushal, Neeraj, Yao Lu, Nicole Denier, Julia Shu-Huah Wang, and Stephen J. Trejo. 2015. Immigrant employment and earnings growth in Canada and the U.S.: Evidence from longitudinal data. Working Paper 21591, NBER.
- Kim, Seik. 2012. Economic assimilation of foreign-born workers in the United States: An overlapping rotating panel analysis.
- LaLonde, Robert J. and Robert H. Topel. 1992. *Immigration and the workforce: Economic consequences for the United States and source areas*, chap. The Assimilation of Immigrants in the U.S. Labor Market. University of Chicago Press, 67-92.
- LaLonde, Robert J. and Robert H. Topel. 1997. *Handbook of Population and Family Economics*, vol. 1B, chap. Economic Impact of International Migration and the Economic Performance of Migrants. Elsevier Sci, 799-847.
- Li, Peter S. 2003. Initial earnings and catch-up capacity of immigrants. *Canadian Public Policy/Analyse de Politiques* 29, no:319-337.
- Lofstrom, Magnus. 2002. Labor market assimilation and the self-employment decision of immigrant entrepreneurs. *Journal of Population Economics* 15:83-114.
- Lubotsky, Darren. 2007. Chutes or ladders? A longitudinal analysis of immigrant earnings. *Journal of Political Economy* 115, no. 5:820-867.
- Munshi, Kaivan. 2003. Networks in the modern economy: Mexican migrants in the U.S. labor market. *The Quarterly Journal of Economics* 118, no. 2:549-599.
- Passel, Jeffrey S. 2006. The size and characteristics of the unauthorized migrant population in the U.S. Tech. rep., Pew Hispanic Center.
- Picot, Garnett and Patrizio Piraino. 2013. Immigrant earnings growth: Selection bias or real progress? *Canadian Journal of Economics* 46, no. 4:1510-1536.
- Ruggles, Steve, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. 2015. Integrated Public Use Microdata Series: Version 6.0 [machine-readable database]. Minneapolis: University of Minnesota, 2015.

- Turra, Cassio M. and Irma T. Elo. 2008. The impact of salmon bias on the Hispanic mortality advantage: new evidence from Social Security data. *Population Research and Policy Review* 27, no. 5:515-530.
- Turra, Cassio M. and Noreen Goldman. 2007. Socioeconomic differences in mortality among U.S. adults: Insights into the Hispanic paradox. *Journal of Gerontology* 62, no. 3:S184-S192.
- United States Census Bureau. 2012. Census Bureau releases estimates of undercount and overcount in the 2010 Census.
https://www.census.gov/newsroom/releases/archives/2010_census/cb12-95.html.
- United States Citizen and Immigration Services. 2013a. L-1a intracompany transferee executive or manager. <https://www.uscis.gov/working-united-states/temporary-workers/l-1a-intracompany-transferee-executive-or-manager>.
- , 2013b. L-1b intracompany transferee specialized knowledge.
<https://www.uscis.gov/working-united-states/temporary-workers/l-1b-intracompany-transferee-specialized-knowledge>.
- United States Department of Homeland Security Office of Inspector General. 2013. Implementation of L-1 visa regulations. Tech. rep., U.S. Department of Homeland Security, Washington, D.C.
- United States Immigration and Naturalization Service. 2000. 2000 Statistical yearbook of the immigration and naturalization Service. Tech. rep., U.S. Department of Justice, Washington, D.C.

Appendix 1

Table A1 Sources of Variables

Variables from 2000 Decennial Census Data
Citizen (immigrants only)
Disabled
Education
English Ability (immigrants only)
Hours worked per week in 1999
Manager in 1999
Married in 1999
Place of Birth
Self-reported Earnings in 1999
Self-Reported Year of Migration (immigrants only)
State of Residence in 1995
State of Residence in 2000
Student in 1999
Weeks worked in 1999
Variables from Longitudinal Employer-Household Dynamics Data
Quarterly Earnings
Industry of firm
Variables from Social Security Administration Data
Year of application for SSN
State of application for SSN

Appendix 2

As discussed above, the characteristics of immigrants in our LEHD sample differ from that of immigrants in the 2000 Decennial Census that self-report arriving in the U.S. in 1995-1999. Here, we consider a subset of our LEHD sample and develop weights to better represent individuals in the 2000 Census. We report baseline earnings assimilation results from regressions using these weights.

The reweighting is complicated by the fact that we did not construct our sample entirely based on responses to the 2000 Census. As described previously, the LEHD sample was drawn from workers who are found in the 2000 Census 1-in-6 long form data and who have a valid SSN. A valid SSN is necessary because the match between the Decennial Census and LEHD earnings is through a Personal Identification Key (PIK), which is based on an individual's SSN number. Additionally, immigrants in the LEHD sample are those who were matched to SSN application data and who applied for an SSN in 1995-1999 in the 27 states for which we have LEHD earnings information. Natives must have self-reported living in one of the 27 states in 1995 in the Census to be in the LEHD sample. Finally, workers must have at least two quarters of earnings in the LEHD data to be included in the analysis.

Because our sample of immigrants are made up of those who applied for an SSN in 1995-1999 rather than those who self-reported arriving in the U.S. in those years, there is no straightforward group of immigrants in the Decennial Census that our sample can be made to represent by reweighting.⁴⁵ In order to facilitate a meaningful comparison between immigrants in the Census and those in the LEHD, we consider a Census sample of male immigrants who self-report arriving in the U.S. in 1995-1999, who live in the 27 states of interest, and are ages 25-57 at the time of the 2000 Census. This is the broader immigrant population we seek to represent with the reweighting. The Census sample of natives is a random sample of those who live in the 27 states in 1995 and 2000 and are ages 25-57 at the time of the 2000 Census.

Our original LEHD sample is not a subset of this Census sample so we modify the LEHD sample to only include natives and immigrants who are in the Census sample to

⁴⁵ As shown in Table 1, some immigrants in our sample who applied for an SSN in 1995-1999 self-reported arriving in the U.S. outside of that time period.

facilitate a comparison.⁴⁶ This is a subset of our original LEHD sample and we will refer to this as the modified LEHD sample. An immigrant in the Census sample is not in the modified LEHD sample if he does not have a PIK, if he has a PIK but did not apply for an SSN in 1995-1999 in one of the 27 states with LEHD data, or if he does not have at least 2 quarters of earnings in the LEHD. A native in the Census sample is not in the modified LEHD sample if he does not have a PIK or if he does not have at least 2 quarters of earnings in the LEHD.

Weights are computed for the modified LEHD sample to reflect the observable characteristics of the Census sample. Let w_i denote the original weight for individual i and $p(x_i)$ denote the probability that a person with characteristics x_i is in the modified LEHD sample. The observations for a person in the modified LEHD sample will have a weight given by $w_i/\hat{p}(x_i)$, the product of w_i and the inverse of the estimated probability of being in the modified LEHD sample. Like Lubotsky (2007), we estimated the probability of being in the modified LEHD sample with a logit model in which the probability of matching is a function of variables from the 2000 Census. We include years of education, a square in potential labor market experience, weeks and hours worked, a square in reported earnings, and indicators for Hispanics, nonwhites, Hispanic nonwhites, Asians, and those who reported not working, were self-employed, or worked in agriculture. We estimate the logit model separately for the three education groups for both natives and immigrants.

Table A2 reports the match rate between our Census sample and the modified LEHD sample by education level and immigrant status. For all education levels, the match rate is substantially higher for natives than immigrants. As mentioned above, the reasons for nonmatch for natives is not having a PIK and/or not having LEHD earnings in 1995-2008 in the relevant states. In addition to these reasons, immigrants may not match because they did not apply for an SSN in 1995-1999 in one of the 27 states considered.

⁴⁶ Our original LEHD sample does not exclude immigrants based on their self-reported arrival date. It also does not exclude immigrants based on their state of residence but rather based on the state in which they applied for an SSN. Our full LEHD sample does not require natives to be in the 27 states in 2000 but only in 1995.

Table A2: Match Rates From Decennial Census Sample to LEHD Sample

	Natives	Immigrants
All	84.32	31.4
n	498000	176200
Less than College	82.84	22.97
n	338800	121500
Bachelor's Degree	88.53	48.1
n	80600	28200
Graduate Degree	87.35	50.52
n	78600	26500

Note.— Sample sizes are rounded to the nearest hundred due to confidentiality requirements.

To gauge how well the final weights adjust for nonmatches between the Decennial Census and the LEHD data, Table A3 reports the level of self-reported earnings and the immigrant-native earnings gaps for the Census sample and the matched subsample with and without the reweighting. Comparing Columns 1 and 2, we see a substantial difference in the self-reported earnings of immigrants with less than a college education between the Decennial Census and the subset that is in the modified LEHD sample; those in the modified LEHD sample report higher earnings than those who are not matched. Because of this, the immigrant-native earnings gap for this education group is smaller in the modified LEHD sample than when considering the entire Decennial Census sample. This is also the case for those whose highest level of education is a bachelor's degree, but to a lesser extent. For both of these groups, reweighting the modified LEHD sample to better represent the Decennial Census sample results in an immigrant-native earnings gap that is closer to what is found in the Decennial Census.

Table A3: Log Earnings from 2000 Decennial Census

	Decennial Census (1)	Modified LEHD (2)	Modified LEHD reweighted (3)
Less than College			
Natives	10.2347	10.2573	10.2360
n	286200	249700	249700
Immigrants	9.5908	9.7146	9.6582
%Gap	-64.39	-54.274	-57.78
n	97700	23200	23200
Bachelor's Degree			
Natives	10.7437	10.7524	10.7396
n	74100	66600	66600
Immigrants	10.3311	10.3915	10.3408
%Gap	-41.26	-36.09	-39.88
n	23400	12200	12200
Graduate Degree			
Natives	10.9888	11.0096	10.9849
n	71300	63600	63600
Immigrants	10.3959	10.4060	10.3858
%Gap	-59.29	-60.36	-59.91
n	22900	12200	12200

Note.— Sample sizes are rounded to the nearest hundred due to confidentiality requirements.

Table A4 presents cross-sectional assimilation regressions like those in our main analysis for the modified LEHD sample. Columns 1-3 display results from regressions using the original weights. The results are similar to those from our full LEHD sample which are reported in Columns 1-3 of Table 6. Ideally we would reweight the full LEHD sample to better represent the sample of interest in the Decennial Census as a robustness check. However, as discussed above, this is not a possibility because of the nature of our sample selection for the full LEHD sample. Instead, we reweight the modified LEHD sample to better represent the sample of immigrants in the 2000 Decennial Census. The results from the reweighted assimilation regressions are reported in Columns 4-6 of Table A4. The results indicate that the patterns of assimilation remain the same after the reweighting; immigrants with less than a college education appear to experience the most substantial relative earnings gains with time in the U.S. while immigrants whose highest level of education is a bachelor's degree appear to experience a relative earnings decline in the cross-section. The similarity between the regression results with the original

weights and the results with the new weights suggests that selection into the LEHD sample does not substantially drive our results. For the rest of the paper, we will use the full LEHD sample with the original weights rather than the modified LEHD sample with the new weights.

Table A4: Log Quarterly Earnings Regressions- Modified LEHD Sample

	Modified LEHD Sample			Modified LEHD Sample Reweighted		
	High School Dropouts/ Graduates	Bachelor's Degree	Graduate Degree	High School Dropouts/ Graduates	Bachelor's Degree	Graduate Degree
	(1)	(2)	(3)	(4)	(5)	(6)
Age at Migration:						
20-25	-0.0119* (0.0029)	0.3362* (0.0043)	0.2942* (0.0049)	0.0537* (0.0041)	0.2866* (0.0046)	0.2635* (0.0052)
26-30	-0.1284* (0.0025)	0.1159* (0.0034)	-0.0107* (0.0037)	-0.0971* (0.0034)	0.0655* (0.0037)	-0.0400* (0.0040)
31-35	-0.2285* (0.0027)	-0.0944* (0.0041)	-0.2296* (0.0041)	-0.2450* (0.0038)	-0.1569* (0.0044)	-0.2571* (0.0044)
36-40	-0.3143* (0.0030)	-0.2090* (0.0052)	-0.4027* (0.0054)	-0.3232* (0.0040)	-0.2765* (0.0056)	-0.4463* (0.0058)
41-45	-0.4030* (0.0032)	-0.3993* (0.0062)	-0.5262* (0.0073)	-0.4054* (0.0042)	-0.4524* (0.0065)	-0.5754* (0.0081)
>=46	-0.4556* (0.0030)	-0.5145* (0.0063)	-0.5664* (0.0082)	-0.4306* (0.0040)	-0.5649* (0.0064)	-0.6159* (0.0090)
Year Since migration						
6-10 YSM	0.1014* (0.0023)	-0.0484* (0.0036)	0.0184* (0.0038)	0.0993* (0.0031)	-0.0397* (0.0038)	0.0221* (0.0042)
11+ YSM	0.1203* (0.0037)	-0.0783* (0.0060)	0.0402* (0.0062)	0.1086* (0.0050)	-0.0639* (0.0063)	0.0554* (0.0067)
Observations	11773300	3376000	3156100	11773300	3376000	3156100
R-squared	0.0613	0.0731	0.0713	0.0635	0.0696	0.0678

Note.— Robust standard errors in parentheses. Regressions control for year-quarter dummies and a quartic in age. Columns 3 and 6 also include an indicator for professional/Ph.D. degrees. Sample sizes are rounded to the nearest hundred due to confidentiality requirements. + $p < .10$ ** $p < .05$. * $p < .01$.

Appendix 3

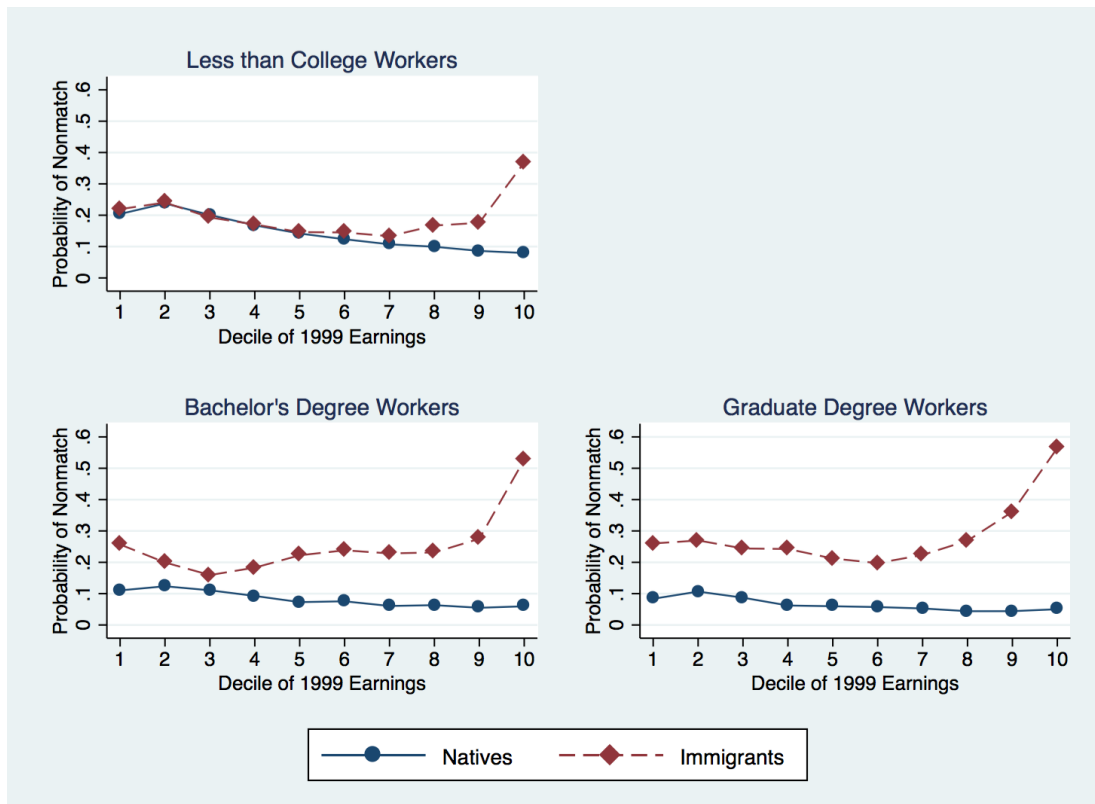


Figure A1: Predicted Probability of Matching to the 2010 Decennial Census by Decile of the 1999 LEHD Earnings Distribution
 Note.—Estimates are for the 1960 birth cohort.

Appendix 4

Table A5: Earnings Assimilation by Labor Force Attachment

	High School Dropouts/Graduates			Bachelor's Degree			Graduate Degree		
	Percent of Total Quarters Worked			Percent of Total Quarters Worked			Percent of Total Quarters Worked		
	100% (1)	85% or more (2)	Less than 85% (3)	100% (4)	85% or more (5)	Less than 85% (6)	100% (7)	85% or more (8)	Less than 85% (9)
Age at Migration:									
20-25	0.1333* (0.0031)	-0.0437* (0.0028)	-0.1666* (0.0053)	0.4667* (0.0061)	0.2426* (0.0056)	0.0733* (0.0105)	0.3838* (0.0088)	0.3510* (0.0066)	0.1251* (0.0099)
26-30	0.0004 (0.0027)	-0.1924* (0.0024)	-0.3777* (0.0048)	0.2589* (0.0047)	-0.0182* (0.0048)	-0.2463* (0.0087)	0.1549* (0.0058)	-0.0579* (0.0053)	-0.2880* (0.0086)
31-35	-0.1328* (0.0032)	-0.3153* (0.0029)	-0.4613* (0.0055)	0.0347* (0.0064)	-0.2756* (0.0055)	-0.4500* (0.0105)	-0.0883* (0.0063)	-0.3246* (0.0063)	-0.6197* (0.0100)
36-40	-0.2204* (0.004)	-0.3933* (0.0032)	-0.5702* (0.0063)	-0.1059* (0.0081)	-0.4235* (0.0073)	-0.6426* (0.0124)	-0.2594* (0.0082)	-0.5492* (0.0077)	-0.6830* (0.0124)
41-45	-0.3511* (0.0046)	-0.5001* (0.0036)	-0.6388* (0.0072)	-0.2455* (0.0102)	-0.5986* (0.0083)	-0.7416* (0.0148)	-0.3486* (0.0121)	-0.7174* (0.0104)	-0.8139* (0.0173)
>=46	-0.3877* (0.0042)	-0.5662* (0.0034)	-0.7162* (0.0069)	-0.4884* (0.0108)	-0.7158* (0.0082)	-0.8702* (0.0151)	-0.3691* (0.0148)	-0.7600* (0.0106)	-0.9888* (0.0202)
Year Since migration									
6-10 YSM	0.1183* (0.0027)	0.1353* (0.0023)	0.0863* (0.0045)	0.0523* (0.0052)	0.0537* (0.0048)	0.0271* (0.0087)	0.0405* (0.0060)	0.0788* (0.0054)	0.1147* (0.0086)
11+ YSM	0.1328* (0.0043)	0.1548* (0.0036)	0.1171* (0.0070)	0.0300* (0.0094)	0.0371* (0.0077)	-0.0208 (0.0134)	0.0921* (0.0112)	0.1145* (0.0084)	0.1480* (0.0121)
Observations	8208100	8475100	8154300	2479000	2536600	2441000	2269900	2319600	2252500
R-squared	0.0683	0.0693	0.0692	0.0911	0.0912	0.0913	0.0897	0.0903	0.0908

Note.— Robust standard errors in parentheses. Regressions control for year-quarter dummies and a quartic in age. Columns 7-9 also include an indicator for professional/Ph.D. degrees. Sample sizes are rounded to the nearest hundred due to confidentiality requirements. + $p < .10$ ** $p < .05$. * $p < .01$.

Appendix 5

Table A6: Public Use Data Wage Regression Results

	High School Dropouts/ Graduates (1)	Bachelor's Degree (2)	Graduate Degree (3)
Age	0.1039* (0.0071)	0.1620* (0.0133)	0.2510* (0.0195)
Age ²	-0.0016* (0.0002)	-0.0023* (0.0003)	-0.0041* (0.0004)
Age ³	0.0740* (0.0126)	0.0827* (0.0240)	0.2121* (0.0328)
Immigrant	0.0504* (0.0078)	0.3369* (0.0183)	0.2020* (0.0202)
1985-1989 arrivals	-0.0106* (0.0035)	-0.0008 (0.0090)	0.0155 (0.0095)
1990-1994 arrivals	-0.0123* (0.0038)	0.0183+ (0.0096)	0.0705* (0.0098)
1995-1999 arrivals	-0.0205* (0.0039)	0.0986* (0.0097)	0.1282* (0.0100)
2000-2004 arrivals	-0.0442* (0.0044)	-0.0090 (0.0112)	0.0812* (0.0114)
2005-2011 arrivals	-0.0650* (0.0061)	0.0237+ (0.0140)	0.1086* (0.0141)
Year since migration	0.0051* (0.0015)	-0.0176* (0.0035)	0.0137* (0.0035)
(Year since migration) ²	-0.0001 (0.0001)	0.0006** (0.0003)	-2.70E-07 (0.0003)
(Year since migration) ³ *10 ⁻⁴	0.0064 (0.0270)	-0.0546 (0.0676)	-0.0841 (0.0695)
1990 period effect	0.1303 (0.1031)	0.7018* (0.1944)	0.7280** (0.3030)
2000 period effect	0.3308* (0.1036)	-0.9857* (0.1943)	0.2801 (0.3035)
2007 period effect	0.3267 (0.1310)	-0.7431* (0.2402)	-0.5311 (0.3777)
2008 period effect	0.0719 (0.1297)	-0.6798* (0.2383)	-0.9803* (0.3707)
2009 period effect	0.1653 (0.1311)	-0.5759** (0.2366)	0.4070 (0.3644)
2010 period effect	0.1984 (0.1326)	-0.3764 (0.2380)	-0.0643 (0.3740)
Age at migration	-0.0094* (0.0001)	-0.0152* (0.0004)	-0.0154* (0.0004)
Constant	0.7363* (0.0975)	0.2566 (0.1790)	-.9293* (0.2750)
Observations	5,138,978	1,346,637	769,274
R-squared	0.162	0.176	0.176

Note.— Robust standard errors are reported in parentheses. Regressions also interact age variables with period effects. Column 1 also includes an indicator for high school graduate.
+ p<.10. **p<.05. *p<.01.