Economic Assimilation of Foreign-Born Workers in the United States: An Overlapping Rotating Panel Analysis

Seik Kim

Department of Economics

University of Washington

seikkim@uw.edu

`http://faculty.washington.edu/seikkim/`

August 2012

Abstract

This paper presents new evidence on whether foreign-born workers assimilate. We compare cross-section and panel models of the foreign-native gap in wage growth using the CPS for 1994-2004. The cross-section specification replicates the models in earlier cross-section studies and the panel specification simply adds individual fixed effects to the model. While the cross-section results are similar to those of previous studies, the longitudinal results suggest that the foreign-native gap in average wages widens with time since migration. It implies that controlling for fixed unobserved heterogeneity reverses the conventional result of economic assimilation. These patterns are robust to sample attrition and outmigration.

Keywords: Economic Assimilation, Immigration, Overlapping Rotating Panel Data, Attrition

JEL Classification Number: C23, J31, J61

[0]I am especially grateful to my advisor, Joseph Altonji, for his insightful suggestions, support, and encouragement. I also thank Yuichi Kitamura, Fabian Lange, and Mark Rosenzweig for their constructive advice and guidance. I have also benefited from helpful comments made by Donald Andrews, Adriana Kugler, Thomas Lemieux, Darren Lubotsky, Peter Phillips, Paul Schultz, Emiko Usui, and seminar participants at the annual meeting of the Society of Labor Economists, the summer meeting of the Econometric Society, Federal Reserve Bank of New York, Federal Reserve Bank of San Francisco, the UW Center for Studies in Demography and Ecology, the World Congress on National Accounts and Economic Performance Measures for Nations, Purdue University, Seoul National University, University of Washington, Vanderbilt University, Washington State University, and Yale University.

1 Introduction

The large and growing share of foreign-born workers in the United States has heightened interest in the economic impact of immigration. How immigrants fare as they accumulate experience in the U.S. labor market is the key to many of these effects.[1] First and foremost, the earnings of immigrants will directly affect the level and distribution of per capita income in the United States. Second, the better immigrants do on arrival and over time, the greater the extent to which their contributions as tax payers will outweigh their use of government services. Third, the greater the extent to which immigrants who enter the United States in low skill jobs quickly acquire country specific skills and spread into higher skill jobs, the smaller any negative impact on less skilled natives is likely to be.

This paper presents new evidence on whether foreign-born workers assimilate, which we define as the degree to which the wages of foreign-born workers approach those of comparable native-born workers with additional time spent in the United States. Assimilation rates are the net result of several offsetting factors. Upon entry into the U.S. labor market, foreign-born persons may earn lower wages than their native counterparts to the extent that human capital is not perfectly transferable across economies and cultures and because employers are likely to have less knowledge about their productivity. On the other hand, some groups of foreign-born workers might outperform natives if they possess superior skill endowments, stronger work ethics, or more powerful incentives. As immigrants stay longer in the United States, their wages might converge to those of natives.

The key econometric challenge to measuring assimilation rates is how to distinguish growth in earnings of particular immigrants from variation in initial skill levels associated with age at entry, year of entry, country of origin, and other factors. As Borjas (1985) points out, estimates of assimilation based on a single cross-section, such as Chiswick (1978), are biased if the ability and skill endowments of immigrants vary by year of entry. Studies using repeated cross-sections can control for the variation in skill composition by tracking the groups of individuals with same

[1] In U.S. immigration law the term "immigrant" or "permanent resident alien" denotes a person admitted to this legal classification. For expositional convenience, we use the terms "foreign-born person" and "immigrant" interchangeably, although our sample possibly includes aliens in an illegal status.

year of entry.[2] However, such studies are vulnerable to bias from individual heterogeneity within an immigration year cell. If migrant workers who arrive at older ages are more skilled than those who arrive at younger ages conditional on the year of entry, analyses of immigrant wage growth based on repeated cross-section studies may be biased upward by fixed unobserved heterogeneity.

This paper compares cross-section and panel analyses of assimilation using the same data set from the Current Population Survey (CPS) for 1994-2004. While the longitudinal model exploits the two-year panel aspect of the sample, the cross-section model ignores its panel structure. In this paper, the former is specified by simply adding individual fixed effects to the latter. We find that controlling for this heterogeneity reverses the conventional results of assimilation. While the cross-section results are similar to those of previous studies, the longitudinal results suggest that the foreign-native gap in average wages widens with time since migration. This finding is consistent with the hypothesized relationship between unobserved skills and age at migration. This is supported by explanations based on the Roy model and human capital theory.

Overall, we find little evidence of a narrowing of the foreign-native gap in economic performances with time since immigration for 1994-2004 in contrast to the literature based on repeated cross-sections for earlier years. New immigrant workers from Latin America earn lower wages than natives, and the wage gap widens with time spent in the United States. The wages of new immigrant workers from Europe and Asia exceed those of natives, and there is no strong evidence of subsequent convergence. These patterns across national origin groups are consistent with those reported in Schoeni (1997). The main findings of this paper on the comparison between the cross-section and panel results are robust to sample attrition and outmigration. These findings, however, are not directly comparable with longitudinal estimates reported in Duleep and Regets (1997a) or Lubotsky (2007) because they use different sampling periods.

The paper proceeds as follows. Section 2 provides an overview of the paper, defines economic assimilation, and outlines issues in the identification and estimation of economic assimilation. Section 3 introduces the data set and presents summary statistics. In addition, this section outlines a method of accounting for sample attrition when outmigration is not observed. Section 4 discusses

[2] See Borjas (1985) for a critique of studies based on single cross-sections and for the first application of a synthetic cohort analysis based on repeated cross-sections. See Borjas (1995) and LaLonde and Topel (1992) for studies using repeated cross-section analysis.

the main results. It contrasts the cross-section and panel results related to economic assimilation and demonstrates the key role of individual fixed effects in the estimation. This section also presents assimilation results separately for foreign-born workers by their continent of origin. Finally, Section 5 offers conclusions.

2 Issues in Measuring Economic Assimilation

2.1 An Overview

The current paper is motivated by the different foreign-native wage gaps observed in Figure 1. The figure depicts the mean hourly wages of cohorts of foreign-born and native-born male workers of various age groups during 1994-2004.[3] The sample is drawn from the CPS. The foreign-born workers in the figure are confined to those who arrived between 1980 and 1991. For the time being, assume that selective return migration is negligibly small. The three thicker lines with larger symbols indicate the mean wages of native-born workers and the three thinner lines with smaller symbols indicate the mean wages of foreign-born workers. The solid lines with squares track the mean wages of those who were 20-24 years old in 1994. These individuals become 21-25 in 1995, and so on. The dashed lines with triangles are the mean wages of those who were 30-34 years old in 1994. The dotted lines with circles correspond to the mean wages of those were 40-44 years old in 1994. Therefore, changes in the gaps between the thicker and the thinner lines of same type with identical symbols measure economic assimilation.

Figure 1 illustrates an idea of how economic assimilation can be measured. The wage gap between the immigrants and the natives in the "20-24 in 1994" cohort widens as the foreign-born workers stay longer in the United States. Foreign-born workers who were 20-24 years old in 1994

[3] The native-born workers in this paper are whites, but there are several alternative ways of choosing a native sample. One may compare wages of foreign-born individuals with those of native-born individuals regardless of ethnic origins, with wages of their ethnically similar native-born counterparts, or compare wages between earlier and later arrivals within the foreign-born population. We use native-born non-Hispanic white individuals because it gives the most conservative assimilation measure. Even with the most conservative definition, we show that cross-section results imply faster wage growth for immigrants than natives, which is consistent with the results in previous literature, but longitudinal results are against economic assimilation. Another reason we use the non-Hispanic white sample is that it is a solid reference group as the racial/ethnic composition of the native population has changed dramatically in recent years.

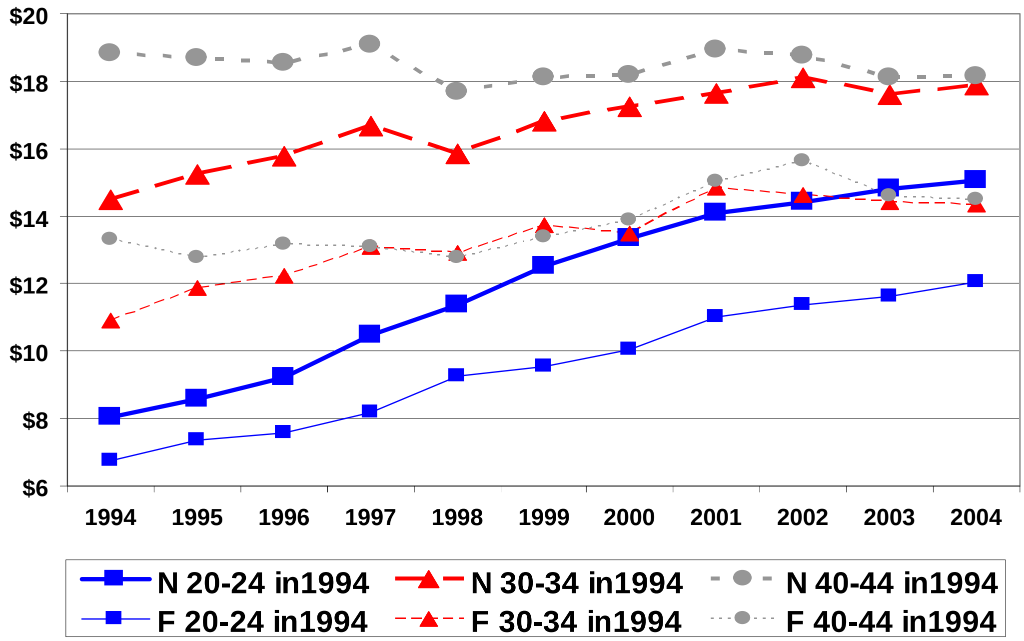

Figure 1: Average Wages (in 1994 Dollars) of Native-Born and Foreign-Born Workers

fail to assimilate economically during the 1994-2004 period. The foreign-born workers in the "30-34 in 1994" cohort also fail to catch up over the 1994-2004 period–the wage gap remains stable. The foreign-born workers in the "40-44 in 1994" cohort experience economic assimilation over the 1994-2004 period as the wage gap narrows. Overall, we observe that there is no assimilation for the younger cohorts, while for the older cohorts there seems to be some assimilation.

The main focus of this paper is to examine using our data whether estimation strategies employed in earlier repeated cross-section studies find the same patterns in the foreign-native wage gap that are apparent in Figure 1. The answer we find is that they do not. We replicate empirical specifications used in previous papers using our data and find that repeated cross-section approaches fail to detect the widening of the foreign-native gap in wages for the younger cohorts. Instead, the results of these specifications suggest that economic assimilation occurs across all age cohorts. One of the key findings of this paper is that it is only when we control for individual fixed effects that the estimation results are consistent with the trends in Figure 1.

We note the fact that our cross-section estimates based on the CPS for 1994-2004 are of a similar magnitude as those of the previous literature, but when we control for fixed unobserved heterogeneity the conventional results of assimilation are reversed. These findings, however, do not imply that the previous assimilation estimates that do not control for individual heterogeneity are all incorrect since these papers focus on earlier periods, while our results are from more recent samples. While we would like to observe such patterns in data for earlier years, it is not possible to obtain the equivalent of Figure 1 due to a lack of data for those years. Nonetheless, we are certain that, for 1994-2004, there is substantial bias in repeated cross-section estimates due to the neglected fixed unobserved heterogeneity. This claim is empirically verified in Section 4, and the next several sections provide some theoretical background for why the cross-section approach can be misleading.

2.2 Definition of Economic Assimilation

In this paper, economic performance is measured by hourly wages. Economic performance of an immigrant worker i is generated by

$$y_{it} = h_{imm}\left(age_{it}, ysm_{it}, edu_i, \mu_i, t, \varepsilon_{it}\right), \tag{1}$$

and that of a native worker n by

$$y_{nt} = h_{nat}\left(age_{nt}, edu_n, \mu_n, t, \varepsilon_{nt}\right), \tag{2}$$

for some known functions $h_{imm}(\cdot)$ and $h_{nat}(\cdot)$, where y is the logarithm of the hourly wage, age is the worker's age, ysm is the number of years since migration, edu is the number of years of education, μ reflects ability or skill endowment, t is calendar year, and ε captures idiosyncratic errors. Years since migration combined with age reflects an immigrant's gain such as information acquisition, human capital accumulation, and employer learning. Ability or skill endowment is not observed but may be correlated with year of entry, country of origin, and age at migration. We do not control for geographic or occupation variables as these are are outcomes, rather than determinants of assimilation.

The economic performance of a foreign-born worker relative to a native-born worker at time t

can be measured by

$$EA\,(age, ysm; t) = \frac{d}{dt}E\,[h_{imm}|age, ysm, t] - \frac{d}{dt}E\,[h_{nat}|age, t]\,. \tag{3}$$

Roughly speaking, $EA\,(age, ysm; t)$ is a difference-in-difference estimator. It reflects the rate of convergence in wages between foreign-born and native-born workers. Many studies find that average foreign-born workers initially earn lower wages than average native-born workers.[4] In this case, wage convergence from below toward the higher native mean, $EA\,(age, ysm; t) > 0$, means economic assimilation. For example, consider a 30 year-old foreign-born worker who has lived in the United States for 5 years. Suppose that his wage grows faster than the wage of a 30 year-old native-born worker. This represents economic assimilation because the wage gap between these two individuals will narrow in the following year. This paper also considers the case where foreign-born workers initially earn higher wages than native-born workers. In this case, a narrowing of the foreign-native gap in wages means wage convergence from above toward the lower native mean, $EA\,(age, ysm; t) < 0$.

Consider an empirical version of (1) and (2) that allows for different returns to skill between immigrants and natives:

$$\begin{aligned} h_{imm}\,(age_{it}, ysm_{it}, edu_i, \mu_i, t, \varepsilon_{it}) &= (\alpha_{nat} + \alpha)\,age_{it} + \delta ysm_{it} + (\beta_{nat} + \beta)\,edu_i + \mu_i + \gamma_t + \varepsilon_{it}, &(4)\\ h_{nat}\,(age_{nt}, edu_n, \mu_n, t, \varepsilon_{nt}) &= \alpha_{nat}age_{nt} + \beta_{nat}edu_n + \mu_n + \gamma_t + \varepsilon_{nt}. &(5) \end{aligned}$$

where γ_t reflects market conditions, business cycles, or economic shocks. We will generalize this simple model in several ways in Section 4. Then $EA\,(age, ysm; t)$ is given by

$$\begin{aligned} EA\,(age, ysm; t) &= \left(\alpha_{nat} + \alpha + \delta + \frac{d}{dt}\gamma_t\right) - \left(\alpha_{nat} + \frac{d}{dt}\gamma_t\right)\\ &= \alpha + \delta, \end{aligned}$$

and economic assimilation is equivalent to $\alpha + \delta > 0$. It is obvious that $\alpha + \delta$ is identified when

[4] Although we focus on the mean wages, the technique developed later in this paper can be applied to the entire distribution of wages.

a panel sample is available. In the next section, we discuss the identification of $\alpha + \delta$ when panel samples are not available.

2.3 Identification of the Measure of Assimilation

Probably the most popular data set used in the literature of assimilation is the U.S. Census. When we have repeated cross-section data, the measure of assimilation, $\alpha + \delta$, is identified under an assumption that μ_i is not correlated with observable variables conditional on the year of entry. This is a strong identification restriction, but is employed in almost all repeated cross-section studies:

$$E\left[\mu_i | i, t, age, ysm\right] = E\left[\mu_i | c\right] \quad \text{w.p.1 for all } t \text{ and } i \in c, \tag{6}$$

where c is the arrival year cohort. Technically, this is equivalent to simply including entry year fixed effects in a pooled cross-section model. In this section, we drop education and calendar year fixed effects from (4), (5), and (6) since excluding them does not change the theoretical results.

Suppose that we have two cross-sections, and individuals i and j are in the same arrival year cohort, but are observed in different cross-sections. From (4) and (6), we have

$$\begin{aligned} E\left[y_{it_1}^{imm} | i, t_1, age_1, ysm_1\right] &= \left(\alpha_{nat} + \alpha\right) age_{it_1} + \delta\left(t_1 - c\right) + E\left[\mu_i | c\right], \\ E\left[y_{jt_2}^{imm} | j, t_2, age_2, ysm_2\right] &= \left(\alpha_{nat} + \alpha\right) age_{jt_2} + \delta\left(t_2 - c\right) + E\left[\mu_j | c\right], \end{aligned}$$

where $t_2 > t_1$. Note that $E\left[\mu_i | c\right] = E\left[\mu_j | c\right]$ since $i, j \in c$ and that $ysm = (t - c)$. Taking the difference, we obtain

$$E\left[y_{jt_2}^{imm} | j, t_2, age_2, ysm_2\right] - E\left[y_{it_1}^{imm} | i, t_1, age_1, ysm_1\right] = \left(\alpha_{nat} + \alpha\right)\left(age_2 - age_1\right) + \delta\left(t_2 - t_1\right).$$

Therefore, $(\alpha_{nat} + \alpha)$ and δ are separately identified. Since α_{nat} is identified from native equations using a similar logic, $\alpha + \delta$ is identified.

In practice, the conventional identification assumption in (6) is not likely to hold. For example, age at migration, $age_{it} - (t - c)$, may be correlated with μ_i conditional on the year of entry.

Suppose that the correlation between μ_i and age at migration is given by

$$E\left[\mu_i | i, t, age, ysm\right] = \mu_c + \eta_i \left(age - (t - c)\right), \tag{7}$$

where η_i is individual specific. Since each of individuals i and j is observed only once, taking the difference does not eliminate the incidental parameter, η. In principle, a pseudo-panel approach can be used to identify $\alpha + \delta$, but in practice, averaging over cohorts will result in a very small sample size.

Suppose that we follow a group of persons with the same year of entry, c, and the same age at migration, aam, so that $age_1 - (t_1 - c) = age_2 - (t_2 - c)$, in different cross-sections. Technically, this is equivalent to conditioning on (c, aam) rather than on i. As a result, we have

$$\begin{aligned} E\left[y_{it_1}^{imm} | \left(c, aam\right), t_1, age_1, ysm_1\right] &= \left(\alpha_{nat} + \alpha\right) age_1 + \delta\left(t_1 - c\right) + \left(\mu_c + E\left[\eta_i | c, aam\right] \cdot aam\right), \\ E\left[y_{jt_2}^{imm} | \left(c, aam\right), t_2, age_2, ysm_2\right] &= \left(\alpha_{nat} + \alpha\right) age_2 + \delta\left(t_2 - c\right) + \left(\mu_c + E\left[\eta_j | c, aam\right] \cdot aam\right). \end{aligned}$$

Note that $E\left[\eta_i | c, aam\right] = E\left[\eta_j | c, aam\right]$ provided there is no outmigration between t_1 and t_2. Taking the difference, we obtain

$$\begin{aligned} E\left[y_{jt_2}^{imm} | c, aam, t_2, age_2, ysm_2\right] - E\left[y_{it_1}^{imm} | c, aam, t_1, age_1, ysm_1\right] &= \left(\alpha_{nat} + \alpha\right)\left(age_2 - age_1\right) + \delta\left(t_2 - t_1\right) \\ &= \left(\alpha_{nat} + \alpha + \delta\right)\left(t_2 - t_1\right), \end{aligned}$$

where the second equation holds due to $age_1 - (t_1 - c) = age_2 - (t_2 - c)$. Since α_{nat} is identified from native equations, $\alpha + \delta$ is identified, although α and δ are not separately identified.[5]

The pseudo-panel approach, however, requires the grouping of individuals into some cohorts, which results in a very small sample size. More importantly, when the functional form of $E\left[\mu_i | i, t, age, ysm\right]$ is not known, it is unclear how to group individuals. One possible grouping criterion may be by year of entry, country/continent of origin, and age at migration, but the resulting group size will be too small. In sum, the pseudo-panel approach for repeated cross-sections is certainly a possible

[5] It is not possible to control for age at migration in repeated cross-section analyses where age and years since migration are used as control variables. For example, Friedberg (1992) controls for age at migration at a cost of assuming that $\alpha = 0$.

option, but longitudinal data make it much easier to estimate assimilation. Before continuing with a discussion of available longitudinal data sets, the next section shows how age at migration may be correlated with μ_i.

2.4 Age at Migration and Fixed Unobserved Heterogeneity

We claim that among the immigrants of the same year of entry, older persons have higher μ_i than younger ones.[6] There are possibly two explanations for the positive correlation between μ_i and age at migration.

The first one is based on the Roy model with search. Suppose that an individual receives random wage offers from the United States. It has to be the case that for an older individual to migrate, the wage offer from the U.S. has to be larger than an offer for a younger individual because the remaining working life of the older person is shorter. Therefore, older new immigrants, on average, will have higher wages than younger new immigrants.

The second explanation relies on the human capital hypothesis. Suppose that schooling is more costly in the United States than in the home country. Then individuals with higher skill endowment or ability will invest in their human capital in their home country and delay the timing of immigration. Foreign-born workers will still initially earn lower wages when they migrate than native-born workers because human capital is not perfectly transferrable. An important point related to the second explanation is that this argument is only valid for immigrants who enter as adults. Among immigrants who enter the United States as children, Bleakley and Chin (2004) find that age at migration is negatively correlated with assimilation due to language proficiency. Section 4 shows that our results are consistent with both predictions.

2.5 Advantages and Disadvantages of the Available Longitudinal Data Sets

An ideal sample for estimating economic assimilation would be a longitudinal data set containing a large representative sample of foreign-born and native-born persons. Longitudinal data on native-

[6] This is consistent with other papers including Borjas (1987).

born and foreign-born populations permit one to control for fixed unobserved heterogeneity by tracking specific individuals over time. In practice, longitudinal analysis of U.S. immigrants has been limited by two key factors. First, sample sizes of immigrants in U.S. panels such as the Panel Study of Income Dynamics (PSID) or National Longitudinal Survey of Youth 1979 (NLSY79) are too small. Second, there is a nonrandom attrition problem. Furthermore, outmigration of the immigrants poses a fundamental problem for both panel and cross-section analyses to the extent that it is related to wage growth. It is a complicated problem since the data does not reveal who emigrated from the United States.

Several studies do use longitudinal samples, but most of the panels have few foreign-born workers or are for non-representative samples. For instance, Chiswick (1980) uses the National Longitudinal Survey (with 98 male immigrants who all arrived before 1965) and Borjas (1989) uses a longitudinal survey of scientists and engineers. A representative random sample of permanent residents from the Immigration and Naturalization Service for fiscal year 1971 used by Jasso and Rosenzweig (1988) does not include wage information. Duleep and Regets (1997a) use the matched June 1987 and June 1988 CPS, but the sample period is short and the sample size is small. More recently, Lubotsky (2007) constructs a longitudinal sample by linking the 1990/1991 SIPP and the 1994 March CPS to Social Security earnings data for 1951-1997. He collects samples of individuals with known social security numbers from cross-sections and connect time series of their past social security earnings. It is possible that these linked data underrepresent immigrants working in unreported sectors or the underground economy.

Given that an ideal sample is not available, it is desirable to have a longitudinal data set which enables one to control for sample attrition and outmigration. As we show, one may do so with an overlapping rotating panel data set, such as the CPS. The data structure of an overlapping rotating panel is depicted in Figure 2, where the short panels are represented by the blocks. Vertical circles symbolize its longitudinal feature, which means that one may use usual panel data tools, such as the first difference or the fixed effects models. The estimators based on multiple short panels are consistent, although they are less efficient than estimators based on a longer panel. Horizontal circles illustrate the overlapping feature of the short panels, and the sample is a representative cross-section of the target population for any given time period. While a major disadvantage of

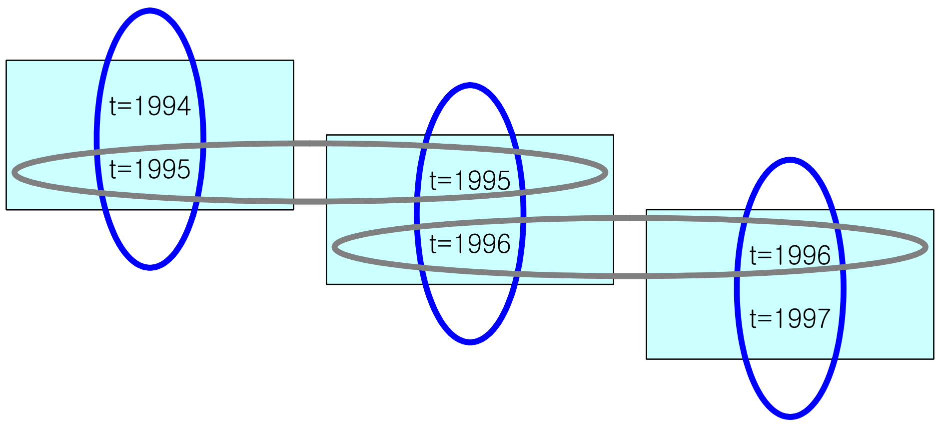

Figure 2: Data Structure of an Overlapping Rotation Panel Data Set

this sample is high attrition, the availability of representative cross-sections enables one to correct for attrition. Later, we discuss how to address attrition problems.

3 Data Description

3.1 The Current Population Survey and its Merged Outgoing Rotation Group

The CPS sample is a collection of representative cross-sections. As of July 2001, the CPS collects a sample of approximately 56,000 housing units from 792 sample areas. Each month, data are collected from the sample housing units on demographic and labor force characteristics of the civilian non-institutional population 16 years of age and older. Since 1994, the CPS includes information on international migration, such as year of entry into the United States and country of birth along with demographic and labor market information, such as age, schooling, marital status, earnings per hour or week, usual hours of work, and labor market status.[7]

The design of the CPS is as follows. A housing unit is interviewed for 4 consecutive months, is dropped out of the sample for the next 8 months, is brought back in the following 4 months, and then is retired from the sample.[8] If a household is included in either the first or the last 4

[7]Prior to 1994, CPS supplements on immigration were administered to all households participating in the survey in November 1979, April 1983, June 1986, June 1988, and June 1991.

[8]About 3/4 of the first and fifth interviews are conducted by visiting. In other interview months, almost 90% of the interviews are conducted over the phone. The rotation scheme ensures that in any 1 month, one-eighth of the

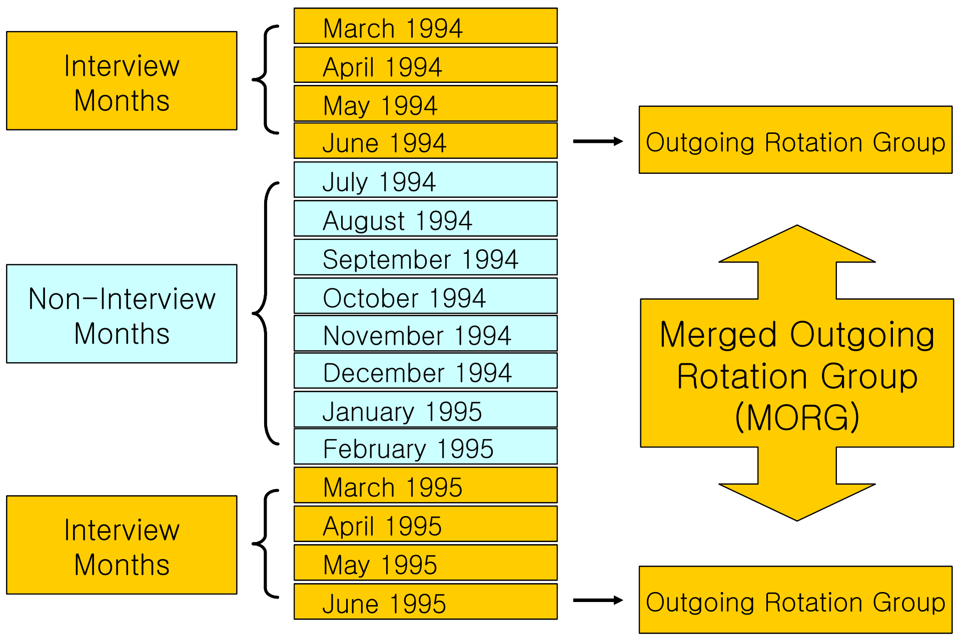

Figure 3: Sample Design of the CPS and its Merged Outgoing Rotation Group

months of the interview periods, it is said that the household is in the rotation group. Figure 3 demonstrates the sample design for a housing unit which, for instance, joins the survey in March 1994. This housing unit is interviewed from March to June in 1994 and 1995. The pre-selected housing units are kept unchanged over the interview periods. If the occupants of a dwelling unit move, the new occupants of the unit are interviewed. Although the interviewees may be replaced by new occupants within the sampling periods, the CPS provides a representative cross-section of the target population because the random sample of housing units is kept fixed.

An interesting feature of the CPS sample is its rotation scheme. Selected questions on labor market information, such as usual weekly earnings and usual weekly hours worked, are asked only in the last interview of each 4-month rotation group. The sets of households in the fourth or eighth month are called the outgoing rotation groups. If records from the 4th and 8th interviews are appended, we get repeated observations on the same individuals. The appended sample is called the Merged Outgoing Rotation Group (MORG) data. (See Figure 3.) By construction, an

housing units are interviewed for the first time, another eighth is interviewed for the second time, and so on. That is, after the first month, 6 of the 8 rotation groups will have been in the survey for the previous month; there will always be a 75 percent month-to-month overlap. When the system has been in full operation for 1 year, 4 of the 8 rotation groups in any month will have been in the survey for the same month, 1 year ago; there will always be a 50 percent year-to-year overlap.

individual appears only once in a year, but may reappear in the following year. Due to the 4-8-4 rotation scheme, the CPS MORG is an overlapping rotating panel data set comprised of multiple panels two years in length. The 1994-1995 panel, for instance, contains the individuals in the households which enter the survey scheme between October 1993 and September 1994.

3.2 Summary Statistics

The sample used in this analysis is drawn from the CPS MORG between 1994 and 2004. We take a sample of foreign-born and native-born men of ages 18-64.[9] We define an individual as matched if the individual appears twice in the CPS MORG. In order to examine differences based on ethnic origin, we divide the foreign sample into four groups: immigrants from Latin America, from Europe (including Australia, New Zealand, and Canada), from Asia, and from other countries.[10] The group of "other" countries consists of immigrants from Africa, Oceania, and unclassified ones. The last group is of little interest due to its small sample size and heterogeneity. Details on how the data are processed are explained in the Appendix. This section provides a general picture.

Table 1 reports summary statistics for cross-section/matched and all/reported wages samples. The summary statistics for the matched sample are the first year observations. In this section we focus on the cross-section sample with all individuals. Years of education provides a rough measure of skill endowment. Foreign-born persons have a lower mean and a much larger standard deviation of education. In the cross-section sample with all the individuals, the average education level is 13.6 years for native-born persons and is 12.0 years for foreign-born persons. Immigrants from Latin America have 10.0 years of average education, those from Europe 13.7 years, those from Asia 14.2 years, and those from the other countries 13.7 years.

The wage information in the CPS sample is mostly self-reported, but also involves imputed

[9] The foreign sample includes foreign-born men who were not U.S. citizens at the time of birth. Following Warren and Peck (1980), our foreign sample consists of persons born outside the United States, the Commonwealth of Puerto Rico, and the outlying areas of the United States. Foreign-born persons may have acquired U.S. citizenship by naturalization or may be in illegal status. The reference group consists of non-Hispanic native-born white men. The native sample includes persons born in the Unites States, but excludes persons born in the Puerto Rico and the outlying areas.

[10] We combine Australia, New Zealand, and Canada with Europe because of sample size considerations and so that immigrants from countries that are predominantly white and are at a similar stage of political and economic development are grouped together. We refer to the group as Europe. The data do not identify mother tongue. The impact of language proficiency has been studied in a large literature. LaLonde and Topel (1997) provide a survey.

wages. Throughout the sample period, an increasing fraction of workers do not answer questions about wages. When a person is working but does not report the wage, the Census Bureau assigns values for the missing wages using an allocation rule which is known as the cell hot deck match criteria.[11] The native imputation rates are about 17-23% with an increasing trend from September 1995 through 2004. The foreign imputation rates are higher than the native ones by 2-4 percentage points. The imputation rates are homogeneous across different ethnic groups.

In Table 1, we observe that mean characteristics of persons with reported wages are different from those in the entire sample, especially among foreign-born workers. For instance, the imputed wages for those from Latin America are higher than the reported wages and those from Europe and Asia are lower. As the imputation rule does not account for the country of origin, the imputed wages of immigrant workers tend to be biased toward the wages of native workers. Consequently, our preferred way to handle the imputed wages is simply dropping them.[12]

The average hourly wage of native-born workers is \$16.0-16.2, in 1994 dollars, while the average foreign-born worker earns \$12.8-13.0. Immigrants from Latin America make \$9.4-9.8 per hour, those from Europe \$18.4-19.6, those from Asia \$17.0, and from the other countries \$13.9-14.7. The estimates also indicate that foreign-born persons are about 2 years younger than native-born persons on average. Immigrant workers work 1.3-1.4 less hours per week than native workers. 79.0% and 78.7% of the foreign-born and native-born populations are full-time workers, while 5.4% and 5.8% are part-time workers, respectively. Although not reported in the table, the proportions of full-time and part-time workers are relatively stable over the sampling period: 75-82% and 5-7% of the foreign-born population and 76-80% and 5-6% of the native-born population are full-time and part-time workers, respectively. A larger proportion of the foreign-born population is married.

[11] According to the imputation rule, a value of the wage is allocated based on the cell of same gender, age, race, education, occupation, hours worked and receipt of tips, commissions, or overtime. (The numbers of cells are 14976 in 1994-2002 and 11520 in 2003-2004.)

[12] Hirsch and Schumacher (2004) raise the problem of imputed wages. They find that regression estimates including variables not used in imputation rules, such as union status, are biased. As country of origin is not used as imputation criteria, using the whole sample may bias the results. Bollinger and Hirsch (2006) propose a weighting scheme to correct for the bias. These methods do not affect our results qualitatively. We provide results using the entire sample as well as using weights which is suggested by Bollinger and Hirsch (2006) as a robustness check in the Appendix. The results do not change qualitatively.

Table 1. Summary Statistics

	Cross-Section Sample				Matched Sample			
	All Wages		Reported Wages		All Wages		Reported Wages	
	Native	Immigrant	Native	Immigrant	Native	Immigrant	Native	Immigrant
Age	41.1	39.4	41.4	39.4	42.5	40.8	42.8	40.8
	(12.1)	(11.6)	(12.3)	(11.7)	(11.3)	(11.2)	(11.4)	(11.3)
Education	13.6	12.0	13.7	11.9	13.7	12.1	13.7	11.9
	(2.4)	(4.3)	(2.4)	(4.3)	(2.4)	(4.3)	(2.5)	(4.4)
Latin America		10.0		9.9		10.1		9.9
		(4.1)		(4.3)		(4.2)		(4.2)
Europe		13.7		13.8		13.7		13.7
		(3.3)		(3.3)		(3.3)		(3.4)
Asia		14.2		14.2		14.3		14.3
		(3.4)		(3.4)		(3.4)		(3.4)
Others		13.7		13.6		13.7		13.5
		(3.5)		(3.6)		(3.6)		(3.7)
Hourly Wage	16.0	13.0	16.2	12.8	16.5	13.5	16.6	13.5
	(15.5)	(12.9)	(15.2)	(13.1)	(15.3)	(13.5)	(15.4)	(14.4)
Latin America		9.8		9.4		10.2		9.8
		(7.2)		(6.8)		(7.3)		(7.2)
Europe		18.4		19.6		18.9		20.4
		(18.6)		(19.8)		(19.6)		(21.3)
Asia		16.5		17.0		17.0		17.8
		(15.5)		(16.9)		(16.3)		(18.3)
Others		14.7		13.9		14.6		14.7
		(15.9)		(13.8)		(15.0)		(15.2)
Weekly Hours Worked	43.4	42.0	43.6	42.3	43.8	42.3	44.2	42.9
	(10.5)	(9.5)	(10.9)	(9.8)	(10.3)	(9.6)	(10.9)	(10.3)
Full Time	0.787	0.790	0.746	0.750	0.814	0.810	0.767	0.760
Part Time	0.058	0.054	0.058	0.052	0.049	0.050	0.050	0.050
Marital Status	0.640	0.680	0.639	0.682	0.696	0.730	0.699	0.739
U.S. Citizen	1.000	0.385	1.000	0.387	1.000	0.440	1.000	0.434
Latin America		0.513		0.529		0.497		0.508
Europe		0.163		0.161		0.179		0.181
Asia		0.256		0.254		0.265		0.262
Others		0.068		0.056		0.059		0.049
Observations	872598	126240	578519	82630	254837	34018	167981	20718

Standard deviations are reported in parentheses.

All Wages: reported & imputed wages; Reported Wages: reported (non-imputed) wages only.

Marital Status = 1 if married. U.S. Citizen = 1 if citizen.

Matching is directly related to residential mobility and outmigration as the housing units in the sample are kept fixed over the interview periods, provided that the non-interview rate is low.[13] Between 1994 and 2004, the attrition rates are 28-40% among the immigrant samples and 22-32% among the native samples.[14] We observe that persons in the matched samples, regardless of ethnic origins, tend to earn more, work longer, and participate more in the labor market than those in the cross-section samples. It implies that more successful workers are more likely to be matched than unsuccessful ones. Foreign-born persons from Latin America tend to attrite more than those from Europe and Asia. The consequence of nonrandom attrition, however, has not been addressed in immigration studies using the matched CPS.[15] We find substantial sample attrition, and the next section discusses the attrition problem in detail.

3.3 Accounting for Sample Attrition when Emigrants are Unobserved

Suppose that there is no emigration. Then the target population is stationary. Hirano, Imbens, Ridder, and Rubin (2001) and Bhattacharya (2008) develop an attrition correcting method that uses the availability of representative cross-sections as the basis for weighting the persons in a balanced part of the panel. They show that the sample attrition process, as a function of both past and current variables, can be identified under fairly flexible assumptions up to a known link function such as the logit or probit. The attrition correcting weighting function is given by the inverse of one minus the probability of sample attrition.

When international migration is possible and some immigrants go back to their home country, the above method should not be applied since the second period cross-section sample is a nonrandom subset of the first period population. This is called the population attrition. What makes it more complicated is the fact that data do not tell us who emigrated from the United States. More specifically, when a foreign-born respondent is missing in the second period, it is not possible to

[13] The average yearly non-interview rates for the CPS in the early 1990's are as low as 4-7%. This non-interview rate is comparable with the initial non-response rate of the National Longitudinal Survey of Youth 1979 (NLSY79), which is 10%.

[14] In practice, matching is not possible between June 1994 - August 1995 and June 1995 - August 1996 due to sample redesign. If samples in 1994-1995 and 1995-1996 are excluded, the attrition rates are 28-35% among the immigrant samples and 22-29% of the native samples. The gaps between the foreign and native attrition rates are stable in these periods ranging 6-8% points. A part of the gap in the attrition rates may be due to outmigration.

[15] While many papers have used the matched CPS, only two that we are aware of focus on immigration: Duleep and Regets (1997a) and Bratsberg, Barth, and Raaum (2006).

tell whether the person is in the United States or has gone back to his or her home country.

We draw on recent work by Kim (2012) to address the problem of sample attrition in the presence of unobserved population attrition and confirm that accounting for attrition does not alter our results qualitatively. The key idea of attrition correction is generating a counterfactual cross-section where there is no outmigration prior to applying the existing sample attrition correcting scheme. For example, suppose that the two-year panel of 1996-1997 is of interest. The CPS provides 1996 and 1997 cross-sections, but the 1997 cross-section is not representative of the 1996 population due to return migration. First, we use the 1996 cross-section as the basis for generating a representative counterfactual 1997 cross-section. The counterfactual sample is obtained by weighting the second period cross-section by one minus the probability of outmigration. The outmigration process can be identified when repeated cross-sections are available without knowing who emigrated from the United States under some strong assumptions. Then the two representative cross-sections (the 1996 actual and 1997 counterfactual cross-sections) are used as the basis for estimating attrition correcting weighting functions. In this step, existing sample attrition correcting method can be applied. Finally, we assign weights to the persons in the balanced part of the 1996-1997 panel. The exact formulae are in the Appendix, and the attrition correcting weighting function estimates are available upon request.

4 Empirical Evidence of Economic Assimilation

4.1 Estimates of Economic Assimilation

Based on the model given in (4) and (5), this paper considers two sets of empirical specifications. The first specification uses a panel approach, which we call the individual heterogeneity (IH) model. An example of this model is given by

$$\begin{aligned} y_{it}^{imm} &= (\alpha_{nat} + \alpha)\, age_{it} + \delta ysm_{it} + (\beta_{nat} + \beta)\, edu_i + \mu_i + \gamma_t + \varepsilon_{it}, && (8) \\ y_{it}^{nat} &= \alpha_{nat} age_{it} + \beta_{nat} edu_i + \mu_i + \gamma_t + \varepsilon_{it}, && (9) \end{aligned}$$

where γ_t reflects business cycles and ε_{it} captures idiosyncratic shocks.[16] In this specification, both immigrants and natives are indexed by an i for simplicity. The IH model in (8) and (9) allows fixed unobserved heterogeneity such as variation in skill endowments within the groups of individuals who entered in the same year. Estimation of the IH model requires a longitudinal sample.

The second specification uses a repeated cross-section approach, which we call the cohort heterogeneity (CH) model. This model is extensively used in earlier assimilation literature using repeated cross-sections. An example is given by

$$y_{it}^{imm} = (\alpha_{nat} + \alpha)\, age_{it} + \delta ysm_{it} + (\beta_{nat} + \beta)\, edu_i + \mu_c + \lambda_b + \gamma_t + \varepsilon_{it}, \tag{10}$$

$$y_{it}^{nat} = \alpha_{nat} age_{it} + \beta_{nat} edu_i + \gamma_t + \varepsilon_{it}, \tag{11}$$

where μ_c is arrival year cohort effects and λ_b is a birth country indicator. In this specification, year of entry, age at entry, and country of origin control for fixed unobserved heterogeneity. As the individual heterogeneity within an immigration year cell is neglected, the model given in (10) and (11) is the CH model. Estimation of the CH model requires repeated cross-sections.

The IH wage equations, (8) and (9), and the CH wage equations, (10) and (11), are linear in age and years since migration, but we also estimate the wage equations specified by quadratic and cubic polynomials in age and years since migration. These equations are estimated in two ways: with attrition correcting weights and without weights.[17] The coefficient estimates are reported in Tables A2-1 and A2-2 in the Appendix. When age and years since migration enter as polynomials, it is difficult to read the implications of the coefficient estimates. Hence, we present the regression results by predicting the wage path of a foreign-born worker who arrives in the United States at age 20 as many other studies do. This is a reasonable assumption since the average age is about 40 and the average years since migration is about 20 in our data.

Table 2 reports the economic assimilation estimates, $EA\,(age, ysm)$, evaluated at $(age, ysm) = (24, 4)$, $(32, 12)$, $(40, 20)$, and $(48, 28)$. The upper panel presents the IH estimates and the lower

[16] We implicitly make an assumption that aggregate economic shocks affect the wages by the same percentage amount to foreign-born and native-born workers. This restriction is proposed by Borjas (1985). For details see Borjas (1999).

[17] The main (wage) equations use the matched longitudinal sample of workers with positive wages. In this step, we exclude individuals with too high or too low wages and negative potential experience.

panel the CH estimates. These estimates measure the foreign-native difference in wage growth rates. Positive values indicate that immigrant wages grow at a faster rate than native wages at specific (age, ysm). Accompanied by the fact that the mean wage of foreign-born workers is below the native mean, positive estimates implies that wages of immigrants and natives converge. Negative estimates imply that immigrant and native wages diverge. The estimates are reported in percentage points. For example, –0.25 in the first line of the first column is interpreted as each additional year in the United States immigrant wages grow at a slower rate than native wages by 0.25 percentage points when sample attrition and outmigration are accounted for. This estimate is derived from the observation that immigrant wages grow annually by 2.13% and native wages by 2.38% under the assumption that year fixed effects on the level of wages are constant between two adjacent years. The difference is –0.25 percentage points and is not statistically different from zero.[18]

The first three columns in the upper panel of Table 2 present the IH estimates of economic assimilation that accounts for sample attrition and outmigration. The attrition-adjusted estimates from the quadratic specification suggest that wages of foreign-born workers grow slower than those of native-born workers by 1.17 percentage points per year at age 24. When they become 32, the speed of divergence slows down, but immigrant wages still grow slower than native wages by 0.75 percentage points per year. These assimilation estimates are statistically different from zero. From the cubic specification, we find that wages of foreign-born workers grow slower than those of native-born workers by 1.49 percentage points at age 24 and by 0.55 percentage points at age 32. The nonlinear specification results reveal that young foreign-born workers fall behind rather than catch up.

These findings can be described graphically. Using the attrition-adjusted IH estimates from the quadratic specification, it is possible to generate a wage growth path. Let the hypothetical foreign-born and native-born persons have the same wage at age 20. Then the foreign-native difference in wages is zero at age 20. The coefficient estimates suggest that at age 24, the foreign-native difference in log wages is –0.0509. The solid line with circles in Figure 4 plots the foreign-native difference

[18] To be precise, one-sided test should be used instead of a two-sided test, as the alternative hypothesis is given by either $EA\,(age, ysm) > 0$ or $EA\,(age, ysm) < 0$.

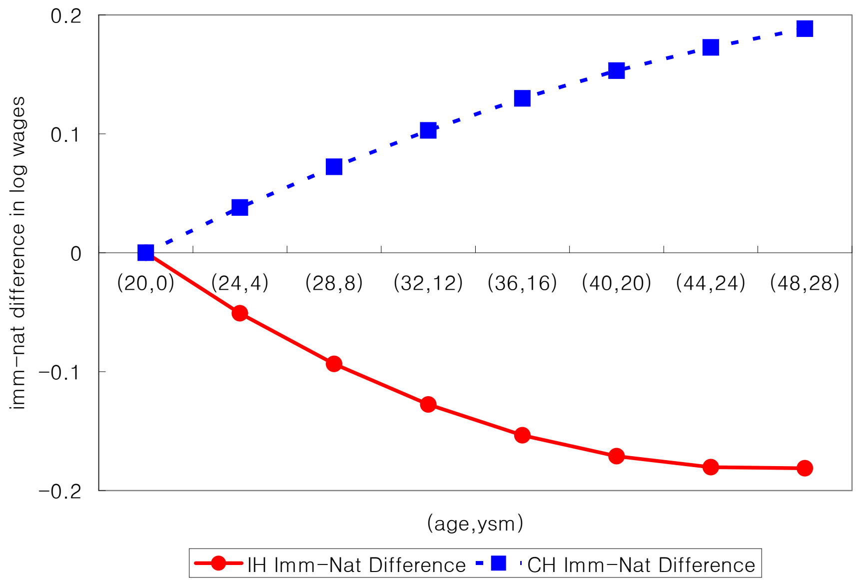

Figure 4: Immigrant-Native Difference in Simulated Log Wages (using the Quadratic Model Estimates)

in log wages, which is comparable to the observed wage difference in Figure 1. By definition, the measure of economic assimilation is the slope of the line. At age 24, the slope is –0.0117 meaning that the wages of immigrants grow slower than that of natives by 1.17 percentage points as reported in Table 2. According to Table 2, the slope estimate is negative and is statistically different from zero. Similarly, at age 32, the slope estimate is –0.0075 and is statistically different from zero. The foreign-native difference in wages stops widening above age 40. The slope or the measure of economic assimilation becomes close to zero.

The last three columns report unadjusted estimates. In general, the unadjusted estimates are not very different from the attrition-adjusted ones. Since the signs of estimated assimilation measures do not change, there is little evidence of assimilation whether or not attrition is corrected for.

Our findings are strikingly different from the results in the previous literature. For instance, using the 1970, 1980, and 1990 Census cross-sections, Borjas (1999) reports that the relative wage growth of immigrants is 0.60-0.76 percentage points higher per year during the first 10 years and 0.38-0.50 percentage points higher per year during the first 20 years based on CH models. Lubotsky (2007) estimates an IH model for 1951-1997 data and finds that the earnings of immigrants have grown 0.50-0.65 percentage points per year during the first twenty years since migration relative to the earnings of native-born workers with similar characteristics.[19] Since our data cover 1994-2004, it is not possible to directly compare Lubotsky's results with ours. Moreover, it is not possible to compare our results with previous CH results because they utilize different methodologies and cover different sampling periods. However, it is possible to investigate whether CH and IH results differ from each other using our data for 1994-2004. To replicate the CH models using our sample, we intentionally ignore its panel structure by dropping the second period observations from the longitudinal samples and construct cross-sectional data.

From the comparison between the CH and the IH results using our data, we conclude that assimilation estimates based on CH models appear to be biased upward, at least for 1994-2004. The lower panel of Table 2 shows the economic assimilation estimates using the repeated cross-section approach using the same data. These results misleadingly suggest that there is significant economic assimilation. However, the estimates are surprisingly similar to the results in the previous literature, which is consistent with the previous repeated cross-section studies. In the first column, applying the linear CH model, the estimate is 0.99 and is statistically different from zero at the 1% significance level. It implies that with each additional year in the United States immigrant wages grow faster than native wages by 0.99 percentage points when sample attrition and outmigration are accounted for. In the second column, the quadratic specification results suggest that wages of a foreign-born worker grow at a faster rate than those of a native-born worker by 0.93 percentage points per year at age 24. When they become 32, immigrant wages still grow at a faster rate than native wages by 0.74 percentage points per year. At age 40, immigrant wages are growing 0.56 percentage points faster than native wages. The cubic specification results in the third column suggest that wages of foreign-born workers grow faster than those of native-born workers by 0.70 percentage points at age 24 and by 0.69 percentage points at age 32. We find similar patterns in models without attrition correcting weights.

[19] His results also suggest that repeated cross-section estimates overstate assimilation.

Table 2. Economic Assimilation Estimates in Percentage Points

	Attrition-Adjusted			Not Adjusted		
	linear	quadratic	cubic	linear	quadratic	cubic
Individual Heterogeneity Model Estimates						
age=24, ysm=4	–0.25	–1.17**	–1.49**	–0.18	–1.15**	–1.44**
	(0.31)	(0.55)	(0.68)	(0.30)	(0.54)	(0.68)
age=32, ysm=12		–0.75**	–0.55		–0.78**	–0.70*
		(0.35)	(0.39)		(0.35)	(0.38)
age=40, ysm=20		–0.33	0.05		–0.40	–0.16
		(0.32)	(0.47)		(0.32)	(0.47)
age=48, ysm=28		0.08	0.33		–0.03	0.18
		(0.48)	(0.53)		(0.47)	(0.52)
Cohort Heterogeneity Model Estimates						
age=24, ysm=4	0.99***	0.93**	0.70	0.95***	1.05***	0.77
	(0.21)	(0.36)	(0.52)	(0.21)	(0.37)	(0.55)
age=32, ysm=12		0.74***	0.69***		0.83***	0.76***
		(0.24)	(0.24)		(0.25)	(0.25)
age=40, ysm=20		0.56***	0.64**		0.60***	0.69**
		(0.21)	(0.27)		(0.21)	(0.27)
age=48, ysm=28		0.37	0.56*		0.37	0.56
		(0.30)	(0.33)		(0.30)	(0.33)

Standard errors are reported in parentheses. Confidence levels: 99% (***), 95% (**), 90% (*).

Estimates represent immigrants' annual percentage wage growth relative to natives' percentage wage growth.

The broken line in Figure 4 with squares uses the attrition-adjusted CH estimates from the quadratic specification to plot the foreign-native difference in log wages. Again, let the hypothetical foreign-born and native-born persons have the same wage at age 20. The estimation results suggest that at age 24, the foreign-native difference in log wages is 0.0380 and is increasing at 0.0093 or 0.93 percentage points per year. According to Table 2, the slope estimate is statistically different from zero. Similarly, at age 32, the slope estimate is 0.0074 and is statistically different from zero. In Figure 4 the simulated wage difference path is increasing even at age 48. In Table 2, the wages of 48 years old immigrants grow faster than that of 48 years old natives by 0.37 percentage points per year, although this estimate is not statistically different from zero.

4.2 Discussion

We address three issues regarding the validity of our estimation results. The first is whether the CH estimates are always biased. The empirical findings in Table 2 indicate a positive correlation between μ_i and age at migration conditional on age, year of entry, and other observables. As we have discussed in Section 2, however, whether there is a correlation between μ_i and other regressors is an empirical question. When μ_i does not cause any problems in estimation, CH specifications may be used. While it is not possible to test the positive correlation for earlier periods using our CPS data, there is some evidence that the IH type models for earlier periods support economic assimilation. For example, Duleep and Regets (1997a), matching the June 1987 and June 1988 CPS, find that the wage growth of foreign-born workers exceeds that of native-born workers by 0.3 percentage points. Therefore, the bottom line is that one needs to be careful in using the repeated cross-section approach since we do not have prior knowledge about the direction of the bias.

Second, Bleakley and Chin (2004) find that among the immigrants who immigrated to the United States as children, those who entered as younger children assimilate faster than those who entered as older children due to language proficiency. To show that our results are consistent with theirs, we estimate economic assimilation after dropping from our sample immigrants who entered as children. We find more negative assimilation estimates, which is consistent with their findings. (See the last three columns of Table A1-2 in the Appendix.) Therefore, this paper shows that among new adult immigrants, those who are older have higher μ_i than those who are younger.

Third, there is an issue of how much we can trust estimates based on annual variation in wages. One may argue that it is difficult to say anything about the nature of changes in wages over time. We disagree. Suppose that observed wages include measurement errors or economic shocks, and they are classical additive errors. Then, due to attenuation bias, the resulting estimates will be biased toward zero making it more difficult to find statistically significant results. This also implies that it is more difficult to find significant results from one-year-interval panels than ten-year-interval panels. Therefore, our CH and IH results which are statistically significant based on a one-year-interval panel are strong results. In addition, if there were no noise in wage observations, both the CH and the IH estimates would be even further away from zero, strengthening our findings.

Finally, a related issue is whether the two-year panels are too short to capture lifetime patterns of economic assimilation. However, this is not a concern. In principle, a single two-year panel is sufficient to identify economic assimilation as the sample covers individuals ranging in ages from their 20's to 60's. Nevertheless, it is useful to have multiple two-year panels because the influence of business cycles on the wage structure can be neutralized by observing many calendar year samples.

4.3 Economic Assimilation by Ethnic Origin

Given that there is little evidence of economic assimilation in general for 1994-2004, a natural and interesting question is whether some ethnic groups do assimilate economically while others do not. Table 3 reports estimates of economic assimilation using reported wages by ethnic origin. In this stage, we use the previously calculated weights instead of estimating them from each ethnic group.

The first panel presents economic assimilation of immigrants from Latin America. From the attrition-adjusted estimates of nonlinear specifications, we learn that their wages grow slower than native wages by 1.41-2.23 percentage points at age 24 and by 0.39-0.76 percentage points at age 32. As they become more experienced, there is no significant difference in relative wage growth compared with native-born workers. Although the gaps in wage growth disappear when they get older, there exists a wage gap. We also find that the assimilation measure estimates of European and Asian immigrant workers are insignificant. It implies that their wage growth paths are parallel with native wage growth path, and there is little evidence to support the hypothesis of economic assimilation.

As a robustness check, Tables A1-1 and A1-2 in the Appendix provide assimilation estimates using different samples and methods. Table A1-1 reports estimates using all the individuals. In addition, following Bollinger and Hirsch (2006), the first six columns in Table A1-2 report estimates when individuals with reported wages are weighted by the inverse probability of reporting wages. The weights correct for nonrandom selection of not reporting wages and are obtained from linear index logit models by country of origin, using age, years since migration, education, citizenship status, and marital status. The last three columns in Table A1-2, as was discussed before, report estimates when we drop foreign-born persons who immigrated before age 18. Dropping these persons significantly diminishes the sample sizes, but the results strengthen our findings.

Table 3. Economic Assimilation Estimates in Percentage Points (from Individual Heterogeneity Models)

	Attrition-Adjusted			Not Adjusted		
	linear	quadratic	cubic	linear	quadratic	cubic
Latin America						
age=24, ysm=4	0.10	–1.41**	–2.23***	0.12	–1.33**	–2.36**
	(0.37)	(0.64)	(0.78)	(0.37)	(0.63)	(0.77)
age=32, ysm=12		–0.76*	–0.39		–0.82**	–0.57
		(0.41)	(0.47)		(0.41)	(0.46)
age=40, ysm=20		–0.11	0.66		–0.31	0.44
		(0.41)	(0.59)		(0.41)	(0.58)
Europe						
age=24, ysm=4	–1.18	–0.96	1.80	–1.09	–1.16	2.54
	(0.86)	(1.74)	(2.49)	(0.84)	(1.77)	(2.63)
age=32, ysm=12		–0.85	–1.21		–0.95	–1.00
		(1.20)	(1.19)		(1.23)	(1.23)
age=40, ysm=20		–0.73	–2.64**		–0.74	–2.68**
		(0.86)	(1.29)		(0.87)	(1.23)
Asia						
age=24, ysm=4	–0.51	–0.84	–0.27	–0.36	–1.12	–0.51
	(0.64)	(1.37)	(1.84)	(0.62)	(1.30)	(1.72)
age=32, ysm=12		–0.52	–0.38		–0.60	–0.47
		(0.82)	(0.87)		(0.79)	(0.85)
age=40, ysm=20		–0.19	–0.29		–0.08	–0.19
		(0.76)	(1.05)		(0.75)	(1.04)

Standard errors are reported in parentheses. Confidence levels: 99% (***), 95% (**), 90% (*).

Sample sizes: Native (89117), Latin America (6438), Europe (1689), Asia (2657), Others (492).

Estimates represent immigrants' annual percentage wage growth relative to natives' percentage wage growth.

5 Concluding Remarks

This study reexamines the evidence of wage convergence of immigrants using a novel research design. The existing literature on immigrant wage convergence suffers from a lack of representative longitudinal data on the foreign-born population with sufficient sample size. We address the sample size problem by using the CPS MORG. The sample forms an overlapping rotating panel data set, which enables one to control for fixed unobserved heterogeneity and account for high sample attrition as well as outmigration.

We compare panel and cross-section models of economic assimilation by exploiting and ignoring, respectively, the two-year panel aspect of the CPS. The results suggest that controlling for individual fixed effects reverses the conventional result of economic assimilation. Overall, there is little evidence of economic assimilation for 1994-2004. New immigrants from Latin America earn lower wages than natives, and this gap widens with time spent in the U.S. labor market. Foreign-born workers from Europe and Asia earn higher wages than native-born workers, but there is no strong evidence of convergence. The results are robust to attrition.

Our cross-section and panel results are qualitatively different from the findings in earlier repeated cross-section and longitudinal studies. We find that the skill endowment or match-specific component of individual workers is positively correlated with age at migration, which is supported by human capital theory or the Roy model, respectively. This bias may not be large in earlier studies. However, whether cross-section results are biased is an empirical question. Therefore, one needs to be careful in applying cross-section approaches to samples for different years or different countries since we do not have prior knowledge about the direction of the bias.

6 References

Baker, Michael and Dwayne Benjamin (1997): "The Role of the Family in Immigrants' Labor-Market Activity: An Evaluation of Alternative Explanations," *American Economic Review*, 87(4), 705-727.

Bhattacharya, Debopam (2008): "Inference in Panel Data Models under Attrition Caused by Unobservables," *Journal of Econometrics*, 144(2), 430-446.

Bleakley, Hoyt and Aimee Chin (2004): "Language Skills and Earnings: Evidence from Childhood Immigrants," *Review of Economics and Statistics*, 86(2), 481-496.

Bollinger, Christopher R. and Barry T. Hirsch (2006): "Match Bias from Earnings Imputation in the Current Population Survey: The Case of Imperfect Matching," *Journal of Labor Economics*, 24(3), 483-519.

Borjas, George J. (1985): "Assimilation, Changes in Cohort Quality and the Earnings of Immigrants," *Journal of Labor Economics*, 3(4), 463-489.

Borjas, George J. (1989): "Immigrant and Emigrant Earnings: a Longitudinal Study," *Economics Inquiry*, 27(1), 21-37.

Borjas, George J. (1995): "Assimilation and Changes in Cohort Quality Revisited: What happened to Immigrant Earnings in the 1980s?" *Journal of Labor Economics*, 13(2), 201-245.

Borjas, George J. (1999): "The Economic Analysis of Immigration," in Ashenfelter, Orley C. and David Card, eds., *Handbook of Labor Economics*, Vol 3A, Ch28.

Borjas, George J. and Bernt Bratsberg (1996): "Who Leaves? The Outmigration of the Foreign-Born," *Review of Economics and Statistics*, 78(1), 165-176.

Bratsberg, Bernt, Erling Barth, and Oddbjorn Raaum (2006): "Local Unemployment and the Relative Wages of Immigrants: Evidence from the Current Population Surveys," *Review of Economics and Statistics*, 88(2), 243-263.

Butcher, Kristin F. and John DiNardo (2002): "The Immigrant and Native-Born Wage Distributions: Evidence from United States Censuses," *Industrial and Labor Relations Review*, 56(1), 97-121.

Card, David and John E. DiNardo (2002): "Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles," *Journal of Labor Economics*, 20(4), 733-783

Chen, Xiaohong, Han Hong, and Elie Tamer (2005): "Measurement Error Models with Auxiliary Data," *Review of Economic Studies*, 72(2), 343-366.

Chiswick, Barry R. (1978): "The Effect of Americanization on the Earnings of Foreign-Born Men," *Journal of Political Economy*, 86(5), 897-922.

Chiswick, Barry R. (1980): *An Analysis of the Economic Progress and Impact of Immigrants*, U.S. Department of Labor, National Technical Information Service, PB80-200454.

Dieleman, Frans M. (2001): "Modelling Residential Mobility; A Review of Recent Trends in Research," *Journal of Housing and the Built Environment*, 16(3), 249-265.

Duleep, Harriet O. and Mark C. Regets (1997a): "Measuring Immigrant Wage Growth using Matched CPS Files," *Demography*, 34(2), 239-249.

Duleep, Harriet O. and Mark C. Regets (1997b): "Immigrant Entry Earnings and Human Capital Growth: Evidence from 1960-1980 Censuses," in Soloman W. Polachek, ed., Vol 16, *Research in Labor Economics*, JAI Press, 297–317.

Gittleman, Maury and Mary Joyce (1996): "Earnings Mobility and Long-Run Inequality: An Analysis using Matched CPS Data," *Industrial Relations*, 35(2), 180-196.

Guell, Maia and Luojia Hu (2006): "Estimating the Probability of Leaving Unemployment using Uncompleted Spells from Repeated Cross-Section Data," *Journal of Econometrics*, 133(1), 307-341.

Hirsch, Barry T. and Edward J. Schumacher (2004): "Match Bias in Wage Gap Estimates Due to Earnings Imputation," *Journal of Labor Economics*, 22(3), 689-722.

Hirano, Keisuke, Guido W. Imbens, Geert Ridder, and Donald B. Rubin (2001): "Combining Panel Data Sets with Attrition and Refreshment Samples," *Econometrica*, 69(6), 1645-1659.

Hu, Wei-Yin (2000): "Immigrant Earnings Assimilation: Estimates from Longitudinal Data," *American Economic Review Papers and Proceedings*, 90, 368-372.

Jasso, Guillermina and Mark 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, Guillermina and Mark R. Rosenzweig (1988): “How Well do U.S. Immigrants do? Vintage Effects, Emigration Selectivity, and Occupational Mobility,” in T. Paul Schultz, ed., *Research in Population Economics*, Vol 6. 229-253.

Jasso, Guillermina and Mark R. Rosenzweig (1990): *The New Chosen People: Immigrants in the United States*, New York: Russell Sage Foundation.

Jasso, Guillermina, Mark R. Rosenzweig, and James P. Smith (1998): “The Changing Skills of New Immigrants to the United States: Recent Trends and Their Determinants,” Working Paper 6764, National Bureau of Economic Research.

Kim, Seik (2012): “Sample Attrition in the presence of Population Attrition,” Working Paper, University of Washington.

LaLonde, Robert J. and Robert H. Topel (1992): “The Assimilation of Immigrants in the U.S. Labor Market,” in George J. Borjas and Richard B. Freeman, eds., *Immigration and the Work Force*, 67-92.

LaLonde, Robert J. and Robert H. Topel (1997): “Economic Impact of International Migration and The Economic Performance of Migrants,” in Mark R. Rosenzweig and Oded Stark, eds., *Handbook of Population and Family Economics*, Vol 1B, Ch 14.

Lemieux, Thomas (2006): “Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?” *American Economic Review*, 96(3), 461-498.

Little, Roderick J. A., and Donald B. Rubin (2002): *Statistical Analysis with Missing Data*, Wiley Series in Probability and Statistics.

Long, Larry (1988): *Migration and Residential Mobility in the United States*, Russell Sage Foundation.

Lubotsky, Darren H. (2007): “Chutes or Ladders? A Longitudinal Analysis of Immigrant Earnings,” *Journal of Political Economy*, 115(5), 820-867.

Neumark, David and William Wascher (1996): “The Effects of Minimum Wages on Teenage Employment and Enrollment: Evidence from Matched CPS Surveys,” in S.W. Polachek, ed., Vol 15, *Research in Labor Economics*, JAI Press, 25–63.

Peracchi, Franco and Finis Welch (1995): "How Representative are Matched Cross-Sections? Evidence from the Current Population Survey," *Journal of Econometrics*, 68(1), 153-179.

Ridder, Geert, and Robert Moffitt (2003): "The Econometrics of Data Combination," in James J. Heckman and Edward E. Leamer, eds., *Handbook of Econometrics*, Vol 6B, Ch 75, 5469-5547.

Schmitt, John (2003): "Creating a consistent hourly wage series from the Current Population Survey's Outgoing Rotation Group, 1979-2002," Working Paper.

Schoeni, Robert F. (1997): "New Evidence on the Economic Progress of Foreign-Born Men in the 1970s and 1980s," *Journal of Human Resources*, 32(4), 683-740.

U.S. Census Bureau (2002): *Technical Paper 63RV: Current Population Survey - Design and Methodology.*

Warren, Robert and Jennifer M. Peck (1980): "Foreign-Born Emigration from the United States: 1960 to 1970," *Demography*, 17(1), 71-84.

Wooldridge, Jeffrey M. (2002): "Inverse Probability Weighted M-Estimators for Sample Selection, Attrition, and Stratification," *Portuguese Economic Journal*, 1, 117-139.

7 Appendix

7.1 Variables used in the Analysis

This section explains in detail how the CPS MORG is processed to generate the sample used in the analysis. The wage measure used in the analysis is the hourly rate of pay. The wage measure is the hourly wage for the hourly workers and the weekly payments divided by the usual weekly hours of work for non-hourly workers. We clean the wage measure by following steps which are similar to those in Lemieux (2006). Both the hourly and the weekly wages are topcoded. For workers paid by the hour, the topcode remains between $99.00-99.99 and only a small fraction of workers have their wage censored at this value. On the other hand, a substantial number of non-hourly workers have topcoded wages. The weekly wage is topcoded by $1923 in 1994-1997 and by $2884 in 1998-2004. Topcoded wages are adjusted by a factor of 1.4.[20] Workers with extreme wages (less

[20] The simplest way of handling topcoded values is to adjust censored values by a factor that approximates the mean for those above the censoring point (typically, a factor like 1.33 or 1.4). According to Schmitt (2003), a

than \$2 and more than \$200 in 1994 dollars) are trimmed. In addition, the sample drops persons with negative potential experience. As a result, 998 out of 35,016 foreign-born and 11,791 out of 266,628 native-born persons are dropped. These trimmed samples are used throughout the paper unless otherwise indicated.

The year of arrival information provided by the CPS MORG lets us identify those who arrived in the United States before 1950, 1950-1959, 1960-1964, 1965-1969, 1970-1974, 1975-1979, 1980-1981, 1982-1983, and so on. The most recent entrants, however, are coded in an inconsistent way. For instance, the arrival year code 13 in the 1994 sample includes the 1992-1994 arrivals, the code 13 in the 1995 sample includes the 1992-1995 arrivals, and the code 13 in the 1996 sample and afterwards include the 1992-1993 arrivals. Therefore foreign-born persons who arrived in the United States in 1992-1993 and are in the 1994-1995 or the 1995-1996 panels cannot be matched. As a consequence, we drop immigrants with the arrival year code 13 in the 1994-1995 or the 1995-1996 panels. So, the most recent immigrants in the 1994-1995 and the 1995-1996 panels are those who entered the U.S. in 1990-1991 with the arrival year code 12. Accordingly in the panels of the subsequent years, we keep immigrants with the arrival year code numbers of the followings:

1994-1995 and 1995-1996 panels: codes 1-12 (1990-1991)

1996-1997 and 1997-1998 panels: codes 1-13 (1992-1993)

1998-1999 and 1999-2000 panels: codes 1-14 (1994-1995)

2000-2001 and 2001-2002 panels: codes 1-15 (1996-1997)

2002-2003 and 2003-2004 panels: codes 1-16 (1998-1999)

where the years in the parentheses indicate the entry years of the most recent immigrants.

Some variables in the CPS MORG are given by intervals. One example is the arrival year. It is given by periods rather than years. In the analysis, the arrival year variable is defined by the mid-point of each period. Immigrants who arrived in the United States before 1950 are coded as 1940. The education measure needs adjustment, too. The values for the education measure are assigned by the following rule:

0 if less than 1st grade

more sophisticated way is estimating the mean above the topcode using the pareto distribution. As the pareto distribution has two parameters, what is mostly done is to fit the pareto distribution through a point high in the observed distribution.

2.5 if 1st-4th grade

5.5 if 5th-6th

7.5 if 7th-8th

10 if 9th, 10th, 11th, or 12th grades with no diploma

12 if high school graduate including GED

14 if some college but no degree or Associate degree

16 if Bachelor's degree

18 if Master's degree, Professional school degree, or Doctorate degree

The estimation results are not very sensitive to the ways of coding year of entry and education.

7.2 Sample Attrition in the Presence of Unobserved Population Attrition: Not for Publication

Denote $D_S = 1$ when an individual is in the sample (or responds) in the second year and $D_S = 0$ when an individual is not in the sample (or does not respond) in the second year. Denote $D_P = 1$ when an individual is in the population (or stays in the United States) in the second period and $D_P = 0$ when an individual is not in the population (or leaves the United States) in the second period. It is possible to construct a balanced longitudinal sample by collecting all the individuals with $D_P = 1$ and $D_S = 1$. This sample is called the matched sample.[21]

Suppose that there is no population attrition. Assume that sample attrition is a function of u_1, u_2, and v, where u_1 and u_2 are vectors of time-varying variables in periods 1 and 2, respectively, and v is a vector of time invariant variables. For instance, u_1 (or u_2) is a vector of the endogenous variable and time-varying exogenous variables and is v is a vector of time-invariant exogenous variables. u_2 is observed because the second period cross-section is available. Specify one minus the sample attrition function by

$$\Pr\left(D_S = 1 | U_1 = u_1, U_2 = u_2, V = v\right) = g\left(v'\phi_0 + u_1'\phi_1 + u_2'\phi_2\right), \tag{12}$$

[21] Similarly, an individual stays in the U.S. but does not respond in the second period if $D_P = 1$ and $D_S = 0$. An individual who leaves the U.S. in the second period is denoted by $D_P = 0$. A combination of $D_P = 0$ and $D_S = 1$, where an individual leaves the country and responds in the second period, is not possible. As a result, being in the matched sample, $D_S = 1$, also implies residing in the U.S. at the same time, $D_P \cdot D_S = 1$.

where v is a vector of a constant, age, education, and dummy variables (marital status, years in the United States, citizenship status, country of birth), u_1 and u_2 are vectors of logged hourly real dollar wages and indicators of "not usually working", and $g(r) = e^r / (1 + e^r)$. Since the $g(\cdot)$ function and $\Pr(D_S = 1)$ are estimable, one can construct the attrition correcting weights by

$$C(u_1, u_2, v) = \frac{\Pr(D_S = 1)}{g(v'\phi_0 + u_1'\phi_1 + u_2'\phi_2)}. \tag{13}$$

Intuitively, this step is equivalent to weighting the individuals in the matched sample with the inverse of one minus the probability of sample attrition, $1/g(v'\phi_0 + u_1'\phi_1 + u_2'\phi_2)$.

In the presence of population attrition, one additional step is required prior to the above procedure. The population attrition function, $\Pr(D_P = 1|u_2, v)$, can be nonparametrically identified when population attrition is solely determined by variables of known transition probability. This is a strong assumption, but necessary because we do not know who emigrated from the United States. Suppose that the transition probability is given by $P(Z_2 = z_2|Z_1 = z_1)$, where z is a vector of variables of known transition probability.[22] For instance, if z is year of entry, the transition probability is given by $P(z_2|z_1) = 1(z_2 = z_1)$, where $1(\cdot)$ is the indicator function. If z is age, the transition probability is given by $P(z_2|z_1) = 1(z_2 = z_1 + 1)$. Specify one minus the population attrition function by

$$\begin{aligned} \Pr(D_P = 1|u_2, v) &= \Pr(D_P = 1|z_2) \\ &\equiv k(z_2'\psi), \end{aligned} \tag{14}$$

where $k(r) = e^r$, and z_2 is a vector of age, years since migration, education (assuming that no additional schooling is obtained), country of origin, and year of entry.[23] Intuitively, weight the individuals in the population (or more precisely the cross-section) with the inverse of one minus

[22] The variables in z_2 must be included in (u_2, v).

[23] These variables have deterministic time paths and satisfy the known transition probability assumption. The assumption, however, is more restrictive than the sample selection model, for instance, because observable variables with unknown transition probability, such as the wage, cannot enter in the selection function. The assumption can be problematic as the transition probabilities of labor market performance variables are usually not known. Intuitively labor market performance will affect population attrition decision. If the assumption is indeed a serious problem in practice, it is required to develop an alternative way of handling population attrition.

the probability of population attrition, $1/k(z_2'\psi)$.

The weights in (13) can be estimated by the conditional moment restrictions given by

$$\begin{aligned} 1 &= E\left[\frac{D_S}{g(v'\phi_0 + u_1'\phi_1 + u_2'\phi_2)} | u_1, v\right] \quad \text{w.p.1}, \\ \frac{1}{k(z_2'\psi)} &= E\left[\frac{D_S}{g(v'\phi_0 + u_1'\phi_1 + u_2'\phi_2)} | u_2, v, D_P = 1\right] \quad \text{w.p.1}. \end{aligned} \tag{15}$$

In the first step, estimate $1/k(z_2)$, which is equivalent to weighting the individuals in the second year cross-section with the inverse of one minus the probability of population attrition. In the second step, estimate (15) and obtain (13). Finally, use (13) to weight individuals in the matched sample and estimate the main model of interest. Since the weights are assigned to individuals, the attrition correcting method is robust to individual fixed effects.

Table A1-1. Economic Assimilation Estimates in % (by Origin): Reported & Imputed Wages

Individual Heterogeneity	Attrition-Adjusted			Not Adjusted		
	linear	quadratic	cubic	linear	quadratic	cubic
Latin America						
age=24, ysm=4	–0.01	–1.52**	–2.09**	0.17	–1.25*	–1.98**
	(0.41)	(0.70)	(0.85)	(0.41)	(0.71)	(0.86)
age=32, ysm=12		–0.91**	–0.66		–0.78	–0.63
		(0.45)	(0.52)		(0.45)	(0.52)
age=40, ysm=20		–0.30	0.24		–0.31	0.19
		(0.46)	(0.65)		(0.46)	(0.65)
Europe						
age=24, ysm=4	–1.69	–3.17*	–1.94	–1.39	–3.14*	–2.21
	(0.90)	(1.76)	(2.30)	(0.90)	(1.79)	(2.38)
age=32, ysm=12		–2.24*	–2.29*		–2.19*	–2.09*
		(1.22)	(1.26)		(1.24)	(1.26)
age=40, ysm=20		–1.32	–2.13		–1.25	–1.68
		(0.90)	(1.47)		(0.91)	(1.47)
Asia						
age=24, ysm=4	0.55	–0.01	0.96	0.91	0.79	2.47*
	(0.69)	(1.40)	(1.81)	(0.69)	(1.38)	(1.74)
age=32, ysm=12		0.36	0.07		0.86	0.41
		(0.85)	(0.92)		(0.84)	(0.91)
age=40, ysm=20		0.72	0.08		0.92	–0.23
		(0.80)	(1.14)		(0.80)	(1.14)
Others						
age=24, ysm=4	0.93	–1.15	–0.42	1.80	–0.29	0.87
	(1.58)	(2.84)	(3.44)	(1.54)	(2.83)	(3.54)
age=32, ysm=12		–0.02	–0.43		0.74	0.26
		(1.83)	(2.01)		(1.81)	(1.95)
age=40, ysm=20		1.11	0.27		1.77	0.59
		(1.70)	(2.51)		(1.69)	(2.46)

Standard errors are reported in parentheses. Confidence levels: 99% (***), 95% (**), 90% (*).

Sample sizes: Native (156241), Latin America (11560), Europe (3392), Asia (5340), Others (1162).

Estimates represent immigrants' annual percentage wage growth relative to the natives' percentage wage growth.

Table A1-2. Economic Assimilation Estimates in % (by Origin): Weighted Reported Wages

Individual Heterogeneity	Attrition-Adjusted			Not Adjusted			Att.-Adjusted, enter ≥18		
	linear	quadra.	cubic	linear	quadra.	cubic	linear	quadra.	cubic
Latin America									
age=24, ysm=4	0.16	–1.29**	–2.13***	0.17	–1.21*	–2.25**	–0.71	–2.77***	–3.80***
	(0.37)	(0.64)	(0.77)	(0.37)	(0.63)	(0.77)	(0.50)	(0.99)	(1.48)
age=32, ysm=12		–0.71*	–0.36		–0.76**	–0.54		–2.05***	–1.43*
		(0.41)	(0.47)		(0.41)	(0.46)		(0.70)	(0.84)
age=40, ysm=20		–0.12	0.64		–0.32	0.41		–1.32*	–0.36
		(0.41)	(0.58)		(0.41)	(0.57)		(0.73)	(0.92)
Europe									
age=24, ysm=4	–1.28	–1.25	1.76	–1.15	–1.50	2.47	–1.39	0.41	7.98**
	(0.85)	(1.71)	(2.49)	(0.82)	(1.74)	(2.62)	(1.18)	(2.45)	(3.69)
age=32, ysm=12		–1.04	–1.34		–1.18	–1.13		–0.94	–2.63
		(1.19)	(1.19)		(1.22)	(1.22)		(1.91)	(2.23)
age=40, ysm=20		–0.83	–2.80**		–0.85	–2.80**		–2.29	–7.53**
		(0.85)	(1.29)		(0.86)	(1.22)		(1.72)	(2.61)
Asia									
age=24, ysm=4	–0.49	–0.75	–0.16	–0.33	–1.04	0.38	–1.53	–3.07*	–2.51
	(0.63)	(1.36)	(1.83)	(0.62)	(1.30)	(1.71)	(0.73)	(1.68)	(2.65)
age=32, ysm=12		–0.49	–0.37		–0.57	–0.46		–2.00*	–2.10
		(0.81)	(0.86)		(0.78)	(0.85)		(1.20)	(1.50)
age=40, ysm=20		–0.23	–0.35		–0.10	–0.26		–0.93	–0.99
		(0.77)	(1.04)		(0.76)	(1.03)		(1.16)	(1.39)
Others									
age=24, ysm=4	–0.82	–0.43	–3.40	–0.15	–0.04	–1.88	–0.63	2.78	–5.48
	(1.81)	(3.30)	(4.16)	(1.66)	(2.93)	(3.80)	(2.05)	(4.83)	(6.28)
age=32, ysm=12		0.10	1.47		0.44	1.06		4.25	0.17
		(2.08)	(2.34)		(1.85)	(2.00)		(3.77)	(3.33)
age=40, ysm=20		0.64	3.25		0.93	2.39		5.72	6.39
		(1.91)	(2.80)		(1.85)	(2.62)		(3.78)	(4.72)

Standard errors are reported in parentheses. Confidence levels: 99% (***), 95% (**), 90% (*).

Sample sizes: Native (89117), Latin America (6438), Europe (1689), Asia (2657), Others (492).

Sample sizes of the last column: Native (89117), Latin America (3530), Europe (979), Asia (1922), Others (355).

Table A2-1. Wage Equation (in First Differenced) Estimates using Non-Imputed Wages

Individual Heterogeneity	Attrition-Adjusted			Not Adjusted		
	linear	quadratic	cubic	linear	quadratic	cubic
Constant	0.024***	0.096***	0.178***	0.034***	0.108***	0.194***
	(0.005)	(0.006)	(0.014)	(0.005)	(0.006)	(0.014)
$\frac{1}{10}$Age		–0.019***	–0.064***		–0.018***	–0.064***
		(0.001)	(0.007)		(0.001)	(0.007)
$\frac{1}{100}$Age^2			0.006***			0.006***
			(0.001)			(0.001)
Imm.	–0.002	–0.019*	–0.036	–0.002	–0.021*	–0.028
	(0.003)	(0.012)	(0.039)	(0.003)	(0.012)	(0.038)
$\frac{1}{10}$Age$_i$		0.003	0.008		0.003	0.005
		(0.003)	(0.021)		(0.003)	(0.020)
$\frac{1}{100}$Age$_i^2$			–0.001			–0.000
			(0.003)			(0.002)
$\frac{1}{10}$YSM		0.002	0.010		0.001	0.008
		(0.003)	(0.009)		(0.003)	(0.009)
$\frac{1}{100}$YSM^2			–0.002			–0.001
			(0.002)			(0.002)

Standard errors are reported in parentheses. Observations = 100393

Imm.: indicator of a foreign-born person; Age$_i$: age × Imm.

Fixed Effects: calendar year

Table A2-2. Wage Equation (in Level) Estimates using Non-Imputed Wages

Cohort Heterogeneity	Attrition-Adjusted			Not Adjusted		
	linear	quadratic	cubic	linear	quadratic	cubic
Constant	0.641***	–0.493***	–1.087***	0.731***	–0.468***	–1.134***
	(0.013)	(0.020)	(0.055)	(0.014)	(0.021)	(0.059)
Age	0.014***	0.081***	0.132***	0.012***	0.078***	0.134***
	(0.000)	(0.001)	(0.005)	(0.000)	(0.001)	(0.005)
$\frac{1}{100}\text{Age}^2$		–0.085***	–0.219***		–0.083***	–0.225***
		(0.001)	(0.012)		(0.001)	(0.013)
$\frac{1}{1000}\text{Age}^3$			0.011***			0.012***
			(0.001)			(0.001)
Education	0.100***	0.094***	0.093***	0.101***	0.096***	0.096***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Imm.	0.501***	0.589***	0.281	0.457***	0.580***	0.337*
	(0.032)	(0.063)	(0.172)	(0.033)	(0.065)	(0.182)
Age_i	–0.009***	–0.019***	0.010	–0.008***	–0.020***	0.003
	(0.001)	(0.003)	(0.014)	(0.001)	(0.003)	(0.015)
$\frac{1}{100}\text{Age}_i^2$		0.013***	–0.064*		0.014***	–0.047
		(0.004)	(0.037)		(0.004)	(0.039)
$\frac{1}{1000}\text{Age}_i^3$			0.006**			0.005
			(0.003)			(0.003)
YSM	0.019***	0.024***	0.015**	0.018***	0.026***	0.017**
	(0.002)	(0.004)	(0.007)	(0.002)	(0.004)	(0.008)
$\frac{1}{100}\text{YSM}^2$		–0.024**	0.027		–0.028***	0.020
		(0.010)	(0.035)		(0.011)	(0.037)
$\frac{1}{1000}\text{YSM}^3$			–0.007			–0.006
			(0.005)			(0.005)
Education_i	–0.043***	–0.038***	–0.038***	–0.041***	–0.037***	–0.037***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)

Standard errors are reported in parentheses. Observations = 100393

Imm.: indicator of a foreign-born person; Age_i: age × Imm.; Educ_i: years of schooling × Imm.

Fixed Effects: birth country, arrival year, calendar year