Title

Immigrants’ Earnings Growth and Return Migration from the U.S.: Examining their Determinants using Linked Survey and Administrative Data

Authors
Akee, Randall
Jones, Maggie R

Publication Date
2019-04-22

Series Name: WPS
Paper No.: 078
Issue Date: 22 April 2019

Immigrants' Earnings Growth and Return Migration from the U.S.: Examining their Determinants using Linked Survey and Administrative Data

Randall Akee, Maggie R. Jones

CEGA
Center for Effective Global Action

Working Paper Series
Center for Effective Global Action
University of California
eScholarship
University of California

IMMIGRANTS' EARNINGS GROWTH AND RETURN MIGRATION FROM THE U.S.:
EXAMINING THEIR DETERMINANTS USING LINKED SURVEY AND ADMINISTRATIVE DATA

Randall Akee
Maggie R. Jones

Any opinions and conclusions expressed herein are those of the authors and do not necessarily reflect the views of the U.S. Census Bureau. The statistical summaries reported in this paper have been cleared by the Census Bureau's Disclosure Review Board release authorization number CBDRB-FY18-478. All results have been reviewed to ensure that no confidential information is disclosed. We would like to thank Dany Bahar, David Card, Michael Clemens, Ira Gang, Chinhui Juhn, Devesh Kapur, Melanie Khamis, Mark Rosenzweig, Neil Ruiz, Todd Sorenson, and Steven Stillman for helpful comments and feedback. Any errors are ours alone.

Immigrants' Earnings Growth and Return Migration from the U.S.: Examining their Determinants using Linked Survey and Administrative Data
Randall Akee and Maggie R. Jones
March 2019
JEL No. F22,J15,J61

ABSTRACT

Using a novel panel data set of recent immigrants to the U.S. (2005–2007) from individual-level linked U.S. Census Bureau survey data and Internal Revenue Service administrative records, we identify the determinants of return migration and earnings assimilation. We show that by 10 years after arrival almost 40 percent have return migrated. We show, for the first time, that return migrants experience downward earnings mobility over two to three years prior to their return migration. This finding suggests that economic shocks are closely related to emigration decisions. As a result, standard calculations of immigrants earnings growth may be understated.

Randall Akee
Department of Public Policy
University of California, Los Angeles
3250 Public Policy Building
Los Angeles, CA 90095
and NBER
rakee@ucla.edu

Maggie R. Jones
Bureau of the Census
4600 Silver Hill Rd.
Washington, DC 20233
margaret.r.jones@census.gov

1 Introduction

The economic impact of immigration has been an important topic of economics research, especially during periods of large immigrant inflows. Economists and policy-makers are concerned about potential negative impacts that newly arriving immigrant cohorts may have on the labor-market outcomes of native-born workers (Borjas, 2003; Card, 1990; Cortes, 2008; Foged and Peri, 2016; Friedberg, 2001; Hunt, 1992; Ottaviano and Peri, 2012; Peri, 2012; Peri and Sparber, 2009). Major questions regarding immigration include how incoming workers of different abilities impact the overall wage structure and skill-based wage inequality. Related research topics have examined the quality of new arrivals in human capital terms and earnings ability (Borjas, 1995; Card, 2005, 2009; Chiquiar and Hanson, 2005; Chiswick, 1978; Lubotsky, 2007). Specifically, the economic question focuses on whether and to what extent new arrivals assimilate into the U.S. economy over time in terms of employment and earnings (generally measured as a convergence in earnings between immigrants and the native born).

In this paper, we use a complete panel of new immigrants to the U.S. Our panel contains a decade's worth of records for those immigrants who remain and those who eventually leave due to return migration over time. We identify new arrivals to the U.S. using 2005–2007 American Community Survey (ACS) data. We restrict analysis to immigrants who were between the ages of 25–45 at the time of entry, who participated in the formal labor market, and who report entering either in the current year (2005, 2006, or 2007) or one year earlier (2004, 2005, or 2006); all other immigrants are excluded from this analysis. A complete entry cohort should have experienced very little return migration between the reported entry year and inclusion in the ACS; thus we take these samples as representative of the complete, formal-market immigrant arrival cohorts in our analysis. We link these observations at the person level to their Internal Revenue Service (IRS) W-2 or 1099 forms for the years 2005–2015. This novel panel data set differs from previous stock-based immigration panel data in that we are able to identify those in the initial arrival cohorts who leave. In other words, these data allow us to examine the forward-looking trajectory of recent immigrant earnings (whereas previous stock-based panels were retrospective), capturing the labor market outcomes of stayers and leavers.

The first generation of research into this topic estimated wage regressions for a single cross-section, examining immigrants' tenure in the U.S. to assess the rate of wage assimilation (Chiswick, 1978). This research indicated a strong convergence of immigrant wages to those of native-born workers, with immigrant wages even eventually surpassing those of the native born. The next generation of research critiqued earlier methods for failing to account for the average differences in immigrant cohorts over time (Borjas, 1985; Borjas et al., 1987). As a result, researchers used data from repeated cross-sections (essentially the U.S. decennial censuses) and showed that the rate of wage assimilation was slower than that shown in a single cross-section. This later research aligns with a story of decreasing average immigrant cohort characteristics in human capital over time.

Our research is related to more recent work that focuses on the use of stock-based panel datasets and forms much of the base of knowledge that we have on the earnings growth and assimilation of immigrant cohorts.[1] These data tend to be merged longitudinal administrative data such as the Social Security Administration's Detailed Earnings Records (DER) and a nationally representative survey such as the Current Population Survey (CPS) or Survey of Income and Program Participation (SIPP).

[1] We provide in Appendix Figure A1 a selected list of existing research using stock-based panel and complete panel data sets in the immigration literature.

These data provide a complete panel for individuals from all prior immigrant arrival cohorts that have survived until the current time period (or the time when the national survey was conducted). Dustmann and Görlach (2015) calls this type of data set "stock-based panel data" because it provides information on the surviving stock of immigrants from an earlier arrival cohort. Duleep and Regets (1997), Duleep and Dowhan (2002), and Duleep and Dowhan (2008), using the DER merged to the CPS, found that immigrants' wage growth is greater than that of natives. They also show that immigrant earnings converge to those of natives by the 10-year mark. In contrast, Hu (2000) used Health and Retirement Survey data merged to DER data and found that wage assimilation takes longer in the longitudinal data than in the repeated cross-sectional data. Lubotsky (2007, 2011) merges DER data to the 1990 and 1991 SIPP and the 1994 CPS to create a panel dataset. He infers that low earners are the most likely to return migrate. Lubotsky (2007) finds that the earnings gap between immigrants and natives takes about 20 years to close (on average) which is about half as fast as is suggested by repeated cross-section analyses. Kaushal (2011) uses the 2010 National Survey of College Graduates to construct a stock-based panel and finds that return immigrants are negatively selected based on earnings data; in more recent work Kaushal et al. (2016) finds strong evidence that there are large gains in employment and earnings growth for immigrant men relative to natives. Other researchers (Hall and Farkas, 2008; Villarreal and Tamborini, 2018) have found differential earnings assimilation outcomes by race and English-language abilities.

This previous research generally improved our understanding of important questions related to immigration, such as: what is the wage assimilation of the immigrants who end up staying for (relatively) long periods of time in the U.S. relative to the wages of the native-born? Stock-based panel data provided improvements to answer these questions compared to repeated cross-section data. Still, stock-based panel data do not allow one to understand the true earnings growth of an initial arrival cohort once return migration is accounted for. A true panel data set, which is forward-looking, is required to answer this type of question.

There are a few complete immigrant panel datasets that researchers have compiled to examine certain aspects of return migration and earnings growth in the U.S.; however, these tend to cover a historical period, be limited to a particular occupational group, lack the employment and earnings data that allow for a study of key labor-market effects, or constitute a limited-duration panel. Jasso and Rosenzweig (1982) and Jasso and Rosenzweig (1988) use administrative data from the U.S. Immigration and Naturalization Service for a 1971 entry cohort of legal permanent residents to identify the emigration rate and occupational status over an eight-year period. They find that the average emigration rate for that cohort is about 50 percent. Both Borjas (1989), examining high-skilled immigrants, and Schwabish (2011) find that return migrants are negatively selected based on earnings data. Borjas (1989) also finds a lack of wage convergence among a set of high-skilled immigrants. Similarly, Van Hook and Zhang (2011), using the two-year panel structure of the CPS ASEC, and Abramitzky et al. (2014), using decennial census data linked at the individual level, find that negative labor market outcomes for migrants in the U.S. are associated with return migration.[2]

[2]European datasets, in contrast, have allowed researchers to examine more recent and representative immigrant cohorts to those countries. Dustmann and Weiss (2007) use the British Labor Force Surveys, which is a complete panel of five consecutive quarters to identify new arrivals in the UK and follow the return migration and the wage assimilation of those who remain in the country. They find high rates of return migration within the first 10 years of arrival. In related work, Dustmann et al. (2010) compare differences in return migration

We contribute to this literature by taking a nationally representative immigrant arrival cohort and linking it to *future* earnings reports via IRS data, in contrast to prior research that has relied on matches to historical earnings reports. This forward-looking link to earnings allows us to examine the labor-market outcomes of immigrants from the time they enter the U.S. until either they reach the end year or leave the U.S. Key to the novelty of our data is our ability to link observations captured in the 2005–2007 ACS to their administrative record for every year that they work in the formal U.S. labor market, up to 2015. This contrasts with survey-DER linkages, where administrative earnings records are linked to survey observations up to the survey year. These *retrospective* studies can only measure labor-market outcomes for immigrants who have remained in the U.S. up to the survey year. Additionally, our sample is large enough to separately categorize the immigrant group into different educational attainment, race, and country of birth categories and follow their differential earnings growth. We also separate the analysis by gender to illustrate the potentially different experiences across these two groups independently. Group-level differences in earnings growth have eluded previous analysis due either to relatively small sample sizes, short-duration panels, lack of forward-looking longitudinal earnings data, or a combination of these factors.

Our data allow us to investigate the determinants of return migration from the U.S. in much greater detail than has been possible in previous research. We show that there is a relatively quick reduction in an immigrant arrival cohort in the first few years in the U.S.—by year 10 almost 40 percent have returned home, with the majority leaving in the first few years. After identifying the immigrants who will end up leaving the U.S., we estimate earnings assimilation relative to the native born for the eventual return migrants, the panel, and the repeated cross-section data. Our analysis shows that the return migrants experience downward earnings mobility in the two to three year prior to their return migration. In general, these individuals are not statistically different at the time of arrival from others in their arrival cohort; therefore, our evidence suggests they must differ from more successful cohort members in their time-variant unobserved characteristics. These characteristics influence their earnings outcomes as well as their decision to return migrate. To our knowledge, this is the first analysis of how success or failure in the U.S. labor market impacts immigrants' choice to return migrate. The results of our analysis have been predicted by previous researchers, who argued that return migrants are negatively selected. However, the data necessary to test this prediction—namely, detailed annual earnings for the return migrants prior to their return—were not previously available.

We use characteristics of the population from the ACS data to identify immigrants' educational attainment, country of birth, and English-language ability at the time of arrival to show how these characteristics are related to return migration over time. We also examine how earnings assimilation differs by an immigrant's initial characteristics. In particular, we examine earnings assimilation across these individual characteristics. One important finding is that earnings assimilation by education level is almost complete by 10 years after arrival for those immigrants who remain in the U.S. We find that earnings assimilation within race groups is nearly indistinguishable across the immigrant–native

from the UK and Germany using two different nationally representative datasets. They find little difference in response to economic shocks across immigrants and natives in earnings, but find that economic shocks have a larger effect on unemployment rates for immigrants as compared to natives. Lehmer and Ludsteck (2015) find that reductions in earnings differences between immigrants and natives in Germany can be attributed to the acquisition of firm-specific human capital over time. Our analysis follows more closely the work conducted with administrative data from European countries.

earnings distribution by 10 years after arrival.

Finally, we examine the earnings growth across the three types of data sets available for studying immigrants. We confirm that repeated cross-section data overstates the rate of earnings growth compared to stock-based panel data. We also find that the stock-based panel analyses may under estimate the true rate of immigrant earnings growth; this may be the case when unobserved time-variant characteristics are related to both immigrants earnings and migration decisions. This is also a novel empirical contribution to the literature; previous researchers have been unable to show how stock-based panel analysis may differ from complete panel data analysis for the U.S.

2 Data Description

We use confidential-use individual-level data from the ACS for the years 2005, 2006, and 2007. The ACS provides characteristics of the population sampled annually. We use the year of entry variable to identify new arrivals to the U.S.[3] In our analysis we include individuals who report entry either in the year prior to the ACS or the current year.[4] In practice this means that we include individuals within a two-year range of the ACS (2004 and 2005 for the 2005 ACS; 2005 and 2006 for the 2006 ACS; 2006 and 2007 for the 2007 ACS).[5]

These immigrant cohorts are linked at the U.S. Census Bureau to their individual IRS data using a process whereby observations in each data set were given a unique, protected identification key, called a "PIK." When a Social Security Number (SSN) is available in a data set, the identifier is assigned based on SSN. For records without an SSN, personally identifiable information such as name, address, and date of birth is used in probabilistic matching to assign PIKs from a master reference file. Personal information is then removed from each data set before they may be used for research purposes. Only those observations that received the unique person identifier are used in the analysis. The IRS W-2 data span the years 2005–2015 and allow us to examine the earnings progression of these individuals over a 10-year time period. It also allows us to identify individuals that start out in the labor force and leave it subsequently either as return migrants or because they enter informal work.

It is important to note that the record-linkage approach we use to link the data introduces some bias. Minorities and people with lower socioeconomic status are less likely to receive a record-linkage key compared to whites and people who have higher levels of socioeconomic status (Bond et al., 2014). Because we focus on incoming immigrants and use IRS return information, we are only able to link and follow those immigrants who enter the U.S. to work in the formal labor market. Any results we report will therefore only apply to immigrants working in the formal sector. We provide some details below on how much the linking procedure covers the new immigrant population in the ACS.

Once these ACS individuals are linked to their earnings records, we identify whether the individuals are present annually in each of the years from 2005–2015. We calculate the two outcome

[3] Grieco et al. (2018) has shown that the single year ACS data provides more accurate reporting than multi-year ACS data for immigrant arrivals. Additionally, they find that there is an undercount the longer the recall period. Our analysis only requires an immigrant to recall arrival for the current or prior year.

[4] The question on the ACS asks: "When did this person come to live in the United States?" and as such may not necessarily be the first time that an individual came to the U.S.; it simply indicates the most recent date of arrival.

[5] While the ACS data provide a snap-shot perspective of the country on average, the sampled individuals are not the same on an annual basis and it is not possible to create a panel data set using the ACS alone.

variables of interest: missing W-2 or 1099 filings in subsequent years (2006–2015) and relative earnings (as compared to natives). We provide these results disaggregated by country of origin, educational attainment, race, and English language ability as reported in the 2005–2007 ACS by gender. This novel data set provides a representative picture of the earnings progression of recent, formal-sector employed immigrants to the U.S. over the time period 2005–2015. This data set is also unique in that there is no top-coding of earnings as is often the case for research using SSA data. We exclude all individuals who report being enrolled in college or graduate school from this analysis.[6]

Finally, for comparison, we take a sample of native-born of both genders from the 2005 ACS in the same age range. We take a 50 percent sample of the 2005 ACS population for comparison in wage assimilation and growth analyses.

Table 1 shows the rounded sample sizes for the panel data used in our analysis.[7] The first row shows that there are 13,000 individuals who fit the criteria of: male; ages 25–45; not enrolled in school of any type; and year of arrival to the U.S. is 2004 or 2005 for the 2005 ACS, 2005 or 2006 for the 2006 ACS, or 2006 or 2007 for the 2007 ACS. Similar criteria exist for women, shown in the first row of Panel B, where we find 12,000 observations. The next rows in Panel A and B show that 6,800 and 6,700 of these individuals, for males and females respectively, were assigned a PIK by the U.S. Census Bureau. The assignment of PIKs are in large part determined by whether the individual has a presence in administrative data and records.[8] By definition, an undocumented immigrant will be much less likely to be found in U.S. administrative records.[9] As a result, our panel data likely identify documented immigrants to the U.S. who have valid visas or work permits. The PIK assignment rate of 52 percent for men and 56 percent for women is very similar to what others have found for the immigrant population using the ACS data (Bond et al., 2014). In the third row of both panels we show that the link to the IRS W-2 and 1099 data results in a total sample size of 5,800 observations for men and 4,500 observations for women.[10]

[6]See Akee and Jones (2018) for analysis that focuses exclusively on immigrants enrolled in graduate school in the U.S. and their earning outcomes compared to the native-born.

[7]U.S. Census Bureau data dissemination requires rounding of count data as well as regression coefficients. Those rules have been employed throughout this paper and results have been approved for release by the Bureau's Disclosure Avoidance Review Board.

[8]Bond et al. (2014) have evaluated the assignment of PIK numbers to individuals from the ACS and find an assignment rate for the foreign-born population that is similar to ours. Subsequent versions of PIK assignment (2010 ACS) accounted for Individual Taxpayer Identification Numbers (ITINs) which may be available in IRS data when an individual does not have a Social Security Number (Bhaskar et al., 2016). However, this is not possible for the older assignment of PIKs to earlier data since the ITIN is not available for the data. As a result, the individuals studied in this analysis are more likely to be employed in the formal sector and have a strong participation in government programs in order to be identified in the Personal Identification Validation System, which assigns PIKs to individuals. The reference files used for matching contain all variants of a person's name, date of birth, and sex, as well as current and recent addresses(Bond et al., 2014).

[9]In recent research, Foster et al. (2018) found a higher match of foreign-born between the Census and IRS data than we do. Their data differ slightly in that they are examining all immigrant arrival cohorts, not just the most recent ones as we are. Moreover, they are using a wider variety of IRS data, including the 1040 tax returns; we do not use these data since we are primarily concerned with identifying the individual immigrant's earnings in isolation, which might not be discernible on jointly-filed tax forms.

[10]In Appendix Tables A1 and A2 we show the comparison of several characteristics for the matched and unmatched observations for both matching steps. The data show, as expected, that the matched observations tend to have higher earnings and more education–these are characteristics that are associated with having a larger presence in administrative records used for matching.

Table 1: Table of Dataset Creation by Subsequent Merges for Immigrants in Age Range 25–45

Panel A Men	Count	Percent of Row Above
Total Observations for Arrival Cohorts (2005-2007)	13,000	-
Not Missing PIK	6,800	0.52
Found in W-2s (2005-2007)	5,800	0.85

Panel B Women	Count	Percent of Row Above
Total Observations for Arrival Cohorts (2005-2007)	12,000	-
Not Missing PIK	6,700	0.56
Found in W-2s (2005-2007)	4,500	0.67

Source: ACS 2005–2007 (top two rows) matched to IRS W-2s (2005–2007). Sample is all immigrants ages 25–45 who are recent arrivals (see text for definition). Numbers have been rounded to comply with the Census Bureau's disclosure-avoidance guidelines.

3 Research Methods

We conduct three empirical analyses in this paper. The first examines the nature of return migration of recent U.S. arrivals. Our panel data allow us to identify individuals who do not have reported W-2 or 1099 data for the years 2005–2015 in the U.S. The second analysis examines the wage assimilation of the recent arrival cohorts of immigrants to that of native-born individuals by educational attainment, race, and country of birth (of the immigrants). The third analysis examines the wage growth by years since migration for the arrival cohort in the repeated cross-section data, the stock-based panel data, and the complete panel data with selection correction.

3.1 Estimating the Return Migration of New Immigrant Arrival Cohorts

For each year of data we count up the total number of individuals who are found to have a W-2 or 1099 form from the original entry cohorts (2005–2007). These data are then plotted in a figure for each cohort over the time period 2005–2015. Additionally, we separate our data into three different groups. The first is the attriters (or return migrants); the second is the repeated cross-section, which is the surviving portion of the initial arrival cohort in the data in a given year; and the third group is the panel of immigrants that remain for at least 10 years in our data.

We identify those whom we classify as leavers or return migrants based on whether they have reported W-2 or 1099 forms for a particular year. Individuals who have a W-2 or 1099 reported in 2005 or 2006 but are not present in the data for the years 2007–2015 are coded as part of the two-year cohort. We create a three-year cohort for those who are not present in the years 2008–2015. We create similar cohorts for leavers up to and including a ten-year cohort which were not present in 2015.[11] See Table 2, which provides the detail for how we code individuals in the 2005 cohort as missing by

[11] It is not possible to remove individuals who leave after eleven years as we do not have information on whether individuals present in 2015 may leave in the subsequent year.

length of time in the U.S.(later cohorts are treated analogously). Individuals are only coded as leavers when they are missing in consecutive years until 2015. Our methods allow individuals to have short unemployment spells and not be counted as leavers as long as they are observed to have a W-2 or 1099 again at some later time period. In Appendix Table A3 we attempt to find individuals who only report 1 to 6 years of administrative IRS data (for the years 2005-2010) in the 2010 U.S. Census records. Our results indicate that our assumption that the foreign born with missing administrative records have return migrated is a good one: for example, we find only four percent of individuals with a single year of administrative records (2005 alone) in the 2010 U.S. Census. This panel data set provides us with the main data for our analysis of earnings regressions over time.

Table 2: Creation of Return Migrant Counts

First W-2 Observation	Final W-2 Observation	Number of Years in U.S. Before Emigration
2005	2006	2 Years
2005	2007	3 Years
2005	2008	4 Years
2005	2009	5 Years
2005	2010	6 Years
2005	2011	7 Years
2005	2012	8 Years
2005	2013	9 Years
2005	2014	10 Years

We tabulate the mean of the missing observations as coded according to Table 2 above for each year. In further analysis, we tabulate the mean of the missing observations in the panel data set for each year separately by each category of the characteristic of interest. For example, we sum the missing by year for each of five educational categories (less than high school, high school degree, some college, college degree, and more than a college degree) for each year and plot those in a single figure. We do a similar analysis for country of birth and for English language ability.

3.2 Estimating the Wage Assimilation of New Immigrant Arrival Cohorts

Next we estimate the wage assimilation for the new arrival cohort of immigrants relative to natives over this time period. We replicate the standard repeated cross-section analysis of previous research. Given that we can further separate the data into the attriters (return migrants) and the panel component, we also show the earnings assimilation of those groups on a single graph with the repeated cross-section data for comparison. The wage assimilation figures are based on the following regression model run for each year (2005–2015) in our analysis:

$$Y_i = \alpha + \beta \times Immigrant_i + \theta_i + \gamma_i + \epsilon_i \tag{1}$$

In our equation, Y_i is the log of earnings for an individual for a single year, where earnings is the sum of all W-2 and 1099 wage reports. We include a state of residence fixed effect θ_i and an age fixed effect γ_i. The coefficient of interest is β, which indicates the difference in log earnings for an immigrant relative to a native-born individual. We also separate out the analysis by the stock-based panel, the repeated cross-section data, and the attriters. We then plot the coefficient β by year by group in the figures.

In subsequent figures, we separate out the analysis by different individual-level immigrant characteristics, such as educational attainment at the time of arrival, country of origin, race, or English-ability categories. We only show the results for the complete panel and the repeated cross-section data (shown in Appendix figures) in such analyses. These estimated coefficients are then plotted on a single figure for each characteristic.

In additional analyses shown in the Appendix, we add in educational category variables as controls for some reported results. In general, those results are similar to our main results; however, they tend to shift observed estimates downward.

3.3 Estimating the Years Since Migration on Earnings Growth

We examine the effect that years since migration has on the earnings growth of immigrants relative to native-born. The discussion below follows closely Dustmann and Görlach (2015). Researchers have been concerned with estimating the earnings growth of the immigrant populations; however, it is notoriously difficult given that the population of immigrants is not a constant one. New arrivals and the possibility of return migrants may confound any estimates. Previous research has had to make several strong assumptions in order to estimate immigrants' earnings growth using repeated cross-section data or with stock-based panel data.

3.3.1 Basic Equations for Log Earnings and Selection for Remaining in the U.S.

It is useful to first define the earnings equation as a function of two separate components. Earnings for a cohort c of immigrant arrivals are comprised of a component μ_t^c (equation 2), which accounts for the cohort average earnings. A second term, ϵ_{it}^c, is the individual i's deviation from his cohort average. This term differs across individuals and time periods denoted by t.

In Equation 3 we further separate out this deviation term into two components (α_i^c and η_{it}^c) that represent time-invariant and time-variant unobserved characteristics. The first term may represent unobserved endowment of skills and the second some type of shock in a time period.

Finally, the decision to remain in the U.S. is based on a number of both observable and unobservable characteristics and shown in Equation 4. The observable characteristics $Z_{i\tau}$ could be income, age, educational attainment, or other similar characteristics. The next determinant of return migration accounts for potential unobserved characteristics and preferences $u_{i\tau}^c$. In this equation, if the term inside the brackets is true, then an individual will remain in the U.S., otherwise, that person return migrates.

This provides both an explicit threshold for return migration as well as allowing for a potential time-varying component of earnings that may also be related to return migration decision-making.

$$\mathrm{W}_{it}^c = \mu_t^c + \epsilon_{it}^c \tag{2}$$

$$\epsilon_{it}^c = \alpha_i^c + \eta_{it}^c \tag{3}$$

$$\mathbf{S}_{it}^c = \prod_{c<\tau\leq t} 1[Z_{i\tau}^{\prime c}\beta + u_{i\tau}^c > 0] \tag{4}$$

3.3.2 Estimating Earnings and Selective Return Migration Using Repeated Cross Section Data

Several previous researchers (Borjas, 1985; Duleep and Regets, 1997; Lubotsky, 2007) have shown that using standard cross-section data to measure the earnings growth of immigrants leads to an upward bias when more recent arrival cohorts of immigrants are lower skilled than previous ones or there is negative selection of return migrants. In equation 5 we show the change in earnings over time periods t_1 and t_2. The first term on the right-hand side of the equation indicates the true growth in earnings for the arrival cohort c in the absence of return migration. The next two terms on the right-hand side provide the bias term given the presence of return migration. If there is negative selection for the return migrants from cohort c over time on unobserved characteristics, this implies that the the two terms on the right will be positive since $\mathbb{E}[\epsilon_{it_2} \mid \mu^c_{t_2}, S_{it_2} = 1] > \mathbb{E}[\epsilon_{it_1} \mid \mu^c_{t_1}, S_{it_1} = 1]$ as the individuals with the lowest unobserved skills leave, and ϵ_{it_2} will be greater than ϵ_{it_1}. Therefore, the estimated earnings growth will be larger than the true earnings growth if there were no return migration from the arrival cohort c.

$$\Delta \text{Earnings}^c_{t_1 t_2} = \{\mu^c_{t_2} - \mu^c_{t_1}\} + \mathbb{E}[\epsilon_{it_2} \mid \mu^c_{t_2}, S_{it_2} = 1] - \mathbb{E}[\epsilon_{it_1} \mid \mu^c_{t_1}, S_{it_1} = 1] \tag{5}$$

3.3.3 Estimating Earnings and Selective Return Migration using Stock-Based Longitudinal Data

One method to improve upon the bias inherent in repeated cross-section data has been to use administrative data merged to existing survey-based data. Several researchers have linked administrative data from the Social Security Administration to the SIPP or CPS and created a stock-based panel. As mentioned in the Introduction, this type of data is a panel of all individuals from arrival cohort c that survived up to some ending date T. While these data cannot provide information on those that return migrated, it does provide a panel data set of the surviving cohort of arrivals. As in equation 6, a researcher can directly estimate the change in earnings between time 1 and time T for the individuals that survived up to time T from arrival cohort c.

$$\Delta \text{Earnings}^c_{t_1 T} = \mu^c_T - \mu^c_{t_1} \tag{6}$$

However, the stock-based panel data does not allow one to estimate the true earnings growth of immigrants except with some strong assumptions. The stock-based panel data provide an accurate depiction of earnings growth for the entire cohort only under the assumption that there is no relationship between the unobserved time-variant elements of earnings and the term $u^c_{i\tau}$ from Equation 4. We again difference the earnings equation across two time periods and provide that below in Equation 7. Taking expectations and assuming the joint normality of $u^c_{i\tau}$ and η^c_{it} gives us Equation 9. In that equation, the true change in cohort c's earnings over time is the first term in the brackets on the right-hand side of the equal sign. The next two terms indicate the correlation between the individual earnings deviation and the $u^c_{i\tau}$ term in different time periods multiplied by each of those expected values in different time periods. Assuming that there is no relationship between the deviation term in the earnings equation and the unobserved characteristics in the return migration equation would mean that those final two terms are equal to zero. In the absence of such an assumption, Equation 9 does not provide an unbiased estimate of the true earnings growth of the arrival cohort c.

$$\Delta\text{Earnings}^c_{t_1t_2} = \{\mu^c_{t_2} - \mu^c_{t_1}\} + \mathbb{E}[\epsilon_{t_2} \mid \mu^c_{t_2}, S_{iT} = 1] - \mathbb{E}[\epsilon_{it_1} \mid \mu^c_{T_1}, S_{iT} = 1] \tag{7}$$

$$\begin{aligned}\Delta\text{Earnings}^c_{t_1t_2} = \{\mu^c_{t_2} - \mu^c_{t_1}\} + \Delta\alpha^c_i + \mathbb{E}[\eta^c_{it_2} \mid u_{i\tau} \text{ for all } \tau \leq t_2] \\ - \mathbb{E}[\eta^c_{it_1} \mid u_{i\tau} \text{ for all } \tau \leq t_1]\end{aligned} \tag{8}$$

$$\begin{aligned}\Delta\text{Earnings}^c_{t_1t_2} = \{\mu^c_{t_2} - \mu^c_{t_1}\} + \sigma_{\eta,u,2}\mathbb{E}[u^c_{it_2} \mid u_{i\tau} \text{ for all } \tau \leq t_2] \\ - \sigma_{\eta,u,1}\mathbb{E}[u^c_{it_1} \mid u_{i\tau} \text{ for all } \tau \leq t_1]\end{aligned} \tag{9}$$

conditional on surviving to period T.

For example, given a negative earnings shock such that $\eta^c_{it} < 0$ we would expect a similar negative realization of $u^c_{i\tau}$ from the selection equation, which would increase the probability of return migration. Therefore, the correlation between η^c_{it} and $u_{i\tau}$ should be positive. However, due to the negative self-selection of return migrants over time, the correlation between these two time-variant characteristics should be decreasing over time; individuals who were closer to leaving have already been induced to do so. This implies that $u^c_{it_2} < u^c_{it_1}$ and that the last two terms on the right could be negative and result in an underestimation of the average earnings of an immigrant cohort. Second, the correlation between the two terms η^c_{it} and $u^c_{i\tau}$ will diminish over time as the selection condition for the return migration decision (Equation 4) becomes more slack (those at the margin and most likely to return migrate have already done so in prior periods).

3.3.4 Estimating Earnings and Selective Return Migration Longitudinal Data

Using complete panel data, however, allows researchers to identify both the growth rate of the surviving individuals of arrival cohort c at time t as well as the true average cohort earnings growth $\mu^c_{t_2} - \mu^c_{t_1}$. This earnings growth rate is of particular interest for understanding the time that it takes for immigrant cohorts (absent return migration) to assimilate to the earnings of the native-born. In equation 10, we reproduce the result from above for the differenced earnings equation for periods 1 and 2. The second term can be re-written in expectations as the correlation between the unobserved time-variant terms from equations 2 and 4 and the inverse Mills ratio computed from the annual selection into the data. We assume that the two unobserved time-variant characteristics from equations 2 and 4 are joint normally distributed. The inverse Mills ratio (IMR) may be computed directly from the complete panel data by estimating a logit model where the outcome variable is equal to one in a given year if the immigrant is observed in the data and zero otherwise. Dustmann and Görlach (2015) suggest using earnings from two periods prior (in this case $t = 0$) as the exclusion restriction for this regression with the assumption that any earnings shocks in two time periods earlier are not correlated with current shocks or other unobserved characteristics. Then the IMR is included for each year in a pooled earnings regression equation as a correction for selective out-migration.

$$\Delta\text{Earnings}^c_{t_1t_2} = \Delta\mu^c_{t_1t_2} + \Delta\epsilon^c_{it} \tag{10}$$

$$\mathbb{E}[\Delta\text{Earnings}^c_{t_1t_2}] = \Delta\mu^c_{t_1t_2} + \mathbb{E}[\Delta\epsilon^c_{it}] \tag{11}$$

given that ϵ_{it}^{c} is jointly normally distributed with $u_{i\tau}$, this can be re-written as:

$$\mathbb{E}[\Delta \text{Earnings}_{t_1 t_2}^{c}] = \Delta\mu_{t_1 t_2}^{c} + \sigma_{\Delta\epsilon,\eta} \times \lambda(Z_{it}^{\prime c}\beta) \tag{12}$$

where λ is the inverse Mills ratio for the selection into the data.

In our analysis, we calculate the earnings growth for the three different data types (repeated cross-section, stock-based panel data, and complete panel) and show how they differ from one another in this case.

4 Return Migration

4.1 Return Migration by Entry Cohort of Immigrants and for Native-Born Men

In Figure 1 we present the most basic results for missing either 1099 or W-2 data for the three entry cohorts 2005–2007 for men in Panel A and for women in Panel B. The data show that there is a pronounced drop after the first year in the U.S. for both genders, and then the missing observations tend to moderate after that. Of note is the increase in filed W-2 or 1099 forms in 2011, implying a return to the labor force, which is when employment began to recover from the Great Recession.

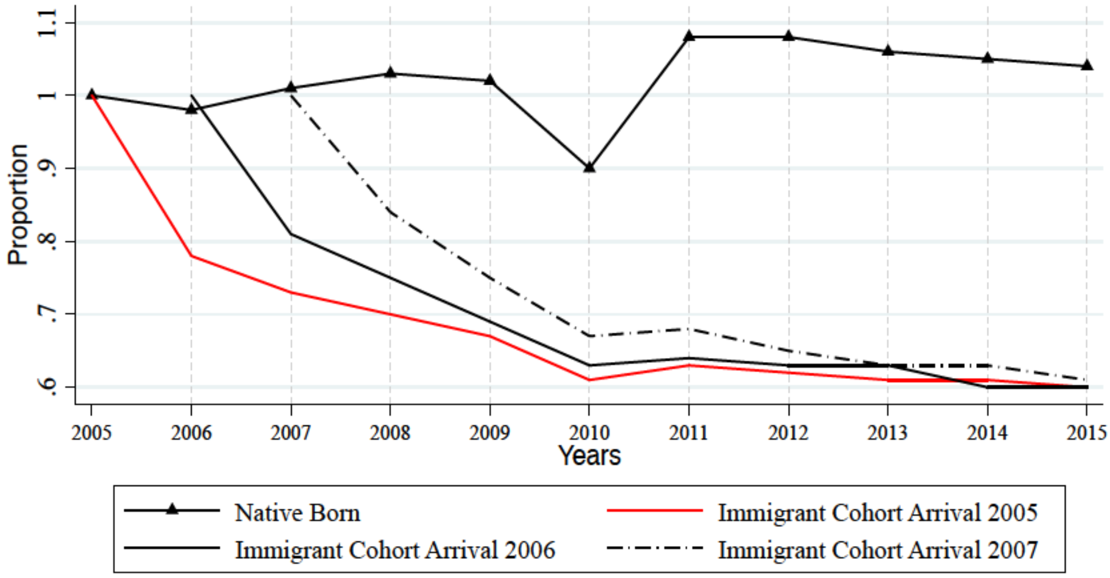
We also provide a measure of changes in the reported W-2s and 1099s for native-born males and females ages 25–45 in Figure 1. This measure includes all native-born in the same age range as the immigrants. We take the labor force participation rate for the native-born in 2005 as the base and normalize all subsequent years for native-born relative to this amount. This measure takes the count of the total individuals with a reported W-2 or 1099 in the year 2005 and then shows any changes in that amount by subsequent year. The line for the native-born shows a increase in labor force participation and then a steep drop in 2010 during the Great Recession.[12] There is a recovery by 2011 that exceeds the original proportion in 2005, and then there is a slight downward trend (which may reflect entry into disability or early retirement). For the native-born we allow for gaps in reporting of W-2s and 1099s, which may explain the increase in labor force participation rates we observe for 2011. For all three arrival cohorts of immigrants, we observe that there is a steep drop between the year of arrival and the next year; almost 20 percent of the observed cohort of immigrants are no longer present in the W-2 or 1099 data.[13] There is a continual decline in the number of immigrants with reported W-2 or 1099 data over the next 10 years, and by 2015 each of the three entry cohorts contains slightly more than 60 percent of persons from the the initial arrival cohort.

[12]We have also re-weighted these observations so that they are similar to the foreign born on demographic characteristics and we observe qualitatively similar results for the native-born results.

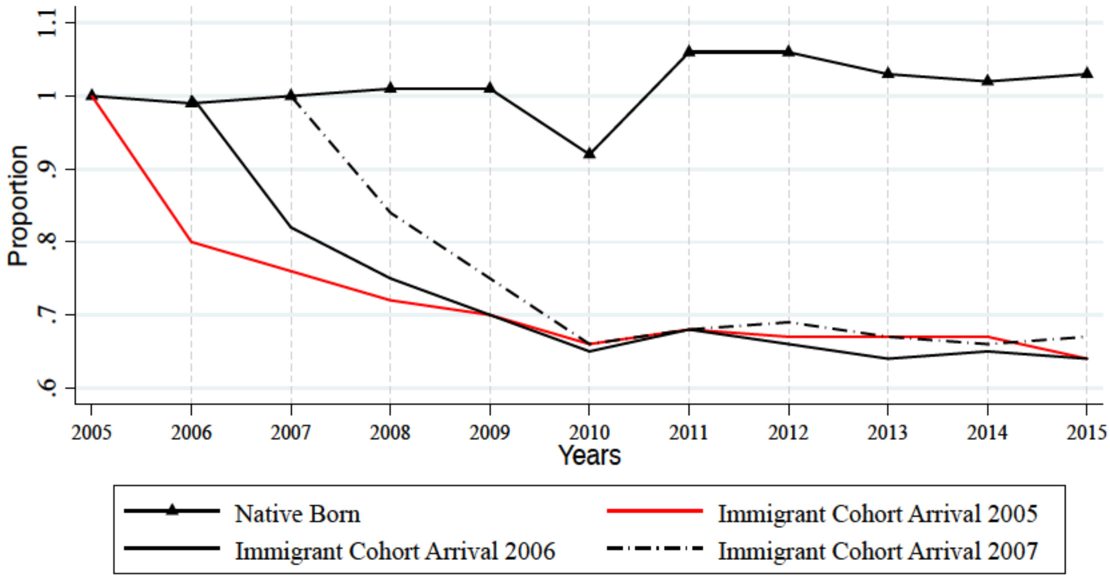
[13]We find qualitatively similar results for other arrival cohort years over and during the Great Recession. In Appendix Figure A2 we provide data for return migration for individuals who arrived in the U.S. for the years 2005–2013.

Figure 1: Presence of W-2 or 1099 for 2005-2007 Entry Cohorts for Ages 25–45

Panel A: Men



Panel B: Women



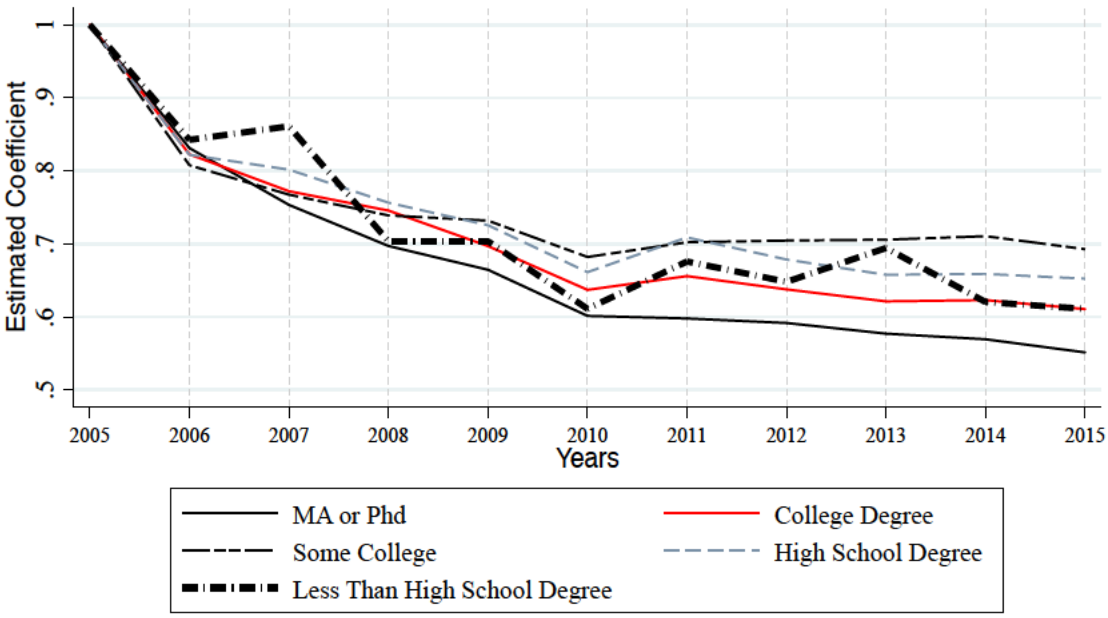
Notes: Each point represents the proportion of each group that is present in the data for each year. We start our analysis in 2005/2006/2007 and take each as the complete immigrant arrival cohort. For the native-born, we take the proportion reporting a W-2 or 1099 in 2005 as the base amount and subsequent amounts are relative to that 2005 rate. Source: ACS 2005–2007 and IRS W-2s or 1099 data (2005–2015).

4.2 Return Migration by Initial Educational Attainment

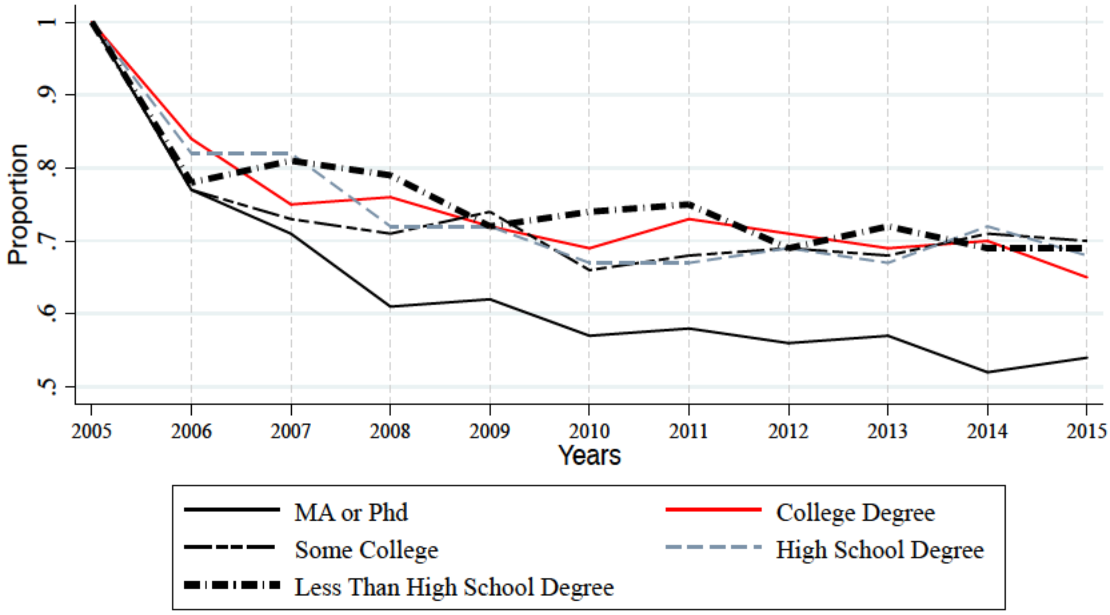
In Figure 2 we show missing tax forms by the educational attainment of the individual immigrant for the 2005 entry cohort alone. We assume that educational attainment occurred abroad because we have selected individuals who are 25-45 years old (who have only recently arrived in the U.S.) and have explicitly excluded any individuals who report that they are currently enrolled in school of any type. Overall, the more educated individuals of either gender are most likely of any education group to return migrate. The data for individuals with less than high school degree are the most volatile, reflecting the higher probability of this education group to be unemployed or perhaps employed in the informal economy. On the other hand, the other educational attainment categories are mostly monotonically decreasing over time for the males. Individuals with an MA or PhD are the most likely to leave, followed by those with a college degree. For women, the very highly-educated have a steep rate of return migration over the first three years, but the rate of return does not differ substantially over the other education categories. These education categories are independent of one another and these results should be interpreted as the percent returning from within each educational category.

Figure 2: Presence of W-2 or 1099 for 2005 Entry Cohort for Ages 25–45 by Educational Attainment

Panel A: Men



Panel B: Women



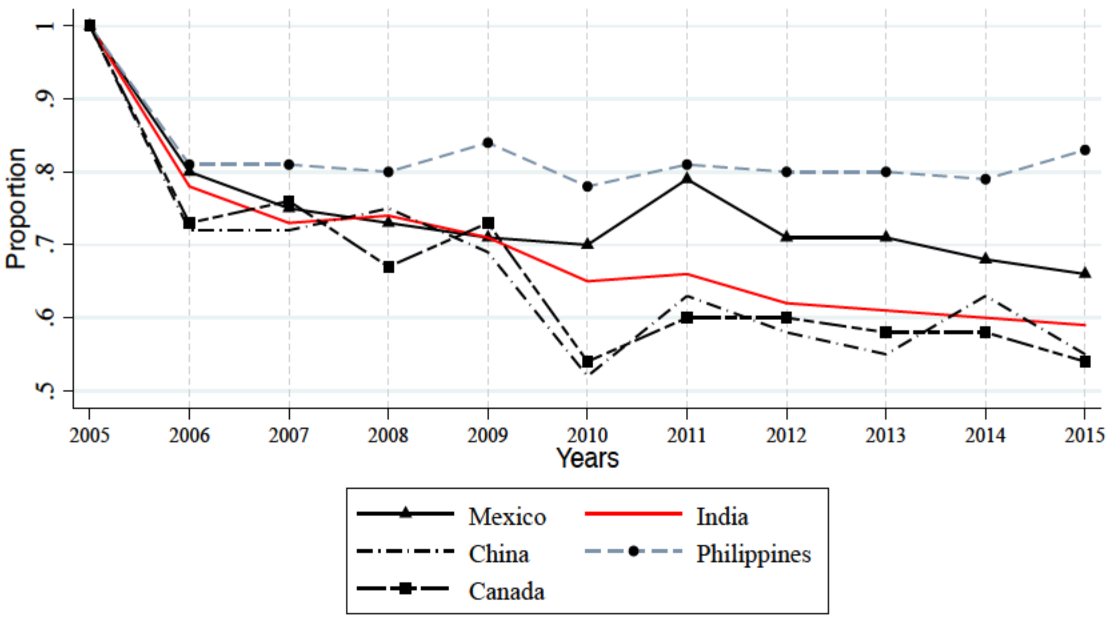
Notes: Each point represents the proportion of each group that is present in the data for each year. We start our analysis in 2005 and take that as the complete immigrant arrival cohort. Source: ACS 2005 and IRS W-2s or 1099 data (2005–2015).

4.3 Return Migration by Country of Birth

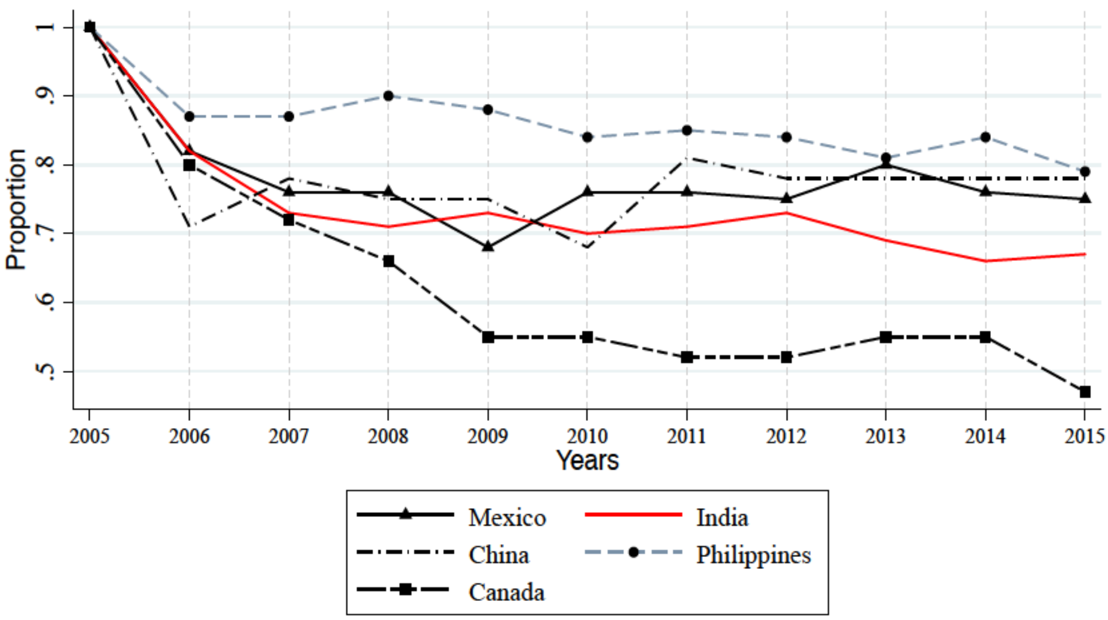
Figure 3 provides the return migration of individuals from the 2005 entry cohort by country of origin for the top five immigrant-sending countries. Panel A presents the results for males and Panel B presents the results for females. In Panel A, the highest return migration is for Canada, followed by China, India, Mexico, and Philippines. There is a steep drop for Canadians and Chinese men in 2010 which may indicate that this group was the most likely to return migrate as a result of the Great Recession. Mexican males rebound in their reporting of W-2 or 1099 in 2011, indicating that they may have remained in the U.S., perhaps working in the informal sector, but that they returned to formal sector employment in 2011 and returned to trend in 2012. Immigrants from the Philippines in this entry cohort are the least likely to return migrate as a percent of their initial arrival cohort. In Panel B Canadian women are the most likely to return migrate followed by Indians, Mexicans, Chinese and then Filipinos. Return migration is largest for Canadians, and then the other four groups tend to be clustered together. Mexican women show a rebound in their reporting of W-2s and 1099s in 2010 and subsequent years; this return to the formal U.S. labor market is sustained over the remaining years, which differs from that of Mexican men in the panel above. Chinese women show a similar pattern to those from Mexico.

Figure 3: Presence of W-2 or 1099 for 2005 Entry Cohort for Ages 25–45 by Country of Birth

Panel A: Men



Panel B: Women



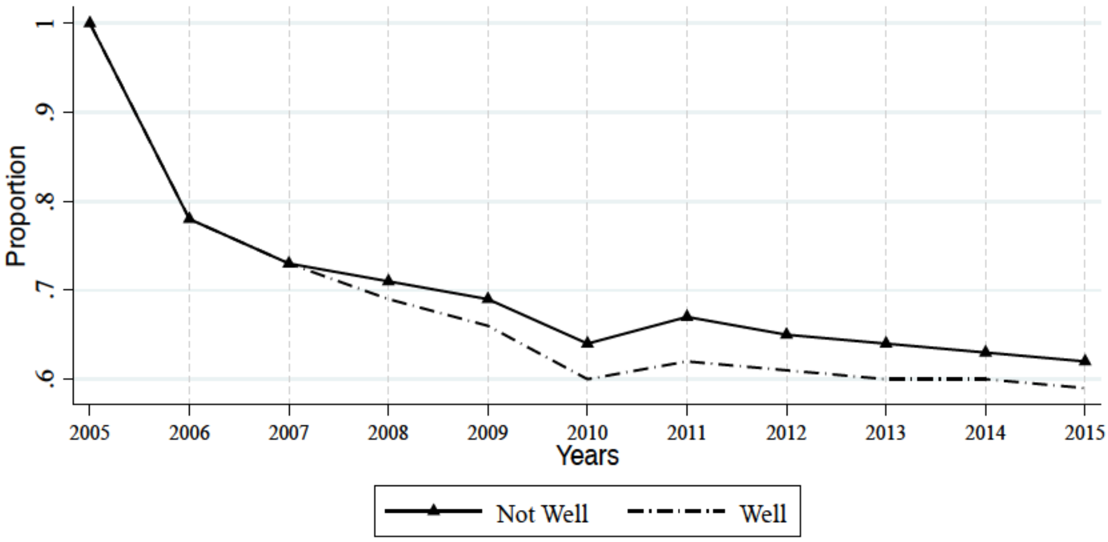
Notes: Each point represents the proportion of each group that is present in the data for each year. We start our analysis in 2005 and take that as the complete immigrant arrival cohort. Source: ACS 2005 and IRS W-2s or 1099 data (2005–2015).

4.4 Return Migration by English Language Ability

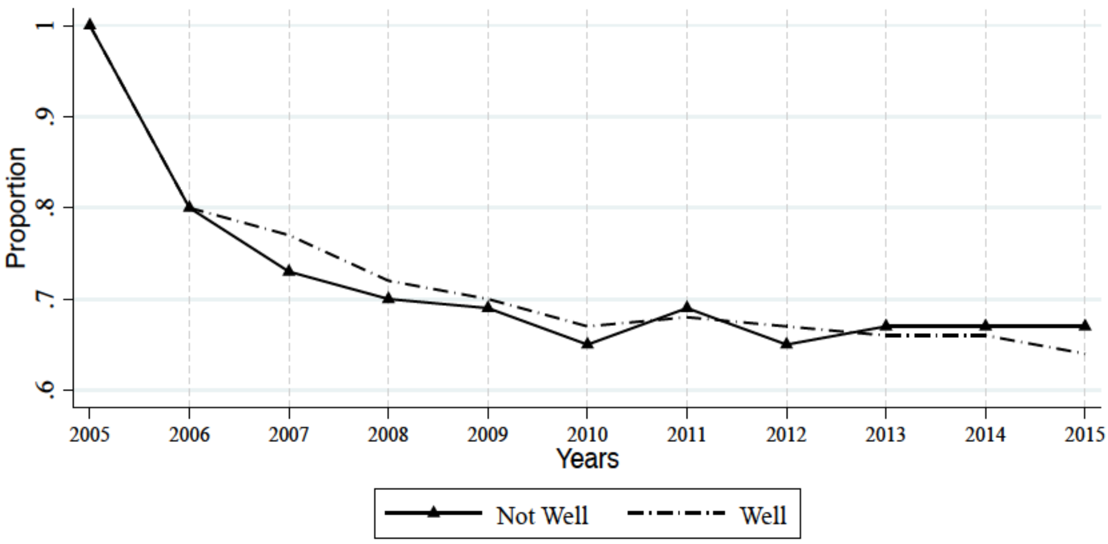
In Figure 4 we show return migration by an individual migrant's English language abilities as reported in the ACS data. We collapse the original four categories into just two to indicate "Speaks English Well" or "Does Not Speak English Well." There is a slight increase in reporting of W-2s or 1099s in 2011 for the group that doesn't speak English well in both panels. For men, there is a slightly higher return of individuals who reportedly speak English well; however, this might be correlated with educational attainment, consistent with what we observed in Figure 2 (the highly-educated are the most likely to return migrate). In Panel B we see that the probability of return migrating doesn't differ between the two categories for women. Overall, return migration doesn't appear to differ by English language abilities for women and differs only marginally so for men.

Figure 4: Presence of W-2 or 1099 for 2005 Entry Cohort for Ages 25–45 by English Language Ability

Panel A: Men



Panel B: Women



Notes: Each point represents the proportion of each group that is present in the data for each year. We start our analysis in 2005 and take that as the complete immigrant arrival cohort. Source: ACS 2005 and IRS W-2s or 1099 data (2005–2015).

4.5 Return Migration by Marital Status

Return migration may also be influenced by an individual's marital status. This might differ by gender in significant ways. In Figure 5 we show the return migration by gender and marital status for the recent cohort of immigrant arrivals. The fact that return migration rates do not differ by marital status for either gender was a surprising finding. Unmarried men were slightly more likely to return migrate in the first few years after arrival in the U.S. than their married counterparts, but after that there is little difference between marital groups over time. For women, there's a slightly higher probability of married women to return migrate in the first few years after arrival, but rates become similar across marital status after about year 5.

Figure 5: Presence of W-2 or 1099 for 2005 Entry Cohort for Ages 25–45 by Marital Status

Panel A: Men

Panel B: Women

Notes: Each point represents the proportion of each group that is present in the data for each year. We start our analysis in 2005 and take that as the complete immigrant arrival cohort. Source: ACS 2005 and IRS W-2s or 1099 data (2005–2015).

4.6 Regression Results for Return Migration by Education and Country of Birth

In Table 3 we present the regression results as specified in equation 1 for men in Panel A and women in Panel B. The outcome variable indicates whether there is a missing W-2 or 1099 by year for an individual based on their educational attainment. We include other control variables in this analysis such as state of residence and age fixed effects; standard errors are clustered at the state of residence in the U.S., according to the ACS. The omitted education category is less than a high school degree. The regression results indicate that more educated immigrants are likely to be missing over time. In Panel A we find that in 2011, all education categories show a higher missing rate compared with the lowest education group, which reflects the finding in Figure 2 that this group had a higher rate of return migration than others in 2010. After 2011, male high school graduates and males with completed college degrees or more (master's or doctorate's), and to a lesser extent those with some college are more likely to be missing in the data than the base group. This is in contrast to earlier years in the time series—specifically, 2007 and 2008, when those with higher educational attainment were less likely to be missing than those from the lowest educational attainment group (although some of those estimates are not statistically significant). In Panel B, the results for women differ from men in that it is only the very highly educated that have a significantly different return migration after 2011, confirming what was found in Figure 2. Overall, these results indicate that the more highly educated are the most likely to return migrate (relative to those with less than a high school diploma) over time.

In Appendix Table A4 we show a similar regression analysis by country of birth. The omitted country category is all other countries (besides the top 5 immigrant-sending countries). As before, we include other control variables in this analysis such as state of residence and age fixed effects; standard errors are clustered at the state of residence in the U.S. In Panel A we find that immigrants from Mexico, India, the Philippines, and China are less likely to return migrate than are immigrants from other countries. Filipinos appear to be much less likely to return migrate at all years as compared to immigrants from other countries. Indian men appear to be less likely to return migrate in the middle years of this ten-year panel. Mexican men are less likely to return towards the end of data years. Chinese men have a statistically significant lower return migration rate only in 2014. Canadian men appear to have an approximately similar rate of return as immigrants from other countries across all years in our data. In Panel B we find somewhat similar results for women from the Philippines and their male counterparts; they are less likely to return migrate in all years over the 10-year period. Mexican women are less likely to migrate towards the second half of the 10-year panel. To a lesser extent there is a reduction in the return migration of Indian women during the middle part of the panel (similar to Indian men) and Chinese women are less likely to return migrate at the end of the panel.

Finally, Appendix Table A5 provides the regression table where we control for initial educational attainment and country of birth and the interaction of those categories. One consistent result emerges from this analysis—the highly educated are the most likely to return migrate. In particular, the highly educated men from Mexico, India, and China are more likely to return migrate towards the end of our time period. Highly educated male immigrants from the Philippines and Canada appear to return migrate at the same rate those from other education groups. For women, there does not appear to be differential return migration across country and education categories that is statistically significant, with the exception of highly-educated Chinese women in the first few years of the data.

Table 3: Missing Tax Data (1099 or W-2) by Year by Educational Attainment

Panel A: Men

VARIABLES	(1) Missing in 2006	(2) Missing in 2007	(3) Missing in 2008	(4) Missing in 2009	(5) Missing in 2010	(6) Missing in 2011	(7) Missing in 2012	(8) Missing in 2013	(9) Missing in 2014	(10) Missing in 2015
High School Degree	0.005	-0.029	-0.040*	-0.029	0.018	0.066***	0.029	0.059**	0.051***	0.055**
	(0.022)	(0.018)	(0.022)	(0.020)	(0.025)	(0.024)	(0.025)	(0.023)	(0.018)	(0.025)
Some College	0.005	-0.052***	-0.025	-0.012	-0.015	0.069***	0.022	0.036*	0.026	0.034
	(0.023)	(0.017)	(0.026)	(0.025)	(0.023)	(0.025)	(0.025)	(0.021)	(0.021)	(0.028)
College Degree	0.024	-0.048**	-0.051**	-0.028	0.012	0.099***	0.081***	0.115***	0.103***	0.086***
	(0.021)	(0.020)	(0.020)	(0.021)	(0.019)	(0.021)	(0.025)	(0.018)	(0.018)	(0.021)
MA or Phd Degree	-0.018	-0.052***	-0.017	-0.001	0.037*	0.142***	0.111***	0.137***	0.127***	0.118***
	(0.020)	(0.019)	(0.019)	(0.025)	(0.020)	(0.018)	(0.020)	(0.017)	(0.019)	(0.023)
Constant	0.338***	0.443***	0.480***	0.397***	0.523***	0.234***	0.345***	0.257***	0.300***	0.414***
	(0.044)	(0.021)	(0.020)	(0.025)	(0.030)	(0.026)	(0.033)	(0.026)	(0.033)	(0.025)
Observations	5,400	5,400	5,400	5,400	5,400	5,400	5,400	5,400	5,400	5,400
R-squared	0.123	0.025	0.023	0.018	0.018	0.026	0.028	0.033	0.025	0.023

Note: Includes age fixed effects, year of entry, and state fixed effects. Standard errors are clustered at the state of residence. Omitted educational attainment category is less than high school. Source: American Community Survey, 2005–2007 and IRS 2005–2015 W-2 and 1099 data. Numbers have been rounded to comply with the Census Bureau's disclosure-avoidance guidelines.

Panel B: Women

VARIABLES	(1) Missing in 2006	(2) Missing in 2007	(3) Missing in 2008	(4) Missing in 2009	(5) Missing in 2010	(6) Missing in 2011	(7) Missing in 2012	(8) Missing in 2013	(9) Missing in 2014	(10) Missing in 2015
High School Degree	-0.033	0.018	-0.002	0.011	0.043*	0.054*	0.015	0.037	0.018	0.037
	(0.035)	(0.021)	(0.024)	(0.028)	(0.023)	(0.031)	(0.025)	(0.032)	(0.037)	(0.030)
Some College	0.003	0.063**	0.023	0.017	0.025	0.030	0.013	0.019	0.003	0.022
	(0.027)	(0.025)	(0.029)	(0.037)	(0.025)	(0.027)	(0.026)	(0.022)	(0.032)	(0.031)
College Degree	-0.036	0.027	-0.040	-0.033	-0.005	0.003	0.001	0.000	0.007	0.037
	(0.027)	(0.021)	(0.024)	(0.033)	(0.018)	(0.025)	(0.021)	(0.019)	(0.026)	(0.026)
MA or Phd Degree	-0.029	0.014	0.019	0.038	0.085***	0.110***	0.079***	0.085***	0.107***	0.128***
	(0.033)	(0.027)	(0.025)	(0.033)	(0.024)	(0.023)	(0.024)	(0.021)	(0.031)	(0.031)
Constant	0.286***	0.383***	0.458***	0.464***	0.420***	0.346***	0.442***	0.379***	0.286***	0.295***
	(0.044)	(0.026)	(0.027)	(0.032)	(0.028)	(0.027)	(0.033)	(0.031)	(0.035)	(0.031)
Observations	4,200	4,200	4,200	4,200	4,200	4,200	4,200	4,200	4,200	4,200
R-squared	0.097	0.036	0.033	0.027	0.026	0.035	0.038	0.042	0.039	0.033

Note: Includes age fixed effects, year of entry, and state fixed effects. Standard errors are clustered at the state of residence. Omitted educational attainment category is less than high school. Source: American Community Survey, 2005–2007 and IRS 2005–2015 W-2 and 1099 data. Numbers have been rounded to comply with the Census Bureau's disclosure-avoidance guidelines.

5 Immigrant Earnings Assimilation Relative to Native-Born Population

In Figure 6, we provide the results from the earnings assimilation analysis. This analysis is based on equation 1 where we regress the log of earnings on immigrant status, age fixed effects, and state of residence fixed effects; we plot the estimated coefficient on the immigrant variable for all years. The black dotted line in Panel A for men reports the results from a standard repeated cross-section analysis, which includes all immigrants present in the data at each time period. This line provides the difference in log earnings of immigrants and natives in the U.S. in each year. This line crosses the horizontal line at 0, indicating earnings assimilation, by 2012 or about 7 years after arrival on average. The dotted line in Panel B shows that there is almost convergence for immigrant women by 2012 to the native-born; however, the line diverges in the years after. We do not include confidence intervals in these figures; however, the observed results are statistically significant from zero except for instances where the estimated results are near to the zero line.

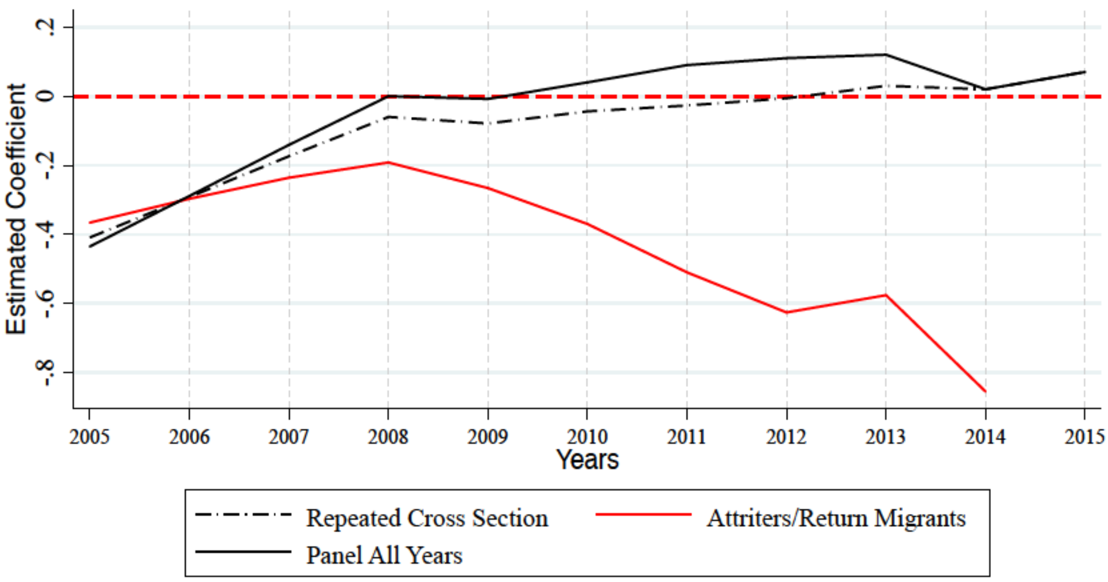
The solid black line "Panel All Years" provides the earnings differences in logs between the immigrants from the 2005–2007 entry cohorts to that of the native born and removes individuals who will return migrate at some point over this time period (2006–2014). This group comprises individuals who remain in the U.S. for all of the years in the study. This analysis replicates the stock-based panel data analysis which links either the CPS or SIPP data (for immigrant survivors of older entry cohorts) to their historical SSA Detailed Earnings Records (see Duleep and Regets (1997); Lubotsky (2007)). Note that the final years (2014 and 2015) necessarily converges to that of the dashed black line (repeated cross-section) as we have no further information on whether an individual will return migrate in future time periods; this would require data from 2016 for instance. This line for both panels shows a quicker convergence to that of the native-born earnings in about 3 years for men and 6 years for women. This provides some evidence that the returning migrants are negatively selected with regard to earnings; at all years those who remain for all time periods have consistently higher earnings than those in the repeated cross-section data.

The bottom line in both panels shows this result explicitly. The red line in both figures provides similar analysis for the individuals who do not remain in the U.S. for all time periods; these are the attriters—or those who emigrate. A description of how we coded these individuals can be found in Table 2. To create the red line we aggregate all of the individuals that remain in the U.S. for less than the full time period of 10 years; this is the complementary set of observations to the solid black line called "Panel All Years" that, when combined, forms the repeated cross-section dotted line in both panels. We find in Panel A that there is evidence of downward earnings mobility for this group of immigrants starting about three years after arrival on average. At the time of arrival the starting earnings for all three groups are relatively similar to one another, but this group of return migrants in particular appear to have a series of negative shocks or experiences in the labor market, different from everyone else in the labor market, that help to increase their probability of leaving the U.S. We present the same analysis in Panel B for women. In this case, the return migrants appear to start out at a significant reduction in their initial earnings compared to those immigrants who stay in the U.S. for the entire time period. They, too, experience downward earnings mobility starting in the third year after migration on average.

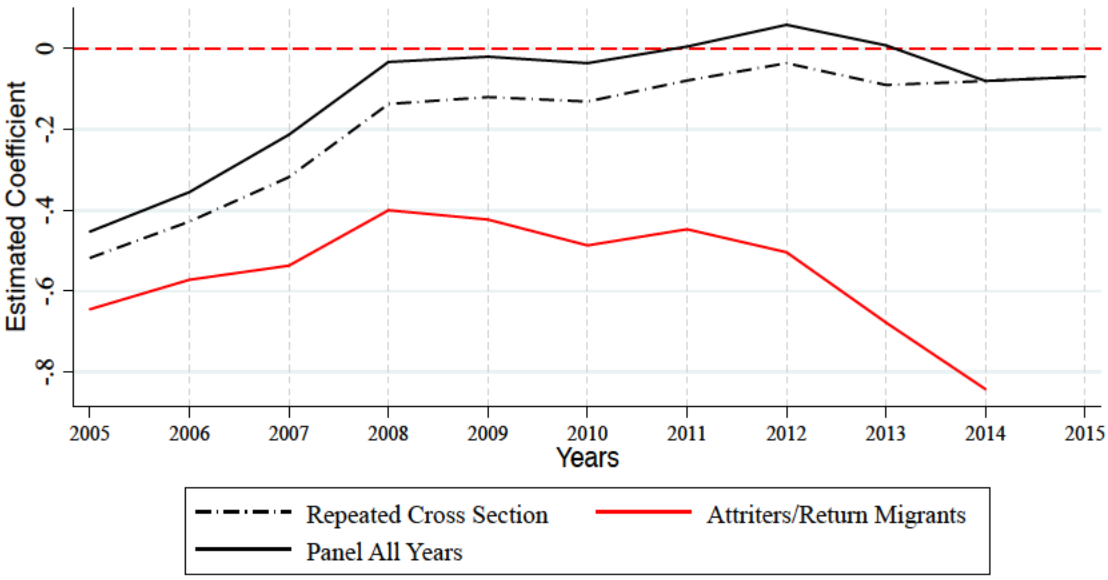
In Appendix Figure A3, we conduct the same analysis and directly control educational attainment. We find that the figures remain qualitatively similar in shape and that there is a general shift downward of all lines. In Panel A, we find that initial earnings for men are slightly higher for the panel observations as compared to the repeated cross-section observations, and the attriters have the lowest initial relative earnings. For women in Panel B, the initial earnings look approximately the same as in Figure 6.

Figure 6: Log Immigrant-Native Annual Earnings, Ages 25–45

Panel A: Men



Panel B: Women



Notes: Each point represents the estimated coefficient on an immigrant indicator variable in a log wage regression conducted separately for each year for each of the three separate groups relative to the native-born population. We include additional control variables in the regression such as state of residence fixed effects and age fixed effects. Source: ACS 2005–2007 and IRS W-2s or 1099 data (2005–2015).

5.1 Immigrant Earnings Assimilation for Return Migrants by Duration of Stay in the U.S.

In Figure 7, we separate out the relative earnings for the return migrants by their duration of stay in the U.S. These lines are the disaggregated form of the single attrition line shown in the red line in Figure 6. Of note is that those who leave have quite low earnings in their year prior to emigration. In many cases, earnings have also experienced a decrease in two years prior to emigration. We believe that this is the first time this feature of immigrant outcomes has been observed for return migrants from the U.S. Overall, negative earnings shocks appear to play an important role in determining return migration for recent arrival cohorts. Note that this downward earnings mobility is a consistent predictor of return migration across all of our years; this is not simply an artifact of the Great Recession as there are downward trends for this group prior, during and after the Great Recession.

In Appendix Figure A4 we decompose the return migrants by two broad education categories—those with a high school degree or less and those with some college or more. The results show that there is a downward trajectory in earnings for both groups that end up return migrating. There is a level difference for those with only a high school degree or less as compared to the other education category for men, but not for women. Overall, it appears that the return migration isn't being driven solely by low-education immigrants. There are relatively highly educated immigrants that experience downward earnings mobility and opt to return migrate.

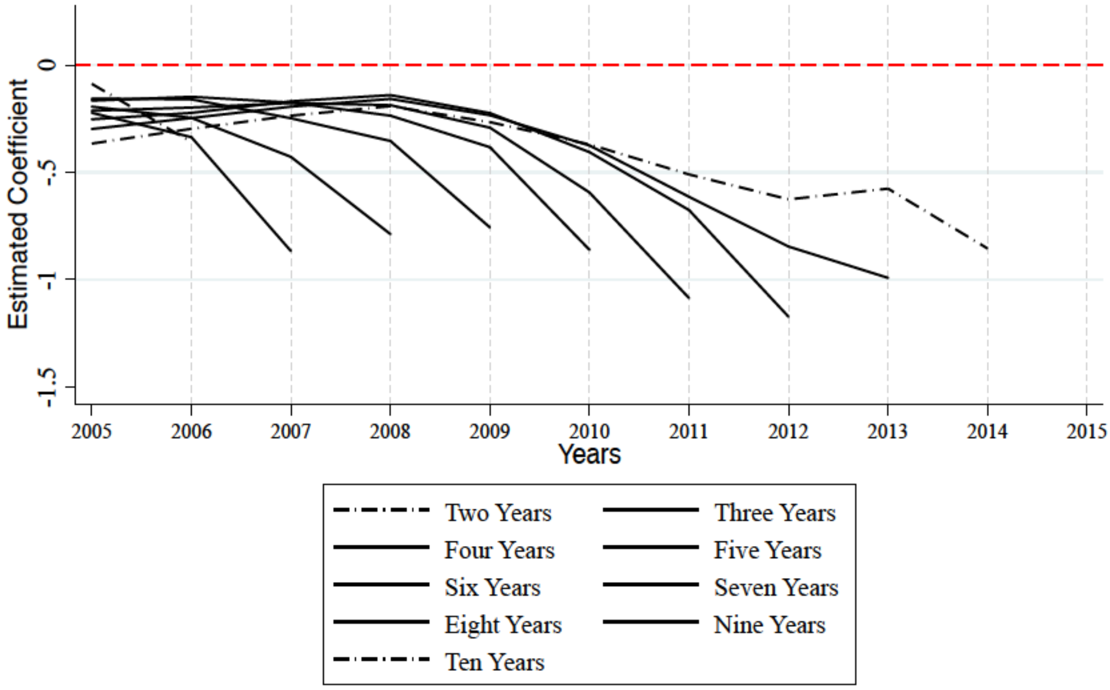
In Appendix Figure A5 we decompose the return migrants by four occupation categories. We observe occupation in the ACS and not in the W-2 data; thus the occupation identified is what the ACS respondent reported at entry into the U.S. The most pronounced downward earnings mobility exists for sales or office occupations and management occupations for men and production occupations for women. In the other categories earnings are mostly stable across time for return migrants who are at least initially employed in these occupations.

In order to explore whether an adverse economic shock affects those who return migrate we investigate one potential shock—job loss or change. In Table 4 we analyze whether return migrants are more likely to have changed employers over time, as identified from the Employer Identification Number (EIN) listed on the W-2. In theory, of course, a change in employers could indicate that a person has found a better employment match and increased earnings. On the other hand, it may also signal a job loss that requires finding another employer (especially if an immigrant is in the U.S. on a work visa) and having to take a reduction in earnings. The estimation reported in Table 4 regresses a binary indicator variable for whether there is a difference in the EINs reported across adjacent years on indicator variables for years since migration and return migration. EINs are unique for employers, and any differences will indicate a change in employer by year.[14] The interaction variable between years since migration and return migration is the variable of interest; this tells us whether an individual that will eventually return migrate and stays for one year longer will change her employer. We also include controls such as age, state of residence, year of entry, and a constant. Standard errors are clustered at the state of residence.

[14] A certain caveat applies: While the EIN is unique to an employer, large employers who have multiple establishments may or may not use a single EIN for all establishments. Employers who make use of franchises are especially prone to multiple EIN use, and a move from one EIN to another may only indicate a change in franchise location and not necessarily employer (but may still indicate a job loss at the original franchise location).

Figure 7: Log Immigrant-Native Annual Earnings for Attriters Ages 25–45

Panel A: Men



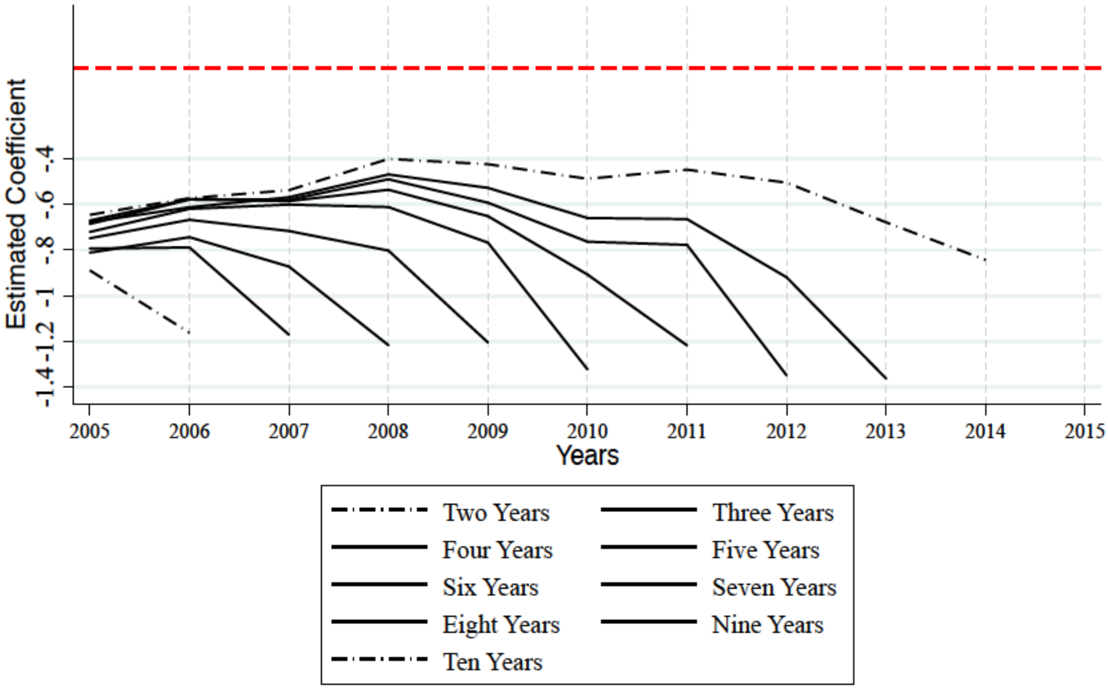
The regression results indicate that, for both males and females, the signs of the estimated coefficients on years since migration and being a return migrant are negative and not always statistically significant. These results are not surprising and indicate that the shorter one stays in the U.S. the lower the chances there are of changing an employer. The estimated coefficient of interest, however, is the interaction of these two variables. This is shown in the first row of Table 4. For men the estimated coefficient is marginally statistically significant, indicating that return migrants who stay longer in the U.S. are more likely to have changed employers. The results are positive for women but not statistically significant. This provides some evidence that job changes (perhaps due to firing) may be a determinant of male immigrant downward earnings mobility that eventually predicts return migration within the first 10 years after arrival. For women, we do not observe a similar direct relationship in the available data.[15]

5.2 Immigrant Earnings Assimilation by Initial Educational Attainment

In Figure 8, we present a similar analysis for the earnings assimilation by an individual's own educational attainment at the time of entry into the U.S. We employ only the panel data, which contain the individuals who remain in the U.S. for the entire time period. The repeated cross-section analysis is

[15] We have also conducted a separate analysis where we include a country of birth fixed effect for the top 5 immigrant sending countries and group the rest in a single category. We have also clustered the standard errors at the country of birth by state of residence categories. These results provide very similar results to those shown in Table 4 and are available upon request. Additionally, we provide simple ordinary least squares regressions in Appendix Table A7 which show a similar result when coding for the number of changes in employment by return migrant status and gender.

Panel B: Women



Notes: Each point represents the estimated coefficient on an immigrant indicator variable in a log wage regression conducted for the group of attriters relative to the native-born population. We include additional control variables in the regression such as state of residence fixed effects and age fixed effects. We separate out the return migrants by their years in the U.S. separately. Note that it is not possible in this data to individuals who have stayed for 11 years as there is no twelfth year of data to indicate the return migration. Source: ACS 2005–2007 and IRS W-2s or 1099 data (2005–2015).

Table 4: Probability of Changing Employer for Immigrants by Length of Stay in U.S.

VARIABLES	Men (1) Changed Employer	Women (2) Changed Employer
Return Migrant x	0.011	0.007
Years Since Migration	(0.006)	(0.006)
Return Migrant	-0.067	-0.040
	(0.031)	(0.036)
Years Since	-0.004	-0.007
Migration	(0.003)	(0.002)
Person Years	15,500	12,500
R-squared	0.018	0.016

Source: American Community Survey, 2005–2007 and IRS 2005–2015 W-2 and 1099 data. Note: Includes age, state fixed effects, education controls, a constant and year fixed effects. The data is transformed to a single observation per individual per year and we report person years. Standard errors are clustered at the state of residence. Numbers have been rounded to comply with the Census Bureau's disclosure-avoidance guidelines.

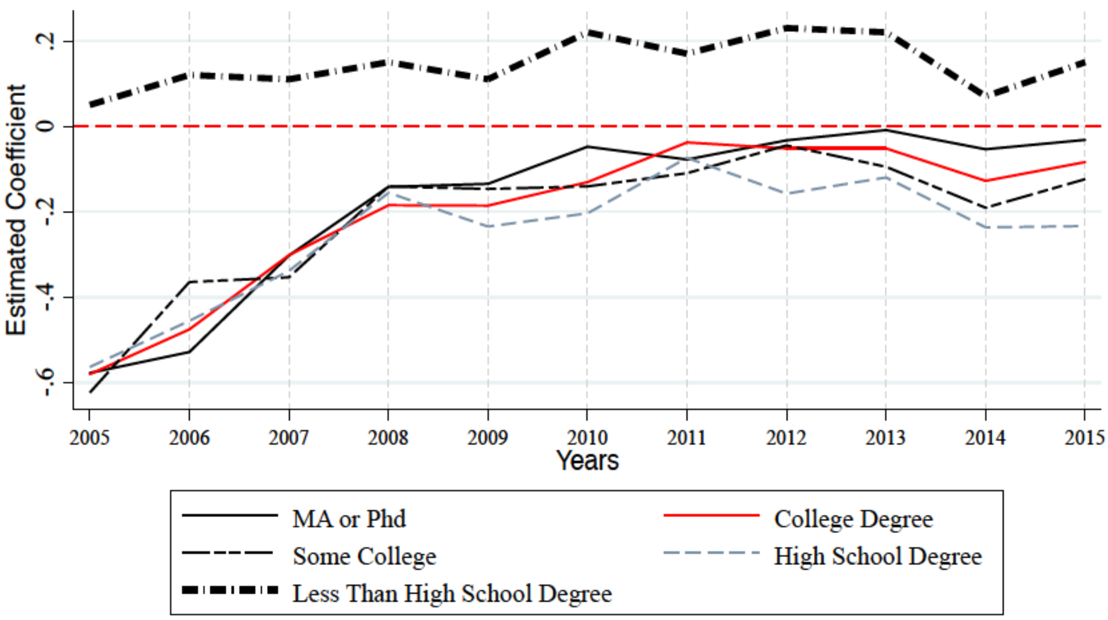
presented in Appendix Figure A6 and is qualitatively similar. Thus, the analysis for the next set of figures answers the question: what is the earnings assimilation of immigrants who remain in the U.S.

for a prolonged period of time by a specific initial characteristic? In each of the subsequent figures we will examine this by different individual-level characteristics of the immigrant. We control for age and state of residence fixed effects in all analyses.

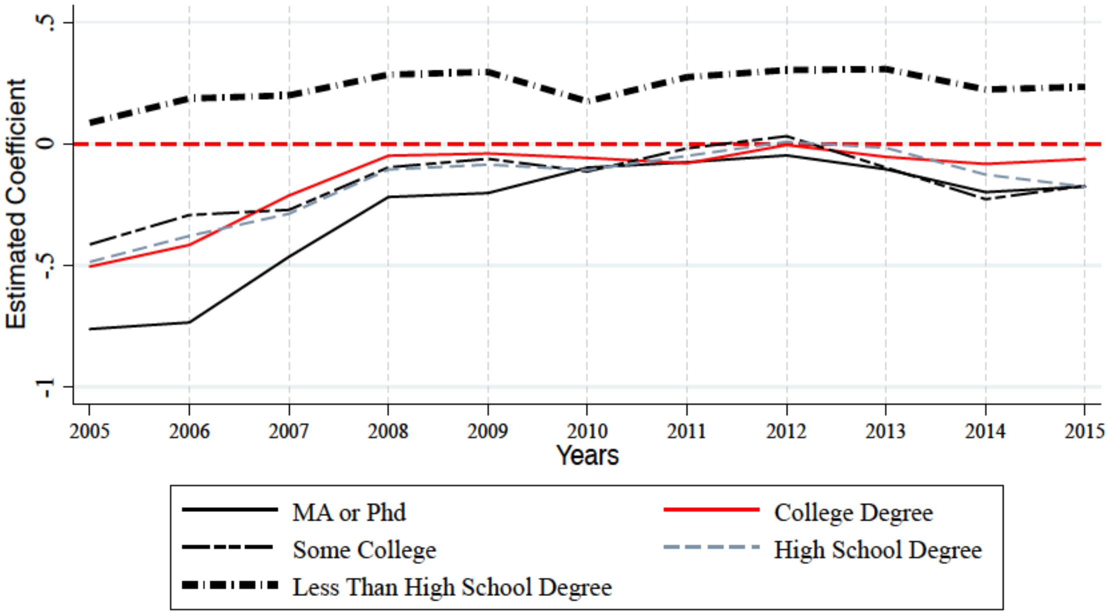
The results indicate a relatively quick assimilation for all educational levels for men in Panel A; while immigrant earnings do not reach parity with the earnings of the native-born it is quite close for those with the highest levels of education. We estimate only a 6 percent difference in earnings between immigrants and native-born workers. This suggests that an important predictor of remaining in the U.S. is a relatively quick (within 8–9 years) earnings assimilation experience. Note that the estimate for individuals with less than a high school education is always above the native-born earnings estimates for all years. Because these immigrants with lower educational attainment are working in the formal labor market–which may indicate that their skills for a particular industry are measurable by something other than a formal degree–comparisons within this group may indicate large difference in abilities (lower skills for the U.S. population in this education category than for those from other countries). Additionally, there is convergence in earnings around the depths of the Great Recession and then a slight divergence after that for those immigrants with less educational attainment. The results for women in Panel B display a more rapid convergence to that of native-born women in the first three to four years after arrival for all education groups other than those with less than a high school degree. After the Great Recession, however, there is some divergence as immigrant earnings decline relative to the native-born. The group with less than a high school education is consistently above that of the native-born, just as for men.

Figure 8: Log Immigrant-Native Annual Earnings for 2005-2007 Arrival Cohorts Ages 25–45 by Educational Attainment in Panel Data

Panel A: Men



Panel B: Women



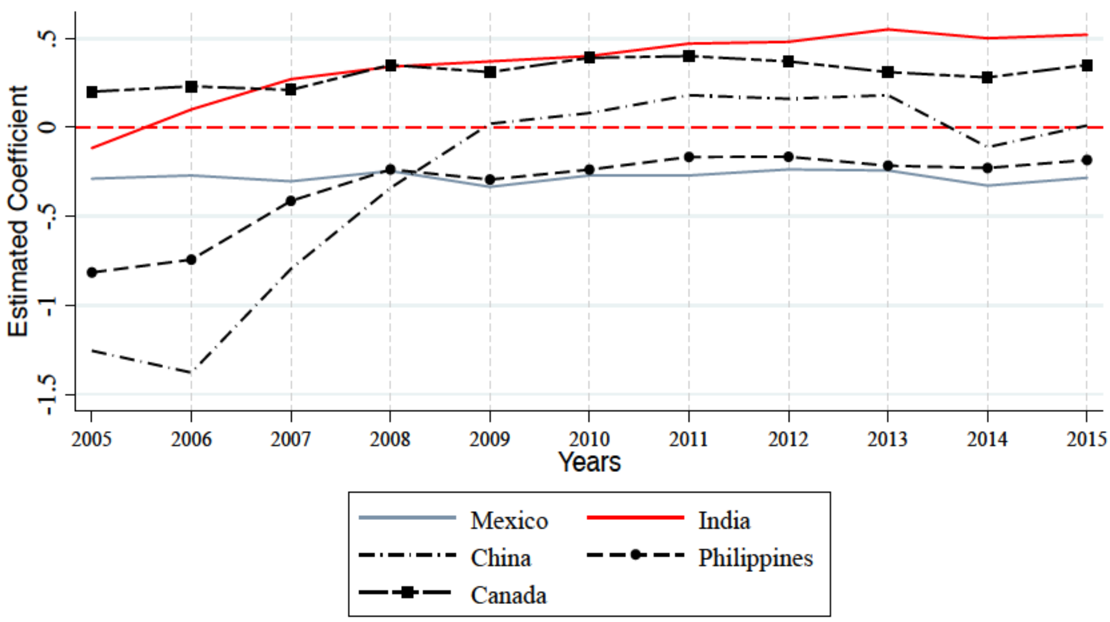
Notes: Each point represents the estimated coefficient on an immigrant indicator variable in a log wage regression conducted separately for each year for each of the five educational attainment groups relative to the native-born population. We include additional control variables in the regression such as state of residence fixed effects and age fixed effects. Source: ACS 2005–2007 and IRS W-2s or 1099 data (2005–2015).

5.3 Immigrant Earnings Assimilation by Country of Birth

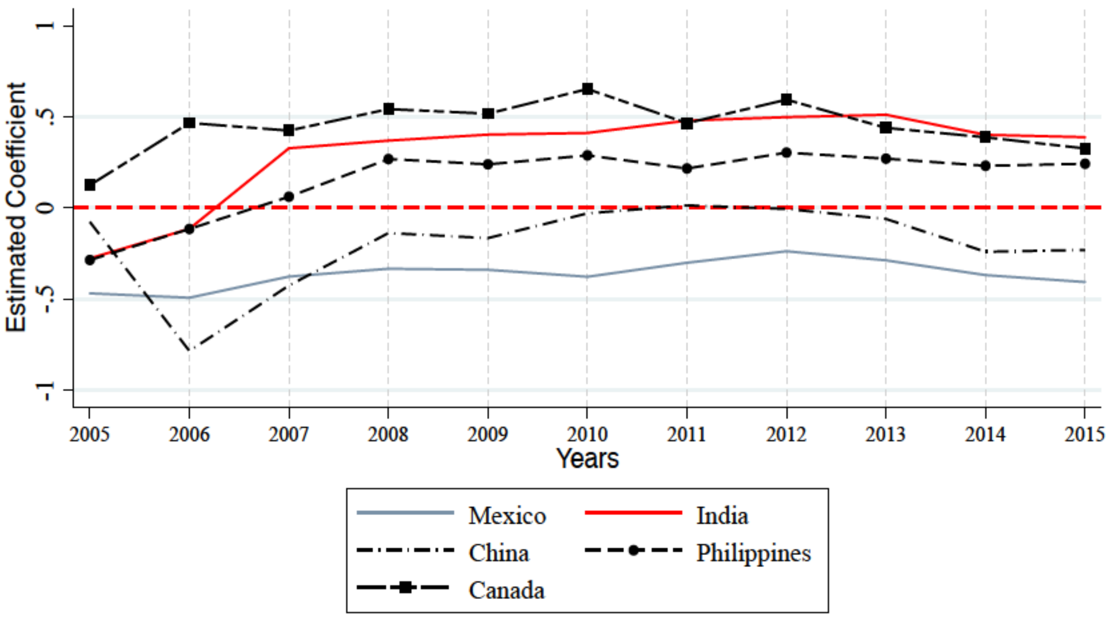
In Figure 9 we show the earnings assimilation for individuals from the top five immigrant-sending countries in an analysis similar to that described in the previous section. We find that Canadians have earnings that are above those of native-born workers in all years for both men and women. By the end of the period, Indian men have the highest earnings (Panel A), followed by Canadians, Chinese, Filipinos, and Mexicans. Mexicans have a very flat earnings profile on average, and the average is consistently below that of the native-born in all years; it is approximately 30 percent lower than native-born earnings. Chinese immigrants tend to make some of the largest earnings gains especially in the first five years: their earnings start at a significant earnings disadvantage but ultimately exceed those of the native-born. Filipinos realize earnings gains over time, but they do not fully converge to the native-born by the final year in the data. For women, reported in Panel B, the results show that Canadians, Indians, and Filipinos experienced rapid earnings gains relative to the native-born shortly after arrival in the U.S. By year two or three all of these immigrants groups have earnings estimates that lie above the native-born estimate. Chinese immigrant earnings converge to those of native-born by year five, but they experience a downward trend after that. Mexican women have a flat earnings profile across all years; their earnings remain at approximately 30–50 percent below those of the native-born in all years. We also show in Appendix Figure A7 the same analysis for the panel data, where we control for a person's educational attainment. The results generally show a downward shift of the observed curves for all countries of birth shown in Figure 9. We show an analysis analogous to that shown in Figure 9 for the repeated cross-section in Appendix Figure A8.

Figure 9: Log Immigrant-Native Annual Earnings for 2005-2007 Arrival Cohorts Ages 25–45 by Country of Birth in Panel Datal Data

Panel A: Men



Panel B: Women



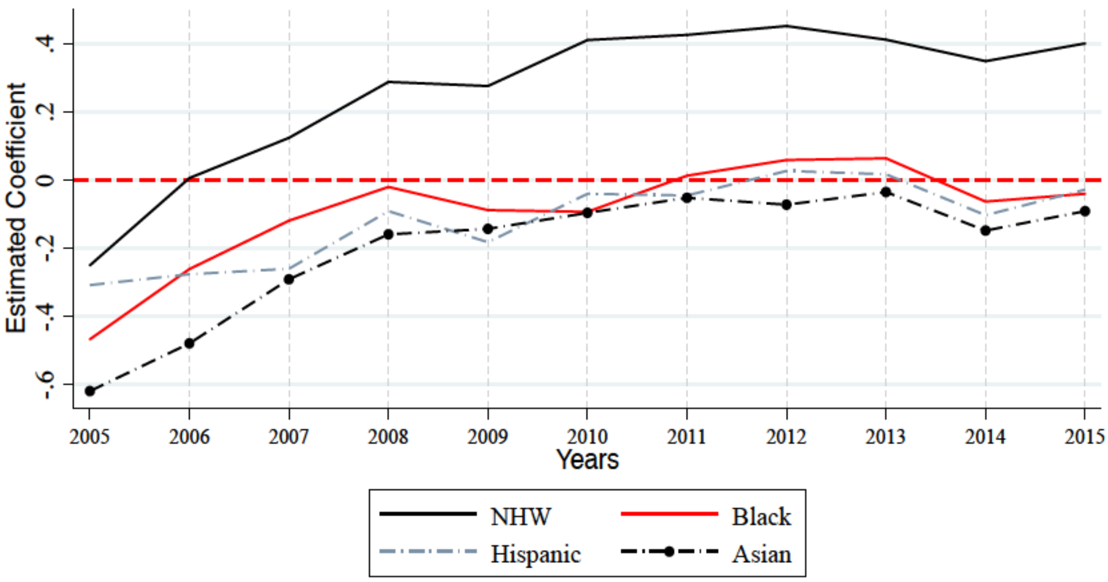
Notes: Each point represents the estimated coefficient on an immigrant indicator variable in a log wage regression conducted separately for each year for each of the five country of birth groups relative to the native-born population. We include additional control variables in the regression such as state of residence fixed effects and age fixed effects. Source: ACS 2005–2007 and IRS W-2s or 1099 data (2005–2015).

5.4 Immigrant Earnings Assimilation by Race

In Figure 10 we separate out the earnings assimilation for the immigrants by their race or ethnic group; all comparisons are with the same race or ethnic groups for the native-born.[16] We show in Appendix Figure A9 the same analysis with the cross-section data for comparison. In contrast to the educational-attainment results, we find that there is strong earnings assimilation within race and ethnic groups for males in Panel A. The one race group that far exceeds the earnings of the native-born is the non-Hispanic white category. Our results are similar to that found by others; Villarreal and Tamborini (2018) find that Black and Hispanic immigrants have slower earnings assimilation than non-Hispanic White immigrants. However, these race groups are able to catch up to earnings within their own race in native-born populations over time. Our additional contribution is showing the earnings assimilation for the Asian race group which is not often reported; the results indicate that there isn't a complete convergence in earnings to native-born Asians, perhaps due to the fact that the native-born population in this category is, on average, a high-earning group. For women in Panel B there are more differences across race and ethnic groups in terms of earnings across time. Asian and Hispanic female immigrants earn significantly less then their native-born co-ethnics over all years. On the other hand, non-Hispanic Whites and Blacks achieve some earnings parity with their native-born counterparts in the first three to four years after arrival. These results highlight the differences across race and gender in immigrant earnings experiences in the U.S. labor market.

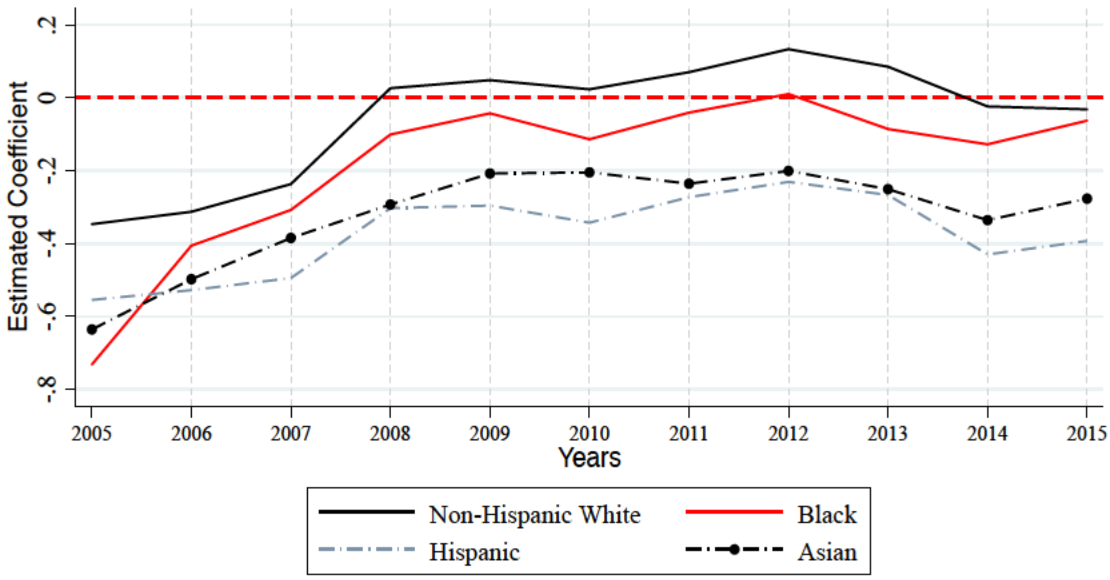
Figure 10: Log Immigrant-Native Annual Earnings for 2005–2007 Arrival Cohorts Ages 25–45 by Race in Panel Data

Panel A: Men



[16]Note that we omit the Alaska Native and American Indian category as well as the Native Hawaiian categories, which are standard race and ethnic groups in U.S. Census data. There are few if any of these individuals who are born abroad by definition and are not classified as immigrants.

Panel B: Women



Notes: Each point represents the estimated coefficient on an immigrant indicator variable in a log wage regression conducted separately for each year for each of the four race and ethnic groups relative to the native-born population. We include additional control variables in the regression such as state of residence fixed effects and age fixed effects. Source: ACS 2005–2007 and IRS W-2s or 1099 data (2005-2015).

6 Immigrant Earnings Growth in Repeated Cross-Sections, Stock-Based Panel and Longitudinal Data Settings

In this section, we examine the role that years spent in the U.S. has on the earnings growth of an immigrant arrival cohort, which we define here as the pooled cohorts between 2004 and 2007 and captured in the 2005, 2006, and 2007 ACS. The focus on this pooled arrival cohort allows us to eliminate earlier arrival cohorts that may already have been subject to selective return migration. It allows us to compute estimates for a single cohort of arrivals over their first decade in the U.S. in a complete manner given the panel nature of the data. Therefore, we do not focus on differences in average cohort quality across the four arrival years since these are clustered so close to one another.

In Table 5 we pool all observations of newly arrived immigrants from the 2005-2007 ACS and regress the log earnings on the number of years that an individual has been present in the U.S. The outcome variable is the change in earnings in column one. We include state of residence and age fixed effects in this analysis; standard errors are clustered at the state of residence. The coefficient of interest—years since migration—is a count variable for the foreign-born population and set to zero for the native-born. The estimated coefficient in column 1 in Panel A for men is positive and statistically significant, indicating that a longer duration in the U.S. will result in greater earnings growth. In particular, these results show that each additional year of residing in the U.S. will increase earnings by 5.2 percent. These results are qualitatively similar to those found by others using cross-section data. In columns 2–9, we provide the estimated coefficient on years since migration in that year. The estimated coefficients are all approximately similar to the estimated coefficient in column 1. An analogous analysis is provided for female immigrants in Panel B. The results show that there is about a 3.4 percent increase in earnings by year in the U.S. in column 1. The duration of years in the U.S. is provided in columns 2–9.

Table 5: Earnings Growth by Year for Repeated Cross-Section

Panel A: Men

VARIABLES	(1) Earnings	(2) Earnings	(3) Earnings	(4) Earnings	(5) Earnings	(6) Earnings	(7) Earnings	(8) Earnings	(9) Earnings
Years Since Migration	0.052***	0.052***	0.052***	0.045***	0.045***	0.045***	0.043***	0.043***	0.042***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Sample Years	Repeated Cross Section All Years	3 Years or More	4 Years or More	5 Years or More	6 Years or More	7 Years or More	8 Years or More	9 Years or More	10 Years or More
OLS	Y	Y	Y	Y	Y	Y	Y	Y	Y
Person Years	15,500	15,500	15,500	15,500	15,500	15,500	15,000	14,500	14,000
R-squared	0.199	0.199	0.199	0.200	0.204	0.208	0.215	0.223	0.227

Panel B: Women

VARIABLES	(1) Earnings	(2) Earnings	(3) Earnings	(4) Earnings	(5) Earnings	(6) Earnings	(7) Earnings	(8) Earnings	(9) Earnings
Years Since Migration	0.034***	0.034***	0.034***	0.023***	0.021***	0.022	0.021***	0.018***	0.015***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Sample Years	Repeated Cross Section All Years	3 Years or More	4 Years or More	5 Years or More	6 Years or More	7 Years or More	8 Years or More	9 Years or More	10 Years or More
OLS	Y	Y	Y	Y	Y	Y	Y	Y	Y
Person Years	12,500	12,500	12,500	12,500	12,500	12,500	12,000	12,000	11,500
R-squared	0.154	0.154	0.154	0.156	0.159	0.161	0.162	0.169	0.181

Source: ACS 2005–2007 and IRS 2005–2015 W-2 and 1099 data. Note: Includes age, state fixed effects, a constant, and year fixed effects. The data have been transformed so that there is an observation for each individual in each year and we report person years. Standard errors clustered at state of residence. Numbers have been rounded to comply with the Census Bureau's disclosure-avoidance guidelines.

In Table 6, we show both the stock-based panel regression results for the years since migration as well as the selection-corrected regression that is possible using the complete panel data for male and female immigrants. In the first and third columns, we replicate the standard analysis conducted by recent researchers using stock-based panel data for males and females respectively. We use the survivors from an arrival cohort and their earnings history to identify the earnings growth over their subsequent time in the U.S. We find that, using an individual fixed-effects regression, the years since migration is still positive and statistically significant; however, it is about 2.7 (1.2) percent per year compared to the 5.2 (3.4) percent per year found in the repeated cross-section analysis in Table 5 for men (women). This is the standard result that has been found by many other researchers in regard to earnings growth between repeated cross-section and stock-based panel data.

In columns 2 and 4, we employ a Heckman selection-correction to the analysis as described in Section 3.3.4.[17] This type of selection-correction can only be conducted when the researcher knows exactly which individuals in the panel data will end up leaving. We use the individual earnings from two periods prior to the departure date (as suggested by Dustmann and Görlach (2015)) as the exclusion restriction for each time period. Then we compute the inverse Mills' ratio and include this correction term into the pooled OLS regression model. The estimated coefficient on years since migration is then presented in columns 2 and 4 (for men and women respectively) in the table below. The key difference between this analysis and the one shown in column 1 and 3 is that we now relax the assumption that the disturbance terms in the earnings equation and the emigration decision equation are orthogonal to one another. Allowing them to be jointly determined results in an even larger estimated coefficient on the years since migration. The estimated coefficient is 6 percent per year for men and 3.3 percent for women. The estimated coefficient is only marginally statistically significant for men and does not reach conventional levels of significance for women in this corrected analysis. Overall, the results for men, at least, indicate that the earnings growth estimated from repeated cross-section analysis may not be overstating the results at all if there is downward earnings mobility for the return migrants.

7 Discussion of Results for Immigrant Selection

Our analysis reveals some novel insights into the determinants of return migration and its relationship to the earnings histories of the immigrants. The first finding is that many immigrants return migrate within a decade of entry, during which almost 40 percent of the arrival cohort has emigrated. Given the immigrant population that we study—i.e., immigrants who are employed in the formal labor market—they presumably have a work visa or are in the U.S. on a family-reunification visa. In light of this, a prior expectation might be that these workers would have had a strong incentive to remain in the U.S. Their high return rates suggests that important unobserved characteristics play a role in the return migration decision of many immigrants. It is especially striking that we find the more highly educated return at higher rates than their lower-educated counterparts, contrary to the expectation that highly educated workers have the most lucrative options and thus the greatest incentive to remain in the U.S. The fact that these individuals are leaving may indicate that other factors play an important role in return migration.

[17] We provide the regression equations for the Heckman selection-correction in Appendix Table A8. The exclusion restrictions are statistically significant at the 1% level for all years 2008-2015; indicating that the earnings lagged by two years is a strong instrument.

Table 6: Earnings Growth by Year for Stock-Based Panel and Completed Panel Data

	Men		Women	
	(1)	(2)	(3)	(4)
VARIABLES	Earnings	Earnings	Earnings	Earnings
Years Since Migration	0.027***	0.060*	0.012***	0.033
	(0.003)	(0.032)	(0.003)	(0.019)
Sample Years	Stock Based	Selection Corrected	Stock Based	Selection Corrected
Individual FE	Y	N	Y	N
Person Years	15,500	15,500	12,500	12,500
R-squared	0.028	0.021	0.012	0.013

Note: Column 2 includes the inverse Mills' ratio for each year. A joint test of statistical significance provides an F-test(100, 15,500)=3.90 for men and F-test(100, 12,500)=1.95 for women which indicates that we can reject the null hypothesis that the IMR variables have no effect on the outcome for men and women respectively. Source: ACS 2005–2007 and IRS 2005–2015 W-2 and 1099 data. Numbers have been rounded to comply with the Census Bureau's disclosure-avoidance guidelines.

One important result that we document is the downward earnings trajectory of the group of immigrants who eventually return migrate. It is well known that employer-employee match plays a strong role in business success and the continued employment of the employee. A mis-match may be particularly difficult for an immigrant who arrives in the U.S. with a specific work visa and employer. Firing by an employer makes it highly likely that the immigrant will have to return home unless he (or she) can find an employer willing to support the work visa. We find in Table 4 that return migrants are more likely to have a different employer over time, even controlling for years since migration, than the migrants who stay in the U.S. for the entire span of our data. This suggests that employment transitions influence return migration, possibly due to firing.

Unobserved time-variant characteristics also likely play an important role in determining the immigrant earnings success. Previous research highlighted on the importance of cultural and social assimilation of immigrants in the U.S. While we have no measures of these characteristics over time, it is possible that non-cognitive-type skills are important in determining employment relationships and success. As noted earlier, these characteristics also must be related to the time spent in the U.S. Therefore, immigrants who remain in the U.S., holding observed characteristics constant, must be those who are most flexible and integrative with respect to the U.S. labor force.

Selective return migration plays an important role in determining earnings assimilation and growth for immigrants. Over time as return migration creates an increasingly high-earning stock of immigrants, earnings assimilation and growth estimates will be upwardly biased. However, if return migration is a standard component of all immigrant arrival cohorts, then perhaps this is the measure that we should be concerned with. Much of the immigration research has been focused on estimating earnings assimilation and growth for a new (entire) arrival cohort of immigrants. However, if it is a standard feature of immigrant arrival cohorts that a relatively large number of immigrants return migrate in the first few years after arrival, it might be more important to estimate earnings for those who remain. This argues in favor of using the stock-based panel of immigrants (in the absence of complete panel data).

On the other hand, stock-based panels can underestimate earnings growth for immigrants if there is time-variant unobserved characteristics that jointly determine earnings and the decision to return migrate. In our analysis, we find this to be the case. It remains to be seen if this is a general condition that persists over different time periods and shocks. Our analysis covers a period including the Great Recession, which was an extreme outlier in U.S. economic history. Therefore, in another period, the importance of time-variant unobserved characteristics may actually play a less important role. Further empirical analysis with complete panel data will shed light on this question.

8 Conclusion

This research uses a novel data set composed of linked survey and administrative data to create a complete panel dataset for a recent arrival cohort of immigrants to the U.S. The linked ACS-IRS data provide information on important characteristics of the population such as age, year of arrival, country of birth, and English language abilities while the IRS data provide a useful panel of earnings linked at the person level. Typically, this type of data does not exist for the U.S. Previous researchers have used stock-based panel data which is a complete panel only for those immigrants who have survived up to a particular point in time. Our data allow us to identify individuals who end up leaving (return migrating) after their arrival to the U.S. and provides forward-looking information not contained in a stock-based panel.

As a result of the additional information provided in the complete panel data, we are able to provide several new insights into the return migration and earnings growth of immigrants. First, we show that the attriters (return migrants) experience downward earnings trajectories in the one to two years prior to their departure. Although several researchers have predicted this negative selection into return migration, a lack of data has hindered confirmation of this feature of the immigrant experience. The findings provide evidence that time-variant unobserved characteristics may play an important role in joint determination of immigrant earnings and the decision to return migrate.

We also find that there is significant earnings assimilation with native-born workers along a number of important individual-level characteristics. We find that there is strong convergence within the race and ethnic groups for immigrants. After large differences in initial earnings by educational attainment, there is significant convergence to the earnings of the native-born by 10 years later, although the convergence is not complete. On the other hand, there are large differences in levels and rates of earnings assimilation by country of birth.

Finally, we examine the earnings growth of the immigrant arrival cohort using three different methods. The first estimates provide a rate of earnings growth per year of 5 percent for men and 3.4 percent for women in the repeated cross-section data. The second replicates the analysis using stock-based panel data and finds, as expected, a much lower rate of earnings growth of 2.7 percent for men and 1.2 percent for women. Finally, we use a selection-correction analysis that accounts for the negative selective return migration over time and finds that the estimated coefficients are larger than the stock-based panel data estimates and closer to those from the repeated cross-section. This finding provides strong evidence that there are important time-variant unobserved characteristics that play an important role in determining immigrant earnings and return-migration decisions.

References

Abramitzky, R., L. P. Boustan, and K. Eriksson (2014). A nation of immigrants: Assimilation and economic outcomes in the age of mass migration. *Journal of Political Economy 122*(3), 467–506.

Bhaskar, R., L. Fernandez, and S. R. Porter (2016). Assimilation and coverage of the foreign-born population in administrative records. *Statistical Journal of the IAOS* (Preprint), 1–11.

Bond, B., J. D. Brown, A. Luque, and A. O'Hara (2014). The nature of the bias when studying only linkable person records: Evidence from the American Community Survey. *Center for Administrative Records Research and Applications Working Paper 8*.

Borjas, G. J. (1985). Assimilation, changes in cohort quality, and the earnings of immigrants. *Journal of Labor Economics 3*(4), 463–489.

Borjas, G. J. (1989). Immigrant and emigrant earnings: A longitudinal study. *Economic Inquiry 27*(1), 21–37.

Borjas, G. J. (1995). The economic benefits from immigration. *Journal of Economic Perspectives 9*(2), 3–22.

Borjas, G. J. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *The Quarterly Journal of Economics 118*(4), 1335–1374.

Borjas, G. J. et al. (1987). Self-selection and the earnings of immigrants. *American Economic Review 77*(4), 531–553.

Card, D. (1990). The impact of the Mariel boatlift on the miami labor market. *ILR Review 43*(2), 245–257.

Card, D. (2005). Is the new immigration really so bad? *The Economic Journal 115*(507), F300–F323.

Card, D. (2009). Immigration and inequality. *American Economic Review 99*(2), 1–21.

Chiquiar, D. and G. H. Hanson (2005). International migration, self-selection, and the distribution of wages: Evidence from Mexico and the United States. *Journal of Political Economy 113*(2), 239–281.

Chiswick, B. R. (1978). The effect of Americanization on the earnings of foreign-born men. *Journal of Political Economy 86*(5), 897–921.

Cortes, P. (2008). The effect of low-skilled immigration on US prices: evidence from CPI data. *Journal of Political Economy 116*(3), 381–422.

Duleep, H. O. and D. J. Dowhan (2002). Insights from longitudinal data on the earnings growth of US foreign-born men. *Demography 39*(3), 485–506.

Duleep, H. O. and D. J. Dowhan (2008). Research on immigrant earnings. *Soc. Sec. Bull. 68*, 31.

Duleep, H. O. and M. C. Regets (1997). Measuring immigrant wage growth using matched CPS files. *Demography 34*(2), 239–249.

Dustmann, C., A. Glitz, and T. Vogel (2010). Employment, wages, and the economic cycle: Differences between immigrants and natives. *European Economic Review 54*(1), 1–17.

Dustmann, C. and J.-S. Görlach (2015). Selective out-migration and the estimation of immigrants' earnings profiles. In *Handbook of the Economics of International Migration*, Volume 1, pp. 489–533. Elsevier.

Dustmann, C. and Y. Weiss (2007). Return migration: theory and empirical evidence from the UK. *British Journal of Industrial Relations 45*(2), 236–256.

Foged, M. and G. Peri (2016). Immigrants' effect on native workers: New analysis on longitudinal data. *American Economic Journal: Applied Economics 8*(2), 1–34.

Foster, T. B., M. Ellis, and L. Fiorio (2018). Foreign-born and native-born migration in the us: Evidence from irs administrative and census survey records.

Friedberg, R. M. (2001). The impact of mass migration on the Israeli labor market. *The Quarterly Journal of Economics 116*(4), 1373–1408.

Grieco, E. M., L. J. Larsen, and H. Hogan (2018). How Period Data Influence the Estimates of Recently Arrived Immigrants in the American Community Survey. *International Migration Review 52*(1), 299–313.

Hall, M. and G. Farkas (2008). Does human capital raise earnings for immigrants in the low-skill labor market? *Demography 45*(3), 619–639.

Hu, W.-Y. (2000). Immigrant earnings assimilation: Estimates from longitudinal data. *American Economic Review 90*(2), 368–372.

Hunt, J. (1992). The impact of the 1962 repatriates from Algeria on the French labor market. *ILR Review 45*(3), 556–572.

Jasso, G. and M. R. Rosenzweig (1982). Estimating the emigration rates of legal immigrants using administrative and survey data: The 1971 cohort of immigrants to the United States. *Demography 19*(3), 279–290.

Jasso, G. and M. R. Rosenzweig (1988). How well do u.s. immigrants do? vintage effects, emigration selectivity, and occupational mobility. In T. P. Schultz (Ed.), *Research in Population Economics*, Volume 6. JAI Press.

Kaushal, N. (2011). Earning trajectories of highly educated immigrants: Does place of education matter? *ILR Review 64*(2), 323–340.

Kaushal, N., Y. Lu, N. Denier, J. S.-H. Wang, and S. J. Trejo (2016). Immigrant employment and earnings growth in Canada and the USA: Evidence from longitudinal data. *Journal of Population Economics 29*(4), 1249–1277.

Lehmer, F. and J. Ludsteck (2015). Wage Assimilation of Foreigners: Which Factors Close the Gap? Evidence From Germany. *Review of Income and Wealth 61*(4), 677–701.

Lubotsky, D. (2007). Chutes or ladders? a longitudinal analysis of immigrant earnings. *Journal of Political Economy 115*(5), 820–867.

Lubotsky, D. (2011). The effect of changes in the U.S. wage structure on recent immigrants' earnings. *The Review of Economics and Statistics 93*(1), 59–71.

Ottaviano, G. I. and G. Peri (2012). Rethinking the effect of immigration on wages. *Journal of the European Economic Association 10*(1), 152–197.

Peri, G. (2012). The effect of immigration on productivity: Evidence from U.S. states. *Review of Economics and Statistics 94*(1), 348–358.

Peri, G. and C. Sparber (2009). Task specialization, immigration, and wages. *American Economic Journal: Applied Economics 1*(3), 135–69.

Schwabish, J. A. (2011). Identifying rates of emigration in the United States using administrative earnings records. *International Journal of Population Research 2011.*

Van Hook, J. and W. Zhang (2011). Who stays? Who goes? Selective emigration among the foreign-born. *Population Research and Policy Review 30*(1), 1–24.

Villarreal, A. and C. R. Tamborini (2018). Immigrants' Economic Assimilation: Evidence from Longitudinal Earnings Records. *American Sociological Review*, 0003122418780366.

A Appendix Figures

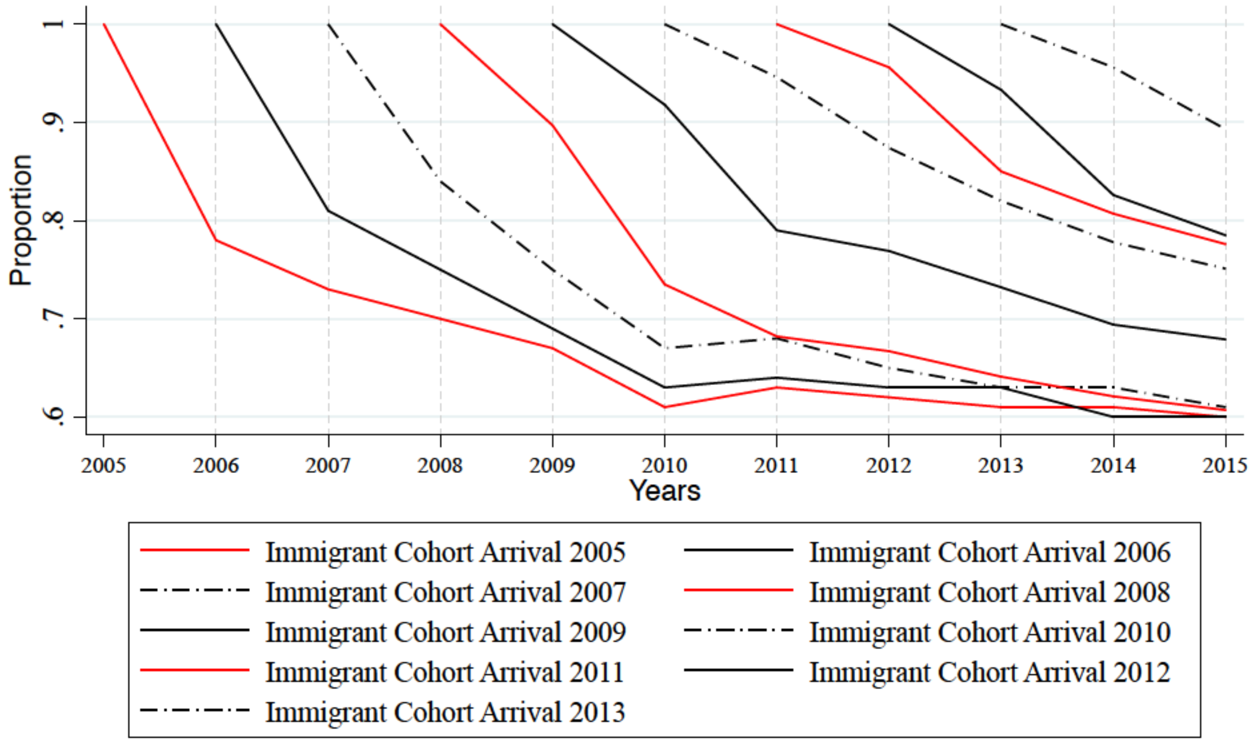
For Online Publication

Figure A1: Previous Research on Immigrant Return Migration and Earnings

Authors	Publication Date	Data Years	Data Sets	Country	Sample Size of Foreign Born
Stock-Based Panel Data					
Duleep and Regets	1997	1988	CPS 1988 Supplement on Immigration for one year (longitudinal).	US	351
Hu	2000	1951-1991	Health and Retirement Survey which is matched to Social Security Earnings for 1951-1991. Again retrospective. Born between 1931-1941	US	<1000
Duleep and Dowhan	2002	1994 backwards	SSA (for all years) and 1994 CPS	US	Not reported
Lubotsky	2007	1951-1997	1951-1997 SSA data; merged to the 1990, 1991 SIPP and 1994 CPS	US	panel; 280,411 for repeated cross section
Hall and Farkas	2008	1996-1999; 2001-2003	4 year SIPP panels (1996-1999, 2001-2003)	US	2,427 for 1996 and 1,813 for 2001
Lubotsky	2011	1980-1997	1990 and 1991 SIPP and 1994 CPS (Cross Sections) merged to SSA	US	1678 in panel
Villarreal and Tamborini	2018	1980-2014	2004 and 2008 SIPP are merged to the SSA data from 1980-2014.	US	1,628
Complete Panel Data					
Jasso and Rosenzweig	1982	1971-1979	resident status matched to the Alien Address report. To identify those that leave by 1979.	US	3,758
Jasso and Rosenzweig	1988	1971-81; 1960-1980	INS New Legal Permanent Residents; 1980 US Census	US	946; 998
Borjas	1989	1972-1978	1972-1978 Survey of Natural and Social Scientists and Engineers.	US	1,166
Dustmann and Weiss	2007	1992-2004	BLFS	UK	10,939
Dustmann, Glitz and Vogel	2010	1982-2001 for Germany; 1981-2005 for UK	BLFS, German IAB 2% sample of soc sec	Germany and UK.	3409 in Germany; 2372 in UK
Schwabish	2011	1978-2003	DER	US	323,896
Van Hook	2011	1996-2009	CPS	US	92,852
Abramitsky	2014	1900, 1910, 1920	Census data	US	20,225

Figure A2: Presence of W-2 or 1099 for 2005-2013 Entry Cohorts for Ages 25-45

Panel A: Men



Panel B: Women

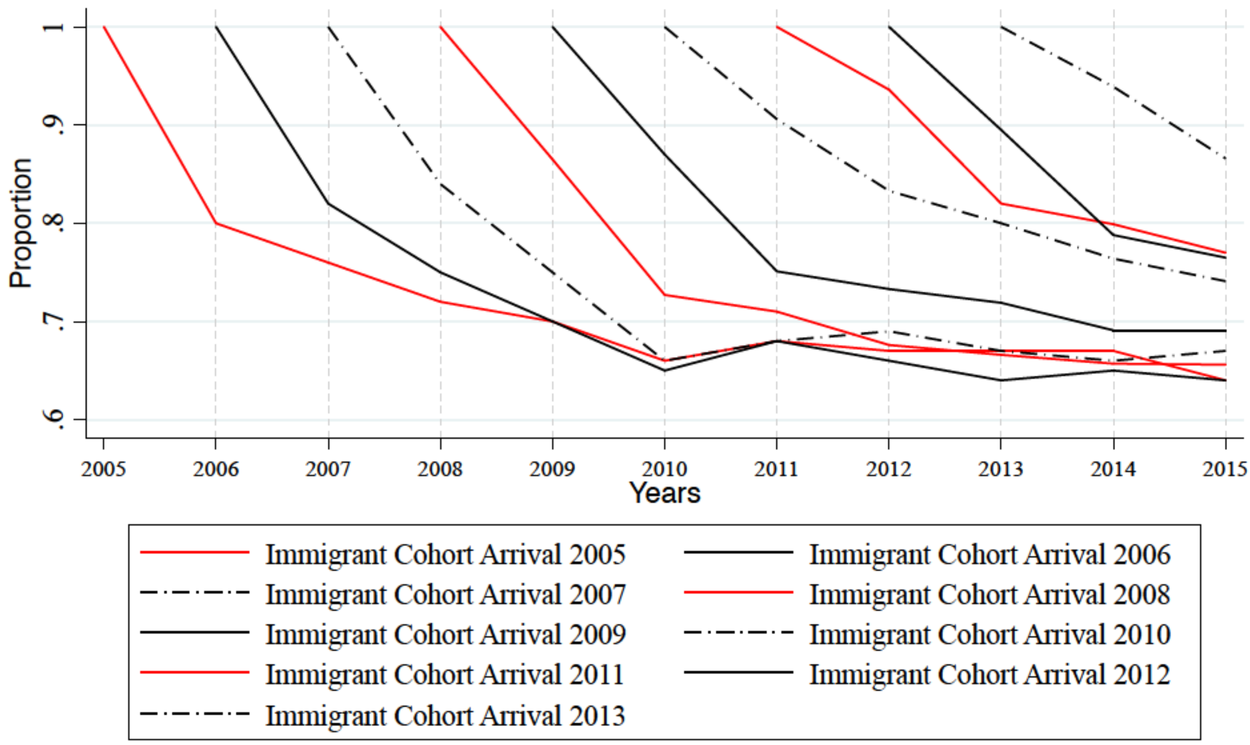
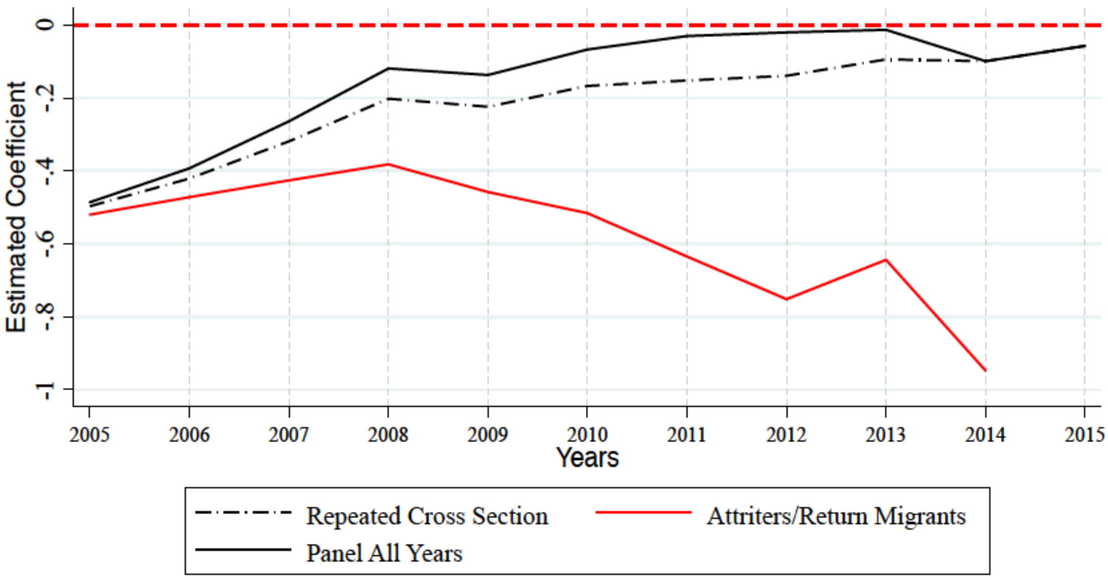
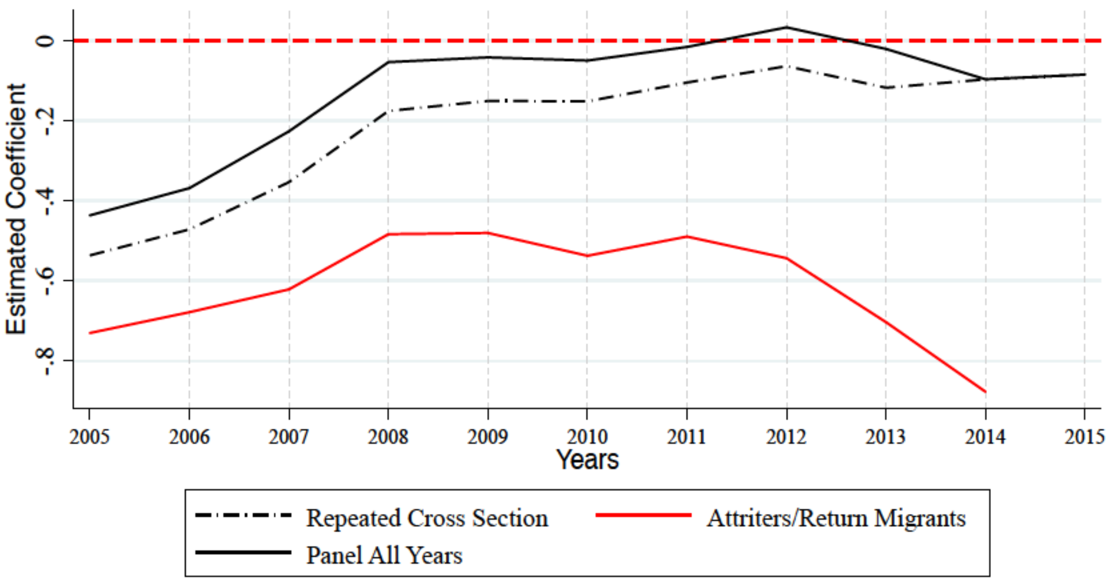


Figure A3: Log Immigrant-Native Annual Earnings for 2005-2007 Arrival Cohorts Ages 25–45 with Educational Controls

Panel A: Men



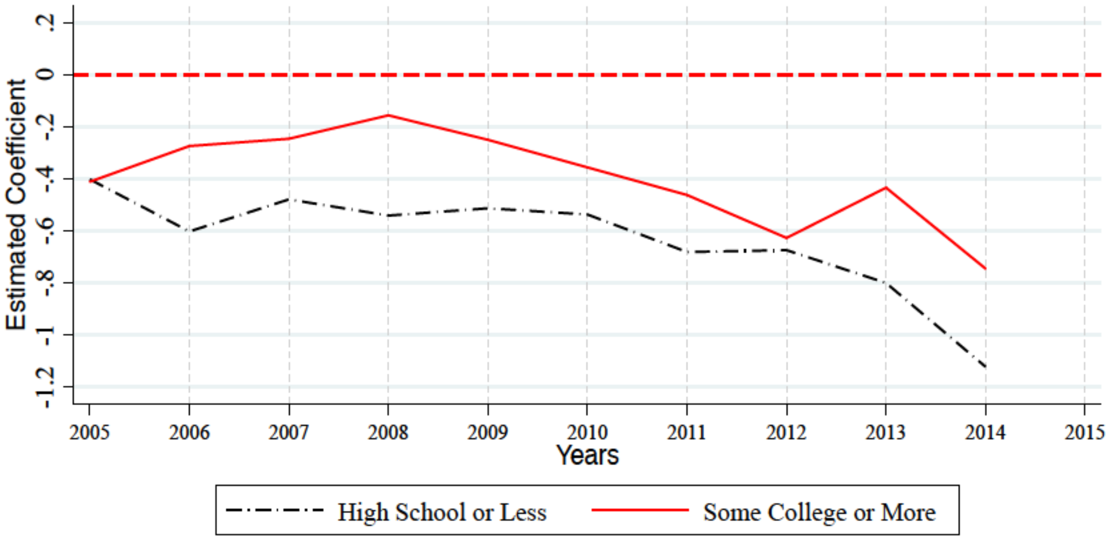
Panel B: Women



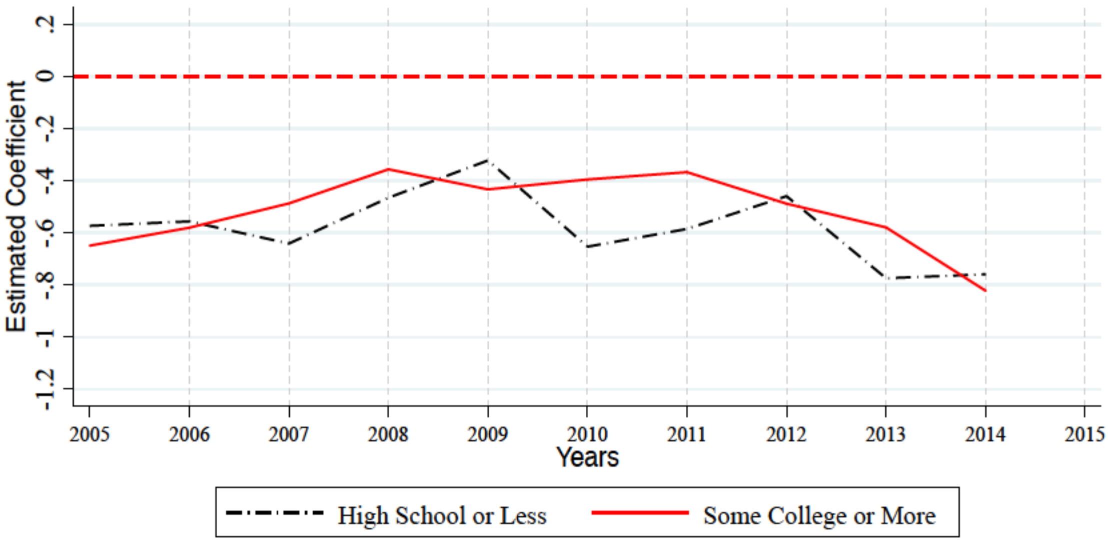
Notes: Each point represents the estimated coefficient on an immigrant indicator variable in a log wage regression conducted separately for each year for the three groups of immigrants relative to the native-born population. We include additional control variables in the regression such as state of residence fixed effects and age fixed effects. Source: ACS 2005–2007 and IRS W-2s or 1099 data (2005–2015).

Figure A4: Log Immigrant-Native Annual Earnings for Attriters Ages 25–45 by Broad Education Categories

Panel A: Men



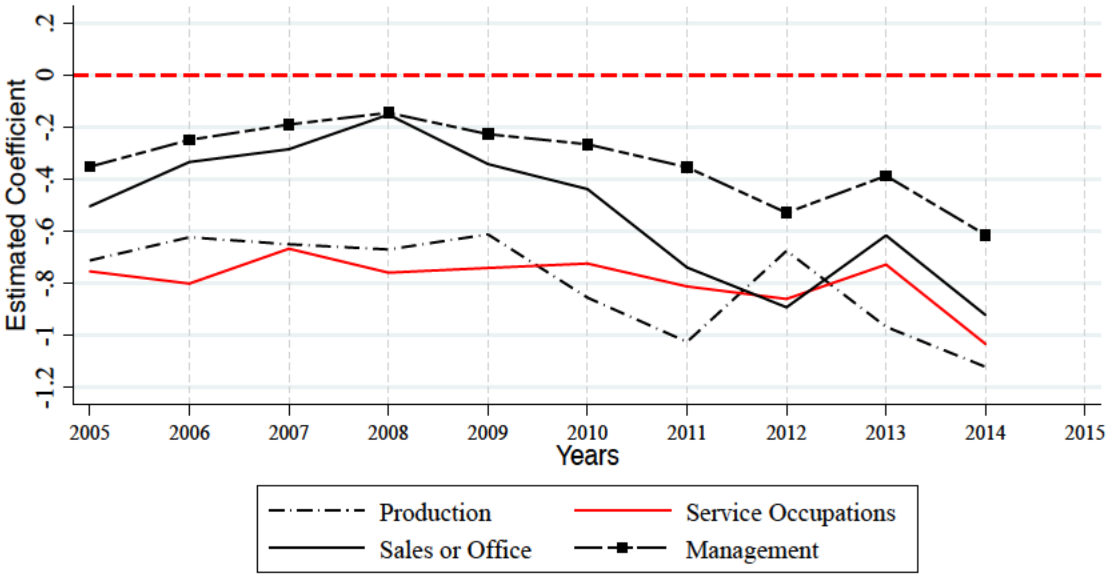
Panel B: Women



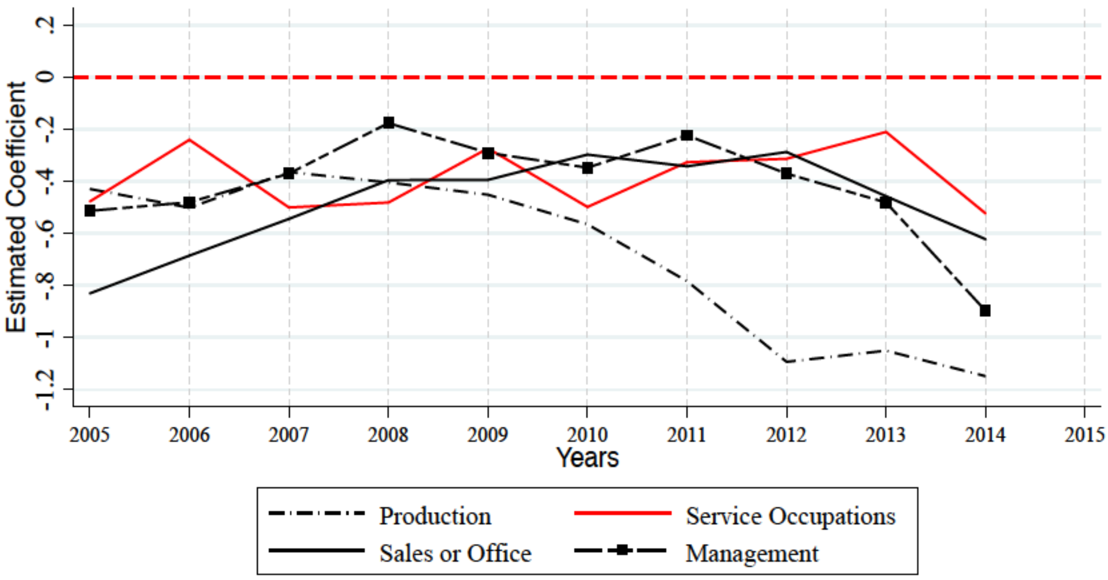
Note: Includes age fixed effects and state of residence fixed effects. Standard errors are clustered at state of residence. Source: ACS 2005–2007 and IRS W-2s or 1099 data (2005–2015).

Figure A5: Log Immigrant-Native Annual Earnings for Attriters Ages 25–45 by Broad Occupation Categories

Panel A: Men



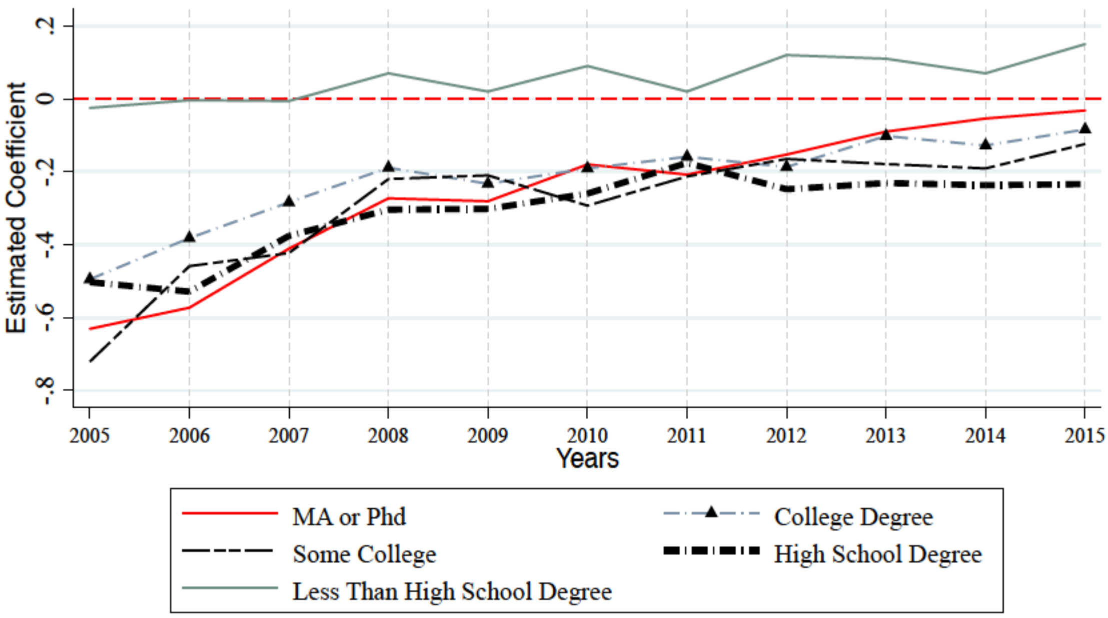
Panel B: Women



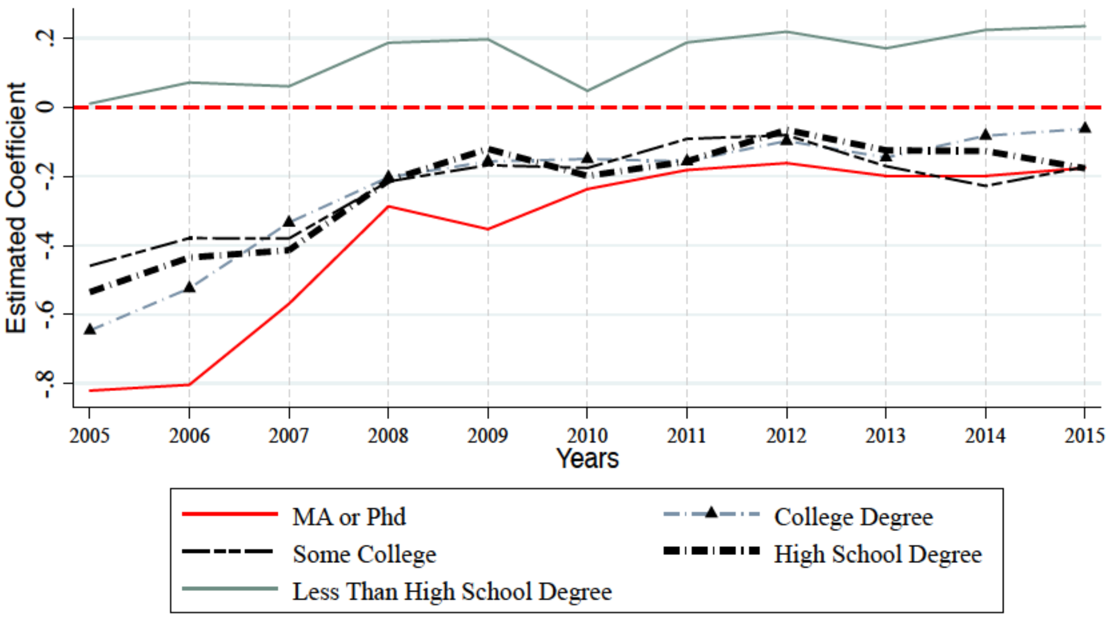
Note: Includes age fixed effects and state of residence fixed effects. Standard errors are clustered at state of residence. Source: ACS 2005–2007 and IRS W-2s or 1099 data (2005–2015).

Figure A6: Log Immigrant-Native Annual Earnings for 2005–2007 Arrival Cohorts Ages 25–45 by Educational Attainment in Cross Section Data

Panel A: Men



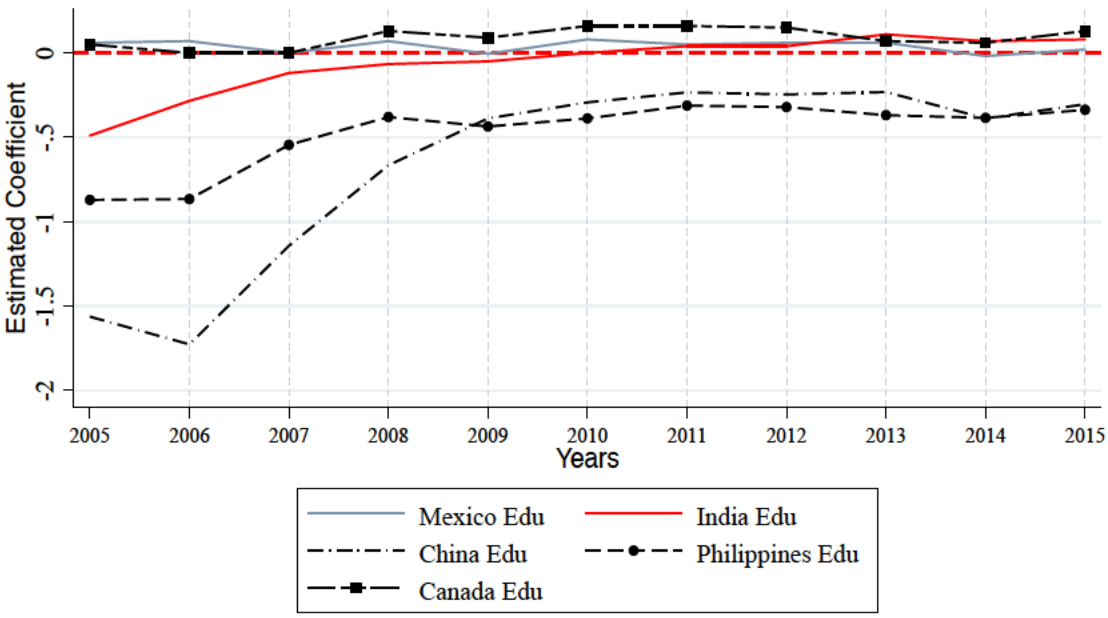
Panel B: Women



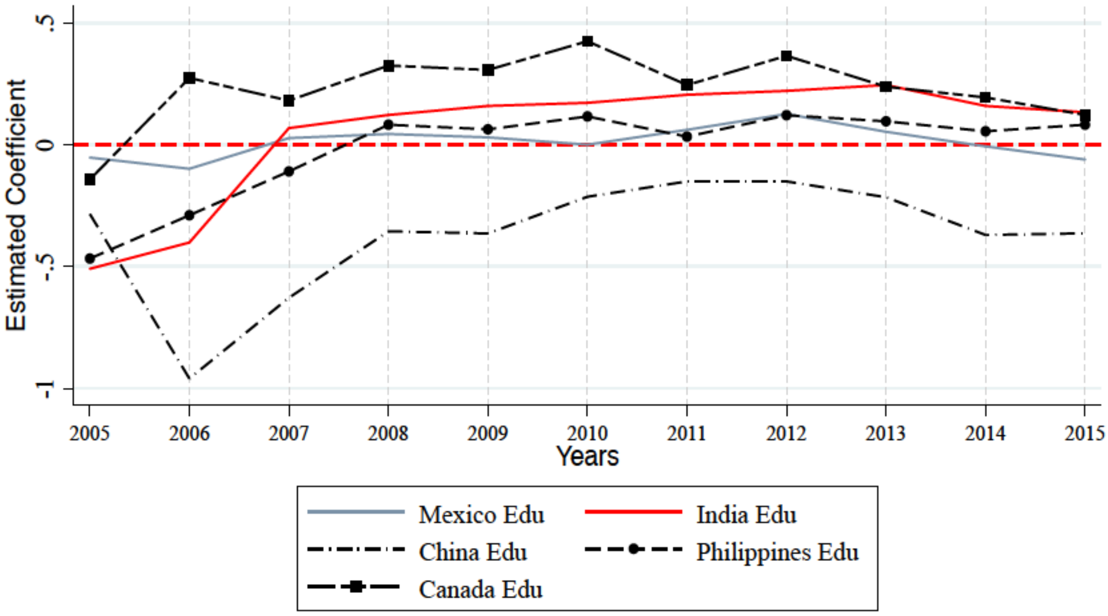
Notes: Each point represents the estimated coefficient on an immigrant indicator variable in a log wage regression conducted separately for each year for the five education groups of immigrants relative to the native-born population. We include additional control variables in the regression such as state of residence fixed effects and age fixed effects. Source: ACS 2005–2007 and IRS W-2s or 1099 data (2005–2015).

Figure A7: Log Immigrant-Native Annual Earnings for 2005–2007 Arrival Cohorts Ages 25–45 by Country of Birth in Panel Data controlling for Educational Attainment

Panel A: Men



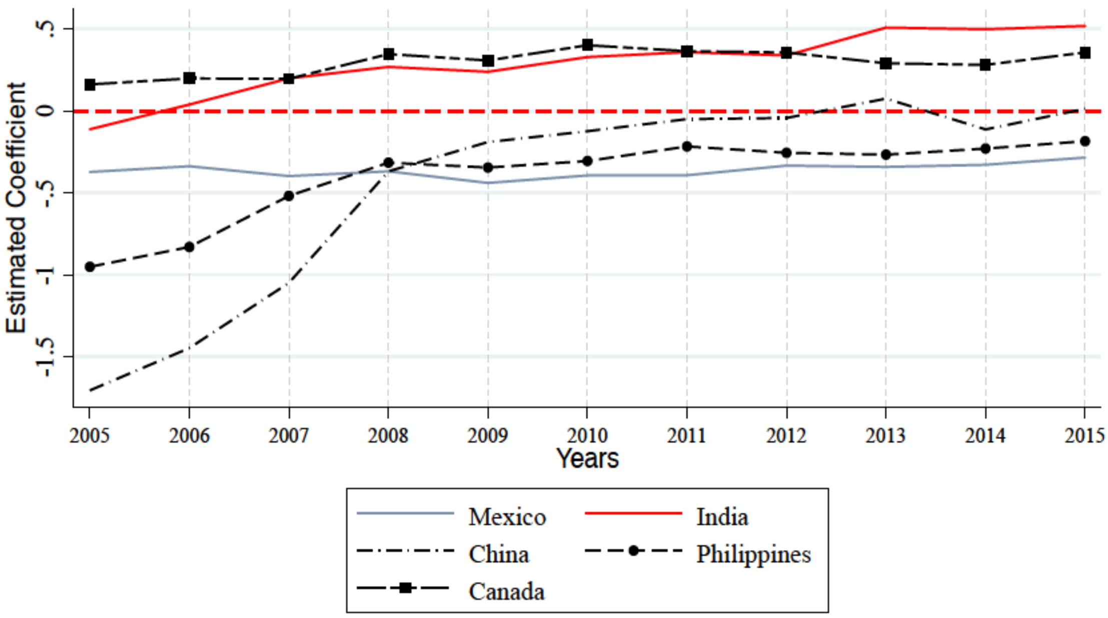
Panel B: Women



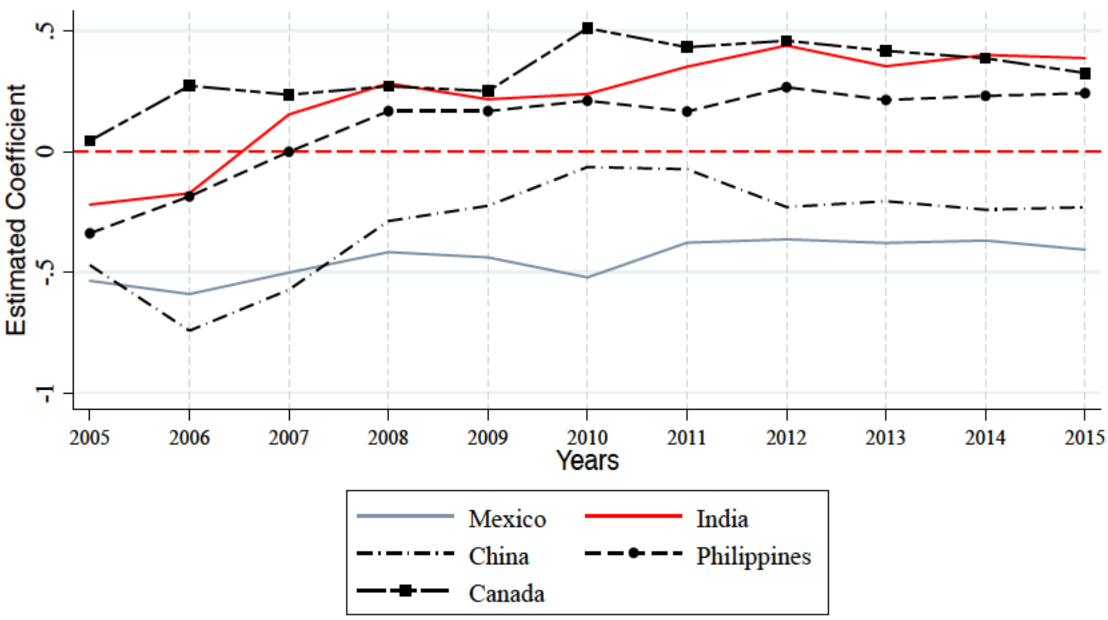
Notes: Each point represents the estimated coefficient on an immigrant indicator variable in a log wage regression conducted separately for each year for the five countries of birth for immigrants relative to the native-born population. We include additional control variables in the regression such as state of residence fixed effects and age fixed effects. Source: ACS 2005–2007 and IRS W-2s or 1099 data (2005–2015).

Figure A8: Log Immigrant-Native Annual Earnings for 2005-2007 Arrival Cohorts Ages 25–45 by Country of Birth in Cross Section Data

Panel A: Men



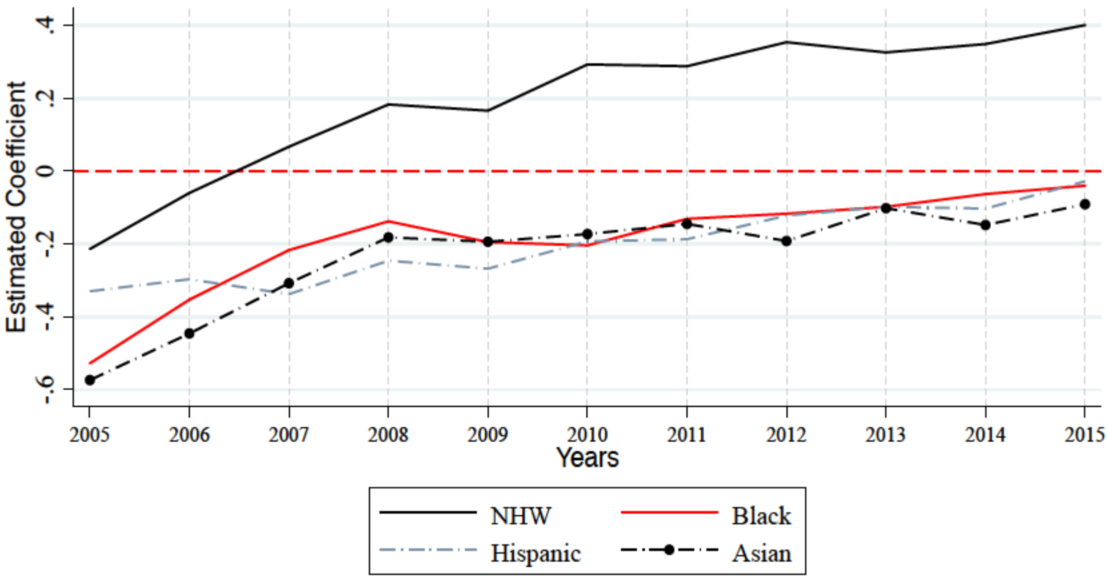
Panel B: Women



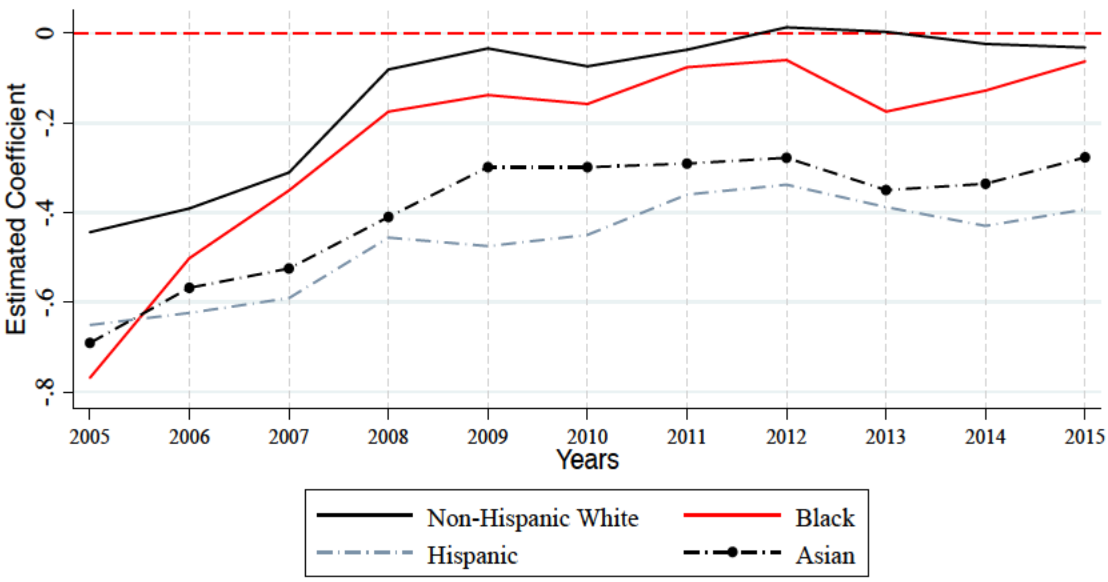
Notes: Each point represents the estimated coefficient on an immigrant indicator variable in a log wage regression conducted separately for each year for the five countries of birth for immigrants relative to the native-born population. We include additional control variables in the regression such as state of residence fixed effects and age fixed effects. Source: ACS 2005–2007 and IRS W-2s or 1099 data (2005–2015).

Figure A9: Log Immigrant-Native Annual Earnings for 2005–2007 Arrival Cohorts Ages 25–45 by Race in Cross Section Data

Panel A: Men



Panel B: Women



Notes: Each point represents the estimated coefficient on an immigrant indicator variable in a log wage regression conducted separately for each year for the four race groups for immigrants relative to the native-born population. We include additional control variables in the regression such as state of residence fixed effects and age fixed effects. Source: ACS 2005–2007 and IRS W-2s or 1099 data (2005–2015).

B Appendix Tables

Table A1: Table of Means for the Matched ACS to No Missing PIK Observations of New Arrivals 2004-2007

Panel A: Men Ages 25-45

	Matched		Non-Matched		
	Mean	Standard Deviation	Mean	Standard Deviation	T-Stat
Total Income	41,000	56,000	17,500	23,000	31.66
Wages or Salary Income	40,000	53,000	16,500	21,000	33.65
Self-Employment Income	1,000	15,000	1,000	8,700	0.00
Age	34	6	32	6	14.85
Male	1.00	0.00	1.00	0.00	0.00
Married	0.68	0.47	0.56	0.50	14.22
Less than High School Education	0.12	0.33	0.45	0.50	-43.48
High School Graduate	0.15	0.36	0.26	0.44	-15.75
Some Post High School Education	0.12	0.33	0.08	0.27	7.76
College Degree	0.30	0.46	0.10	0.30	29.16
MA or PhD	0.30	0.46	0.07	0.25	35.98
Mexico	0.12	0.33	0.58	0.49	-60.81
India	0.15	0.36	0.03	0.17	24.64
Philippines	0.06	0.23	0.01	0.10	15.11
China	0.03	0.18	0.02	0.13	5.80
Canada	0.04	0.18	0.01	0.09	10.66
Other	0.61	0.49	0.36	0.48	29.23
Manufacturing, Transport, Production	0.11	0.31	0.15	0.36	-7.64
Natural Resources, Mining	0.11	0.32	0.40	0.49	-38.42
Office Occupations	0.11	0.31	0.05	0.22	12.39
Service Occupations	0.11	0.31	0.19	0.39	-13.18
Management	0.50	0.50	0.11	0.32	52.41

Note: There are 6,800 observations for the matched and 5,900 for the non-matched samples. Source: ACS 2005–2007 and IRS W-2s or 1099 data (2005–2015). Numbers have been rounded to comply with the Census Bureau's disclosure-avoidance guidelines.

Panel B: Women Ages 25-45

	Matched		Non-Matched		
	Mean	Standard Deviation	Mean	Standard Deviation	T-Stat
Total Income	15000	26500	7000	18100	19.67
Wages or Salary Income	14000	26000	6000	17300	20.23
Self-Employment Income	400	3900	400	3500	0.00
Age	33	6	33	6	6.66
Male	0.00	0.00	0.00	0.00	0.00
Married	0.76	0.43	0.71	0.44	6.30
Less than High School Education	0.12	0.32	0.25	0.43	-18.70
High School Graduate	0.17	0.38	0.19	0.39	-2.56
Some Post High School Education	0.17	0.37	0.13	0.33	5.49
College Degree	0.32	0.47	0.25	0.43	8.57
MA or PhD	0.20	0.40	0.14	0.34	9.47
Mexico	0.09	0.29	0.31	0.46	-30.33
India	0.09	0.29	0.10	0.30	-1.67
Philippines	0.11	0.31	0.02	0.15	20.76
China	0.05	0.21	0.04	0.20	1.61
Canada	0.03	0.17	0.01	0.12	7.49
Other	0.63	0.48	0.49	0.50	15.21
Manufacturing, Transport, Production	0.06	0.24	0.06	0.24	0.23
Natural Resources, Mining	0.01	0.09	0.02	0.15	-5.18
Office Occupations	0.17	0.38	0.10	0.30	11.84
Service Occupations	0.15	0.36	0.18	0.38	-4.43
Management	0.35	0.48	0.16	0.36	25.22

Note: There are 6,700 observations for the matched and 5,400 for the non-matched samples. Source: ACS 2005–2007 and IRS W-2s or 1099 data (2005–2015). Numbers have been rounded to comply with the Census Bureau's disclosure-avoidance guidelines.

Table A2: Table of Means for the Matched to W-2s of New Arrivals 2004-2007

Panel A: Men Ages 25-45

	Matched		Non-Matched		
	Mean	Standard Deviation	Mean	Standard Deviation	T-Stat
Total Income	42,500	56,000	33,000	54,500	5.12
Wages or Salary Income	41,000	54,500	28,500	43,000	8.26
Self-Employment Income	600	8,700	3,200	33,000	-2.48
Age	34	6	34	6	-4.12
Male	1.00	0.00	1.00	0.00	0.00
Married	0.68	0.47	0.70	0.46	-1.67
Less than High School Education	0.11	0.32	0.18	0.38	-4.98
High School Graduate	0.15	0.35	0.16	0.37	-1.29
Some Post High School Education	0.12	0.32	0.13	0.34	-1.40
College Degree	0.31	0.46	0.19	0.39	9.36
MA or PhD	0.29	0.46	0.32	0.47	-1.59
Mexico	0.11	0.32	0.16	0.36	-3.55
India	0.16	0.37	0.06	0.23	12.22
Philippines	0.06	0.23	0.04	0.20	2.15
China	0.03	0.18	0.05	0.21	-2.03
Canada	0.04	0.19	0.03	0.17	0.84
Other	0.60	0.49	0.67	0.47	-4.36
Manufacturing, Transport, Production	0.11	0.32	0.08	0.28	2.92
Natural Resources, Mining	0.11	0.31	0.15	0.36	-3.60
Office Occupations	0.11	0.32	0.10	0.30	1.48
Service Occupations	0.11	0.31	0.10	0.30	0.99
Management	0.51	0.50	0.41	0.49	6.29

Note: There are 5,800 observations for the matched and 1,000 for the non-matched samples. Source: ACS 2005–2007 and IRS W-2s or 1099 data (2005–2015). Numbers have been rounded to comply with the Census Bureau's disclosure-avoidance guidelines.

Panel B: Women Ages 25-45

	Matched		Non-Matched		
	Mean	Standard Deviation	Mean	Standard Deviation	T-Stat
Total Income	19,000	29,000	7,700	19,000	19.07
Wages or Salary Income	18,000	28,500	6,200	17,500	20.87
Self-Employment Income	300	3,000	600	5,200	-2.51
Age	33	6	34	6	-3.35
Male	0.00	0.00	0.00	0.00	0.00
Married	0.71	0.45	0.86	0.35	-14.12
Less than High School Education	0.11	0.31	0.14	0.35	-3.56
High School Graduate	0.17	0.38	0.17	0.38	0.31
Some Post High School Education	0.16	0.37	0.17	0.37	-0.41
College Degree	0.33	0.47	0.30	0.46	2.41
MA or PhD	0.21	0.41	0.20	0.40	1.25
Mexico	0.08	0.27	0.11	0.32	-4.05
India	0.10	0.29	0.08	0.28	1.91
Philippines	0.13	0.34	0.07	0.25	8.74
China	0.04	0.20	0.05	0.23	-1.94
Canada	0.04	0.19	0.02	0.14	3.65
Other	0.61	0.49	0.66	0.47	-4.02
Manufacturing, Transport, Production	0.08	0.27	0.03	0.17	9.05
Natural Resources, Mining	0.01	0.09	0.01	0.09	0.44
Office Occupations	0.20	0.40	0.13	0.33	7.24
Service Occupations	0.17	0.38	0.11	0.31	7.28
Management	0.39	0.49	0.28	0.45	8.65

Note: There are 4,500 observations for the matched and 2,200 for the non-matched samples. Source: ACS 2005–2007 and IRS W-2s or 1099 data (2005–2015). Numbers have been rounded to comply with the Census Bureau's disclosure-avoidance guidelines.

Table A3: Percent Found in 2010 U.S. Census for Return Migrants

Last Year of W2 or 1099 in Data	Years in Data	Men	Women
2005	1 Year	4	4
2006	2 Years	10	11
2007	3 Years	18	25
2008	4 Years	28	41
2009	5 Years	43	60
2010	6 Years	67	81

Note: This table identifies the percent of individuals who have a final administrative record (W-2 or 1099) reported in the years 2005-2010 that can be found in the 2010 U.S. Census. Source: ACS 2005–2007, 2010 U.S. Census and IRS W-2s or 1099 data (2005–2015). Numbers have been rounded to comply with the Census Bureau's disclosure-avoidance guidelines.

Table A4: Missing by Country of Birth

Panel A: Men

VARIABLES	(1) Missing in 2006	(2) Missing in 2007	(3) Missing in 2008	(4) Missing in 2009	(5) Missing in 2010	(6) Missing in 2011	(7) Missing in 2012	(8) Missing in 2013	(9) Missing in 2014	(10) Missing in 2015
Mexico	-0.017	0.030	0.021	0.020	-0.040*	-0.148***	-0.118***	-0.115***	-0.109***	-0.108***
	(0.025)	(0.027)	(0.031)	(0.025)	(0.023)	(0.024)	(0.019)	(0.023)	(0.020)	(0.021)
India	-0.023	-0.032**	-0.025*	-0.057***	-0.065***	-0.082***	-0.059**	-0.025	-0.019	-0.051
	(0.015)	(0.012)	(0.012)	(0.019)	(0.022)	(0.021)	(0.023)	(0.024)	(0.029)	(0.031)
Philippines	-0.024	-0.067***	-0.094***	-0.106***	-0.149***	-0.142***	-0.151***	-0.152***	-0.129***	-0.172***
	(0.030)	(0.017)	(0.023)	(0.025)	(0.025)	(0.022)	(0.027)	(0.032)	(0.019)	(0.022)
China	0.043	0.033	0.028	0.041	-0.001	-0.041	-0.050	-0.040	-0.095*	-0.078**
	(0.044)	(0.027)	(0.032)	(0.029)	(0.032)	(0.034)	(0.053)	(0.040)	(0.050)	(0.036)
Canada	0.051	0.044	0.052	0.006	-0.013	-0.008	0.017	-0.007	0.003	0.034
	(0.033)	(0.048)	(0.033)	(0.038)	(0.038)	(0.048)	(0.047)	(0.054)	(0.047)	(0.052)
Constant	0.350***	0.408***	0.455***	0.387***	0.552***	0.348***	0.427***	0.354***	0.386***	0.500***
	(0.047)	(0.021)	(0.019)	(0.026)	(0.027)	(0.025)	(0.034)	(0.024)	(0.031)	(0.025)
Observations	5,400	5,400	5,400	5,400	5,400	5,400	5,400	5,400	5,400	5,400
R-squared	0.123	0.027	0.026	0.022	0.024	0.031	0.031	0.033	0.025	0.028

Note: Includes age fixed effects, year of entry, and state fixed effects. Standard errors are clustered at the state of residence. Omitted country of birth category is all other countries of birth. Source: ACS 2005–2007 and IRS 2005–2015 W-2 and 1099 data. Numbers have been rounded to comply with the Census Bureau's disclosure-avoidance guidelines.

Panel B: Women

VARIABLES	(1) Missing in 2006	(2) Missing in 2007	(3) Missing in 2008	(4) Missing in 2009	(5) Missing in 2010	(6) Missing in 2011	(7) Missing in 2012	(8) Missing in 2013	(9) Missing in 2014	(10) Missing in 2015
Mexico	-0.029	-0.013	-0.034	-0.018	-0.107***	-0.103***	-0.102***	-0.062**	-0.066**	-0.071***
	(0.033)	(0.027)	(0.037)	(0.037)	(0.026)	(0.020)	(0.023)	(0.026)	(0.027)	(0.022)
India	0.004	-0.002	-0.041**	-0.071***	-0.066***	-0.023	-0.028	-0.029	-0.020	-0.013
	(0.033)	(0.025)	(0.019)	(0.021)	(0.022)	(0.023)	(0.026)	(0.025)	(0.026)	(0.025)
Philippines	-0.056**	-0.063***	-0.115***	-0.154***	-0.203***	-0.199***	-0.174***	-0.187***	-0.208***	-0.193***
	(0.025)	(0.016)	(0.027)	(0.024)	(0.018)	(0.016)	(0.024)	(0.026)	(0.025)	(0.021)
China	0.064*	0.065*	-0.001	0.015	0.001	0.016	-0.053	-0.050	-0.066	-0.086**
	(0.037)	(0.038)	(0.034)	(0.023)	(0.035)	(0.033)	(0.041)	(0.041)	(0.048)	(0.037)
Canada	-0.023	-0.011	-0.025	0.031	0.054	0.053	0.029	0.083*	0.071	0.088*
	(0.045)	(0.040)	(0.041)	(0.040)	(0.043)	(0.047)	(0.047)	(0.049)	(0.051)	(0.046)
Constant	0.269***	0.413***	0.461***	0.467***	0.459***	0.389***	0.477***	0.407***	0.317***	0.346***
	(0.033)	(0.022)	(0.028)	(0.024)	(0.029)	(0.029)	(0.031)	(0.031)	(0.030)	(0.031)
Observations	4,200	4,200	4,200	4,200	4,200	4,200	4,200	4,200	4,200	4,200
R-squared	0.098	0.038	0.037	0.037	0.043	0.049	0.050	0.057	0.054	0.046

Note: Includes age fixed effects, year of entry, and state fixed effects. Standard errors are clustered at the state of residence. Omitted country of birth category is all other countries of birth. Source: ACS 2005–2007 and IRS 2005–2015 W-2 and 1099 data. Numbers have been rounded to comply with the Census Bureau's disclosure-avoidance guidelines.

Table A5: Missing by Combination of Country and Education

Panel A: Men	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10
VARIABLES	Missing in 2006	Missing in 2007	Missing in 2008	Missing in 2009	Missing in 2010	Missing in 2011	Missing in 2012	Missing in 2013	Missing in 2014	Missin 201
Mexico	-0.051	-0.012	-0.01	0.06	-0.008	-0.037	-0.052	-0.007	0.001	-0.01
	-0.043	-0.039	-0.042	-0.038	-0.036	-0.039	-0.032	-0.029	-0.035	-0.03
India	-0.037	0.103	0.016	0.09	0.101	-0.074	0.056	0.15	0.206	0.11
	-0.097	-0.108	-0.117	-0.164	-0.179	-0.126	-0.15	-0.129	-0.14	-0.15
Philippines	-0.134	-0.174***	-0.001	-0.118	0.137**	-0.061	-0.148*	-0.132	-0.004	-0.19
	-0.136	-0.034	-0.11	-0.097	-0.065	-0.073	-0.074	-0.083	-0.154	-0.10
China	0.403***	0.346***	0.052	0.366***	0.225**	0.222*	0.07	0.330***	-0.1	-0.129
	-0.095	-0.074	-0.133	-0.09	-0.101	-0.113	-0.124	-0.081	-0.073	-0.0
Canada	0.493*	0.502*	0.232	0.246	0.474*	0.052	0.012	0.063	0.013	0.00
	-0.255	-0.253	-0.279	-0.26	-0.243	-0.224	-0.23	-0.231	-0.234	-0.24
High School Degree	-0.001	-0.018	-0.047	0.008	0.032	0.088**	0.028	0.089***	0.080**	0.061
	-0.032	-0.026	-0.033	-0.029	-0.03	-0.036	-0.034	-0.032	-0.036	-0.03
Some College	0.01	-0.046**	-0.040*	0.031	-0.003	0.062**	-0.002	0.049*	0.037	0.03
	-0.026	-0.022	-0.023	-0.025	-0.03	-0.029	-0.031	-0.029	-0.028	-0.0
College Degree	0.022	-0.038	-0.050*	0.034	0.062**	0.122***	0.093**	0.157***	0.124***	0.106
	-0.032	-0.033	-0.028	-0.028	-0.027	-0.028	-0.037	-0.033	-0.031	-0.03
MA or Phd Degree	-0.038	-0.033	0.004	0.073**	0.092***	0.196***	0.144***	0.194***	0.181***	0.166
	-0.032	-0.031	-0.03	-0.036	-0.028	-0.027	-0.031	-0.019	-0.026	-0.03
Mexico x HS Degree	0.011	0.011	0.005	-0.015	0.072	-0.079	-0.012	-0.04	-0.05	-0.01
	-0.062	-0.055	-0.055	-0.065	-0.059	-0.061	-0.055	-0.057	-0.085	-0.08
Mexico x Some College	0.035	0.073	0.106	-0.042	0.078	0.024	0.095	-0.018	-0.063	-0.06
	-0.073	-0.072	-0.088	-0.084	-0.064	-0.098	-0.081	-0.089	-0.066	-0.07
Mexico x College	0.098	0.110**	0.080*	-0.021	-0.084	-0.072	-0.059	-0.071	-0.123**	-0.08
	-0.065	-0.046	-0.046	-0.057	-0.077	-0.057	-0.076	-0.079	-0.06	-0.07

Mexico x MA Phd	0.132	0.043	0.031	-0.09	-0.148*	-0.263***	-0.194***	-0.243***	-0.212***	-0.229
	-0.079	-0.077	-0.083	-0.084	-0.078	-0.048	-0.059	-0.057	-0.072	-0.06
India x HS Degree	0.094	-0.167	0.121	-0.081	-0.235	0.007	-0.195	-0.22	-0.285	-0.16
	-0.167	-0.151	-0.151	-0.201	-0.198	-0.138	-0.181	-0.163	-0.172	-0.18
India x Some College	-0.124	-0.147	-0.085	-0.131	-0.133	0.102	-0.014	-0.061	-0.145	-0.03
	-0.123	-0.116	-0.154	-0.161	-0.213	-0.187	-0.187	-0.169	-0.173	-0.18
India x College	-0.015	-0.121	-0.015	-0.14	-0.158	0.002	-0.106	-0.184	-0.199	-0.14
	-0.101	-0.108	-0.116	-0.172	-0.179	-0.126	-0.146	-0.137	-0.133	-0.14
India x MA Phd	0.052	-0.148	-0.086	-0.196	-0.225	-0.092	-0.203	-0.263*	-0.334**	-0.27
	-0.099	-0.11	-0.118	-0.164	-0.169	-0.132	-0.16	-0.135	-0.144	-0.15
China x HS Degree	-0.319	-0.253*	0.165	-0.26	-0.219	-0.335**	0.008	-0.519***	-0.156	0.20
	-0.202	-0.144	-0.254	-0.198	-0.203	-0.142	-0.127	-0.116	-0.111	-0.14
China x Some College	-0.323	-0.234	-0.117	-0.488**	-0.219	-0.216	-0.382***	-0.499***	-0.076	0.24
	-0.282	-0.196	-0.207	-0.194	-0.258	-0.251	-0.131	-0.181	-0.166	-0.21
China x College	-0.303**	-0.295**	0.211	-0.215*	-0.132	-0.079	-0.006	-0.227	0.153	0.11
	-0.138	-0.117	-0.182	-0.118	-0.134	-0.154	-0.099	-0.149	-0.099	-0.12
China x MA Phd	-0.438***	-0.397***	-0.152	-0.447***	-0.329***	-0.412***	-0.229*	-0.502***	-0.058	-0.03
	-0.11	-0.08	-0.132	-0.106	-0.115	-0.128	-0.128	-0.088	-0.083	-0.08
Philippines x HS Degree	0.043	0.008	-0.105	0.004	-0.332***	-0.133	0.007	-0.004	-0.166	-0.0
	-0.123	-0.056	-0.167	-0.131	-0.115	-0.102	-0.093	-0.106	-0.159	-0.08
Philippines x Some College	-0.009	0.129**	0.025	0.089	-0.222***	0.005	0.121	0.073	-0.059	0.09
	-0.148	-0.055	-0.107	-0.106	-0.081	-0.063	-0.093	-0.147	-0.207	-0.15
Philippines x College	0.124	0.131***	-0.093	0.018	-0.306***	-0.078	-0.017	-0.055	-0.132	0.03
	-0.12	-0.046	-0.13	-0.121	-0.081	-0.063	-0.077	-0.085	-0.153	-0.09
Philippines x MA Phd	0.198*	0.120**	-0.18	-0.104	-0.289**	-0.127	-0.121	-0.004	-0.151	-0.00
	-0.111	-0.053	-0.134	-0.146	-0.12	-0.111	-0.099	-0.13	-0.196	-0.16
Canada x HS Degree	-0.398	-0.296	-0.092	-0.26	-0.454*	-0.06	-0.018	-0.101	-0.086	0.03
	-0.309	-0.224	-0.293	-0.272	-0.245	-0.24	-0.247	-0.257	-0.274	-0.27
Canada x Some College	-0.438*	-0.496**	-0.137	-0.217	-0.442*	-0.075	-0.017	-0.049	0.009	0.05

	-0.261	-0.223	-0.309	-0.272	-0.246	-0.234	-0.238	-0.246	-0.238	-0.2
Canada x College	-0.493*	-0.485**	-0.246	-0.31	-0.570**	-0.078	-0.037	-0.135	-0.03	-0.0
	-0.262	-0.239	-0.269	-0.257	-0.241	-0.224	-0.227	-0.243	-0.252	-0.26
Canada x MA Phd	-0.434	-0.461*	-0.141	-0.18	-0.457*	-0.045	0.075	-0.017	0.013	0.07
	-0.281	-0.236	-0.288	-0.257	-0.258	-0.239	-0.244	-0.25	-0.253	-0.25
Constant	0.360***	0.437***	0.472***	0.345***	0.504***	0.234***	0.351***	0.233***	0.284***	0.413*
	-0.045	-0.027	-0.029	-0.033	-0.035	-0.028	-0.038	-0.027	-0.037	-0.03
Observations	5,400	5,400	5,400	5,400	5,400	5,400	5,400	5,400	5,400	5,40
R-squared	0.131	0.033	0.033	0.028	0.032	0.044	0.043	0.048	0.037	0.03

Note: Includes age fixed effects, year of entry, and state fixed effects. Standard errors are clustered at the state of residence. Source: ACS 2005–2007 and IRS 2005–2015 W-2 and 1099 data. Numbers have been rounded to comply with the Census Bureau's disclosure-avoidance guidelines.

Panel B: Women

VARIABLES	(1) Missing in 2006	(2) Missing in 2007	(3) Missing in 2008	(4) Missing in 2009	(5) Missing in 2010	(6) Missing in 2011	(7) Missing in 2012	(8) Missing in 2013	(9) Missing in 2014	(10 Missin 201
Mexico	-0.060	-0.026	-0.047	0.053	-0.012	-0.019	-0.060	-0.004	0.007	0.03
	(0.087)	(0.034)	(0.055)	(0.085)	(0.073)	(0.051)	(0.044)	(0.061)	(0.055)	(0.04
India	0.143	-0.018	-0.183***	-0.140**	-0.099	0.044	0.005	-0.123**	-0.114*	0.00
	(0.132)	(0.059)	(0.062)	(0.059)	(0.083)	(0.072)	(0.076)	(0.053)	(0.061)	(0.08
Philippines	0.017	-0.146**	0.127	-0.056	-0.183	-0.149	-0.049	-0.205**	-0.102	-0.16
	(0.089)	(0.069)	(0.114)	(0.144)	(0.143)	(0.147)	(0.137)	(0.088)	(0.121)	(0.08
China	0.214***	0.271***	0.193*	0.148*	0.098	0.043	-0.070	-0.079	-0.023	-0.01
	(0.073)	(0.087)	(0.109)	(0.083)	(0.091)	(0.097)	(0.081)	(0.107)	(0.139)	(0.10
Canada	0.126	-0.041	-0.073	-0.280***	0.168	0.196	0.162	0.161	0.129	0.19
	(0.322)	(0.146)	(0.228)	(0.056)	(0.292)	(0.327)	(0.298)	(0.277)	(0.268)	(0.27
High School Degree	-0.033	0.011	-0.023	0.038	0.070*	0.055	0.014	0.024	0.032	0.05
	(0.047)	(0.026)	(0.039)	(0.032)	(0.039)	(0.038)	(0.036)	(0.033)	(0.048)	(0.03
Some College	0.016	0.070**	0.033	0.036	0.026	0.038	-0.006	0.027	0.014	0.065
	(0.037)	(0.034)	(0.038)	(0.029)	(0.037)	(0.034)	(0.036)	(0.027)	(0.042)	(0.03
College Degree	-0.019	0.042	-0.015	0.019	0.049	0.061*	0.041	0.048**	0.072*	0.104
	(0.034)	(0.026)	(0.042)	(0.037)	(0.038)	(0.033)	(0.029)	(0.021)	(0.036)	(0.03
MA or Phd Degree	-0.037	0.024	0.044	0.101***	0.132***	0.156***	0.111***	0.108***	0.142***	0.182
	(0.044)	(0.032)	(0.036)	(0.026)	(0.034)	(0.034)	(0.036)	(0.031)	(0.046)	(0.04
Mexico x HS Degree	0.035	0.014	0.031	-0.138	-0.165**	-0.063	-0.080	-0.014	-0.075	-0.146
	(0.155)	(0.070)	(0.067)	(0.096)	(0.068)	(0.068)	(0.053)	(0.077)	(0.074)	(0.06
Mexico x Some College	0.045	0.042	0.038	-0.024	-0.076	-0.024	0.084	-0.092	-0.049	-0.131
	(0.104)	(0.065)	(0.060)	(0.094)	(0.135)	(0.153)	(0.098)	(0.080)	(0.093)	(0.05
Mexico x College	-0.007	0.061	0.007	-0.143	-0.121	-0.183***	-0.091	-0.109	-0.139*	-0.12
	(0.084)	(0.062)	(0.093)	(0.100)	(0.091)	(0.065)	(0.063)	(0.093)	(0.075)	(0.06

Mexico x MA Phd	0.178	0.098	0.085	0.056	-0.083	-0.062	-0.003	-0.031	0.021	0.03
	(0.117)	(0.120)	(0.076)	(0.091)	(0.100)	(0.109)	(0.074)	(0.100)	(0.100)	(0.10
India x HS Degree	-0.148	0.033	0.096	0.039	-0.015	-0.099	-0.062	0.123	-0.081	-0.13
	(0.143)	(0.116)	(0.064)	(0.089)	(0.122)	(0.138)	(0.110)	(0.126)	(0.115)	(0.12
India x Some College	-0.284***	0.037	0.125	0.076	0.001	-0.141	-0.099	0.033	0.029	-0.06
	(0.099)	(0.113)	(0.110)	(0.132)	(0.131)	(0.133)	(0.136)	(0.131)	(0.137)	(0.10
India x College	-0.099	0.021	0.176**	0.118*	0.012	-0.069	-0.030	0.095	0.098	-0.00
	(0.143)	(0.073)	(0.078)	(0.068)	(0.110)	(0.074)	(0.087)	(0.062)	(0.071)	(0.08
India x MA Phd	-0.160	0.006	0.116*	0.000	0.028	-0.118	-0.085	0.066	0.061	-0.08
	(0.126)	(0.071)	(0.068)	(0.064)	(0.085)	(0.074)	(0.083)	(0.063)	(0.056)	(0.09
China x HS Degree	-0.036	-0.087	0.071	-0.130	-0.079	0.032	0.031	0.098	-0.089	-0.09
	(0.144)	(0.133)	(0.137)	(0.100)	(0.173)	(0.178)	(0.141)	(0.171)	(0.149)	(0.17
China x Some College	-0.176	-0.137	-0.212	0.099	0.132	0.213	0.224*	0.109	0.073	-0.07
	(0.147)	(0.130)	(0.162)	(0.144)	(0.141)	(0.136)	(0.127)	(0.126)	(0.141)	(0.11
China x College	-0.185*	-0.245*	-0.186	-0.110	-0.002	-0.031	0.018	0.058	-0.017	-0.01
	(0.099)	(0.125)	(0.122)	(0.102)	(0.127)	(0.124)	(0.095)	(0.128)	(0.146)	(0.13
China x MA Phd	-0.195**	-0.318**	-0.374***	-0.318***	-0.338***	-0.190	-0.105	-0.072	-0.151	-0.18
	(0.086)	(0.119)	(0.122)	(0.089)	(0.087)	(0.127)	(0.120)	(0.118)	(0.139)	(0.09
Philippines x HS Degree	-0.088	0.134	-0.109	-0.019	-0.018	0.123	-0.029	0.146	0.010	0.14
	(0.094)	(0.099)	(0.114)	(0.142)	(0.145)	(0.182)	(0.151)	(0.122)	(0.145)	(0.10
Philippines x Some College	-0.109	0.006	-0.259**	-0.116	0.015	-0.037	-0.041	0.083	-0.033	-0.04
	(0.092)	(0.084)	(0.121)	(0.136)	(0.130)	(0.127)	(0.134)	(0.099)	(0.125)	(0.09
Philippines x College	-0.081	0.087	-0.245**	-0.081	-0.008	-0.061	-0.152	-0.014	-0.145	-0.04
	(0.098)	(0.074)	(0.107)	(0.144)	(0.147)	(0.147)	(0.131)	(0.089)	(0.110)	(0.09
Philippines x MA Phd	0.049	0.062	-0.261**	-0.153	-0.079	-0.115	-0.181	0.020	-0.129	-0.07
	(0.141)	(0.093)	(0.128)	(0.150)	(0.163)	(0.167)	(0.147)	(0.104)	(0.138)	(0.11
Canada x HS Degree	-0.128	0.108	0.264	0.304*	-0.116	-0.002	0.011	0.096	0.024	0.06
	(0.338)	(0.222)	(0.284)	(0.160)	(0.352)	(0.388)	(0.365)	(0.355)	(0.353)	(0.35
Canada x Some College	-0.096	0.071	0.060	0.359***	0.025	-0.092	-0.101	-0.128	-0.069	-0.11

	(0.327)	(0.180)	(0.229)	(0.093)	(0.256)	(0.326)	(0.307)	(0.283)	(0.275)	(0.28
Canada x College	-0.273	-0.036	0.002	0.330***	-0.123	-0.091	-0.058	-0.019	-0.002	-0.08
	(0.341)	(0.151)	(0.232)	(0.088)	(0.280)	(0.311)	(0.287)	(0.267)	(0.262)	(0.26
Canada x MA Phd	-0.081	0.017	0.013	0.250***	-0.259	-0.332	-0.318	-0.180	-0.174	-0.21
	(0.324)	(0.154)	(0.235)	(0.091)	(0.315)	(0.328)	(0.302)	(0.315)	(0.274)	(0.28
Constant	0.289***	0.382***	0.455***	0.433***	0.423***	0.320***	0.438***	0.354***	0.251***	0.251*
	(0.036)	(0.029)	(0.038)	(0.032)	(0.031)	(0.031)	(0.038)	(0.036)	(0.039)	(0.04
Observations	4,200	4,200	4,200	4,200	4,200	4,200	4,200	4,200	4,200	4,20
R-squared	0.104	0.045	0.047	0.047	0.054	0.062	0.059	0.064	0.065	0.05

Note: Includes age fixed effects, year of entry and state fixed effects. Standard errors are clustered at the state of residence. Source: American Community Survey, 2005–2007 and IRS 2005-2015 W-2 and 1099 data. Numbers have been rounded to comply with the Census Bureau's disclosure-avoidance guidelines.

Table A7: Correlations of Return Migration and Changing Employer

	Men		Women	
VARIABLES	Return Migrant	Number of Unique EIN	Return Migrant	Number of Unique EIN
Number of Unique EIN	0.009***		0.010***	
	(0.003)		(0.003)	
Years Since Migration	-0.132***	0.219***	-0.135***	0.229***
	(0.001)	(0.008)	(0.002)	(0.011)
Return Migrant		0.232***		0.243***
		(0.072)		(0.085)
Constant	1.645***	0.086	1.569***	-0.770***
	(0.018)	(0.187)	(0.027)	(0.142)
Observations	5400	5400	4200	4200
R-squared	0.687	0.211	0.656	0.168

Note: Includes age fixed-effects, education controls, and state of residence fixed effects. Standard errors clustered at the state of residence. Source: ACS 2005–2007 and IRS 2005–2015 W-2 and 1099 data. Numbers have been rounded to comply with the Census Bureau's disclosure-avoidance guidelines.

Table A8: Heckman Selection Correction for Return Migration

Panel A: Men

VARIABLES	(1) Missing in 2007	(2) Missing in 2008	(3) Missing in 2009	(4) Missing in 2010	(5) Missing in 2011	(6) Missing in 2012	(7) Missing in 2013	(8) Missing in 2014
Log Earnings 2005	-0.024							
	(0.024)							
Log Earnings 2006		-0.075***						
		(0.029)						
Log Earnings 2007			-0.159***					
			(0.036)					
Log Earnings 2008				-0.204***				
				(0.030)				
Log Earnings 2009					-0.200***			
					(0.036)			
Log Earnings 2010						-0.224***		
						(0.033)		
Log Earnings 2011							-0.281***	
							(0.030)	
Log Earnings 2012								-0.291***
								(0.034)
Observations	2,400	3,500	4,000	4,000	3,700	3,300	3,500	3,400

Note: Includes state of residence fixed effects, age fixed effects, a constant, and educational category controls. Source: ACS 2005–2007 and IRS W-2s or 1099 data (2005–2015). Numbers have been rounded to comply with the Census Bureau's disclosure-avoidance guidelines.

Panel B: Women

VARIABLES	(1) Missing in 2007	(2) Missing in 2008	(3) Missing in 2009	(4) Missing in 2010	(5) Missing in 2011	(6) Missing in 2012	(7) Missing in 2013	(8) Missing in 2014
Log Earnings 2005	-0.106***							
	(0.022)							
Log Earnings 2006		-0.153***						
		(0.023)						
Log Earnings 2007			-0.259***					
			(0.021)					
Log Earnings 2008				-0.367***				
				(0.031)				
Log Earnings 2009					-0.307***			
					(0.022)			
Log Earnings 2010						-0.298***		
						(0.032)		
Log Earnings 2011							-0.261***	
							(0.023)	
Log Earnings 2012								-0.340***
								(0.036)
Observations	1,700	2,600	3,200	3,000	2,900	2,700	2,800	2,800

Note: Includes state of residence fixed effects, age fixed effects, a constant, and educational category controls. Source: ACS 2005–2007 and IRS W-2s or 1099 data (2005–2015). Numbers have been rounded to comply with the Census Bureau's disclosure-avoidance guidelines.