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/

November 24, 2009

Abstract

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 observationally equivalent native-born workers with additional time spent in the United States. We compare cross-section and panel models of foreign-native gap in wage growth using the Current Population Survey (CPS) for 1994-2004. The results suggest that analyses based on repeated cross-section studies are biased upward by fixed unobserved heterogeneity and controlling for this heterogeneity reverses the conventional result of economic assimilation. Overall, we find little evidence of a narrowing of the foreign-native gap in economic performance. New immigrants from Central and South America earn lower wages than natives, and this gap widens with time in the U.S. labor market. The wages of new immigrants from Europe and Asia exceed those of natives and there is no strong evidence of convergence. We account for sample attrition in the presence of nonrandom outmigration and find that our results are robust to attrition.

Keywords: Economic Assimilation, Immigration, Outmigration, Overlapping Rotating Panel Data, Sample Attrition, Population Attrition

---

[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, Peter Phillips, Paul Schultz, Emiko Usui, two anonymous referees, and seminar participants at the summer meetings of the Econometric Society, the annual meetings of the Society of Labor Economists, the UW Center for Studies in Demography and Ecology (CSDE), 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.

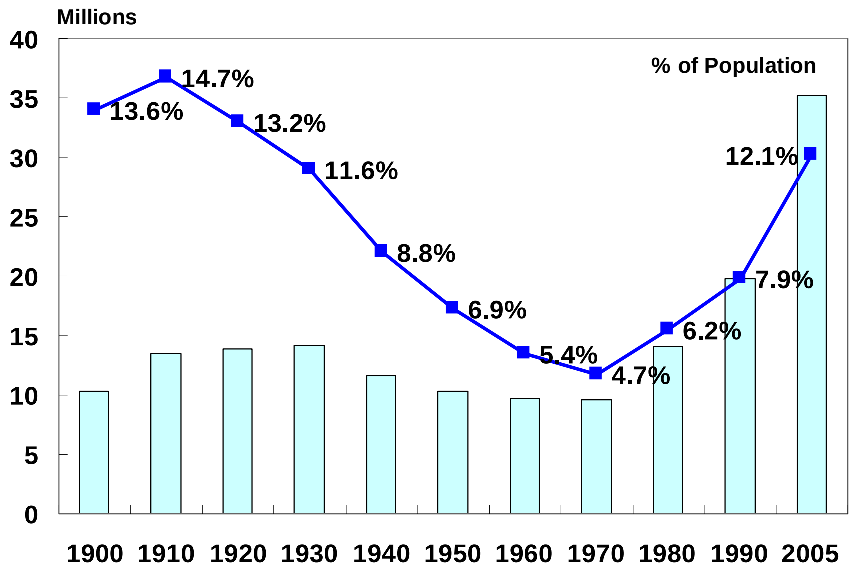

Figure 1: The Number and Share of Foreign-Born Population in the United States (1900-2005)

1 Introduction

The large and growing share of foreign-born workers in the United States has heightened interest in the economic impact of immigration. (See Figure 1.[1]) How immigrants fare as they accumulate experience in the U.S. labor market is the key to many of these effects.[2] 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 U.S. 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.[3]

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

[1]Figure 1 shows the number of foreign-born persons living in the United States over the last 100 years and their share of the total population. Since 1970 the foreign-born population has more than tripled. At the beginning of 2005, 35.2 million foreign-born persons were residing in the United States, making up 12.1 percent of the total U.S. population.

Source: http://www.cis.org/articles/2005/back1405.html, Center for Immigration Studies (CIS), based on the Decennial Censuses for 1900-1990 and CIS Analysis of March 2005 CPS.

[2]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.

[3]See Borjas (1999) and LaLonde and Topel (1997) for discussions of the effect of immigrants on the labor market outcomes of natives.

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 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 year of entry.[4] 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.

In principle, longitudinal data on native-born and foreign-born populations, by tracking specific individuals over time, offers the huge advantage of permitting one to control for fixed unobserved heterogeneity. 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, the use of panel data gives rise to an additional problem: nonrandom sample attrition. For example, if persons with negative (positive) wage shocks are more likely to drop out of the sample, panel estimates will overstate (understate) the growth of wages. In addition, outmigration of immigrants that is related to wage growth poses another attrition problem for panel data analyses as well as for cross-section analyses.[5] For example, if less skilled and/or unlucky immigrants tend to return to their home country, stayers will on average earn higher wages than returning migrants. It is a complicated problem since the data does not reveal who emigrated from the United States.

This paper provides a longitudinal analysis of assimilation by exploiting the two-year panel aspect of

[4] 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.

[5] One needs to account for return migrants when the labor market performance of immigrants is of interest. One does not need to, however, when the impact of immigrant workers on the U.S. economy is of interest. To meet the interest of readers, this paper reports estimates that account for sample attrition and outmigration and that account for sample attrition only. We find that the two sets of estimates are similar.

the Current Population Survey (CPS) to account for individual fixed effects. Individuals are matched between adjacent years in the Merged Outgoing Rotation Groups (MORG) of the CPS to form multiple two-year panel data sets. The estimates from short panels are consistent. The matched CPS has the crucial advantage of being much larger than alternative panel data sets such as the PSID and the NLSY79. The sample also allows one to account for sample attrition in the presence of unobserved outmigration. We confirm that accounting for attrition does not alter our findings qualitatively.

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.[6] New immigrant workers from Central and South 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.

To understand why our results differ from those of the previous literature, this paper compares cross-section and panel models of immigrant and native wage growth rates using the same CPS data. An advantage of using the CPS is that one can construct cross-section samples by ignoring its longitudinal structure. First, we replace the panel model with the repeated cross-section model and find that the estimated measures of economic assimilation flip signs. This implies that the repeated cross-section approach yields misleading conclusion if applied to recent data. Second, to see whether repeated cross-section estimates based on earlier years are biased, we borrow the results from Hu (2000) and Lubotsky (2007). They control for individual heterogeneity by using cross-section linked to Social Security earnings data and find modest positive assimilation for earlier years. The potential bias in earlier repeated cross-section estimates is minimal. In sum, these results suggest that the immigrant earnings process seems to be fundamentally different now than it was in the past and one should not use the repeated cross-section approach to estimate recent assimilation.

The comparison between cross-section and panel estimates yields another interesting finding that relates to heterogeneity of immigrant skills at the time of entry into the United States. Immigrant wage growth estimates based on repeated cross-section studies are biased upward by individual heterogeneity, and this is because older migrant workers are more skilled than younger ones conditional on year of entry and other observable variables.[7] Whether there is bias due to unobserved heterogeneity is an empirical

[6] Changes in the composition of natives who are in the labor market can also lead to different estimates of the age profile for natives relative to what one would get in a repeated cross-section. Failure to account for such changes would bias repeated cross-section estimates. There have been changes in labor force participation rates by age, race, and education in the U.S. population. Note that when estimating sample attrition weights, we include labor force participation status. In the CPS MORG, we find that attrition is negatively correlated with wage growth.

[7] Our findings do not contradict with the findings by Bleakley and Chin (2004). They show that among the immigrants

question and the answer may vary depending on which time periods or which countries are considered. However, one needs to be careful in applying the cross-section approach since it is uncertain ex ante whether fixed unobserved heterogeneity causes a problem even after controlling for the year of entry and the country of origin.

The paper proceeds as follows. Section 2 defines economic assimilation, outlines the models for estimating economic assimilation, and introduces some of the key econometric issues. Section 3 introduces the data set. The summary statistics suggest that sample attrition and possibly outmigration must be taken into account in estimation of economic assimilation. They also provide initial evidence that different ethnic groups may experience assimilation differently. Section 4 presents the main results and Section 5 offers conclusions.

2 Issues in Measuring Economic Assimilation

This section defines a measure of economic assimilation and presents empirical specifications and identification issues. It discusses advantages and disadvantages of the available data sets and outlines the idea of accounting for sample attrition when emigrants are unidentified.

2.1 Definition of Economic Assimilation

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

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

and that of a native worker by

$$y_{it} = h_{nat}\left(age_{it}, edu_i, \mu_i, t\right), \tag{2}$$

for some 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, and t reflects market conditions and economic shocks. 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, age at migration, and country of origin.

who immigrated to the U.S. as children, those who entered as younger children assimilate faster than those who entered as older children due to language proficiency. This paper shows that among the new adult immigrants, those who are older do better than those who are younger.

The economic performance of a foreign-born worker relative to a native-born worker at time t can be measured by

$$EA\left(age, ysm; t\right) = \frac{d}{dt} h_{imm}\Big|_{(age,ysm,t)} - \frac{d}{dt} h_{nat}\Big|_{(age,t)}. \tag{3}$$

Roughly speaking, $EA\left(age, ysm; t\right)$ 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 foreign-born workers initially earn lower wages than average native-born workers.[8] In this case, wage convergence from below toward the higher native mean, $EA\left(age, ysm; t\right) > 0$, means economic assimilation. This paper also considers the case where foreign-born workers initially earn higher wages than average 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, $EA\left(age, ysm; t\right) < 0$.

Figure 2 illustrates an idea of how economic assimilation can be measured. The sample is drawn from the CPS. The figure depicts the mean hourly wages of foreign-born and native-born male workers of various age groups during 1994-2004.[9] 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. 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.

We observe that 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 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.

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

[9] 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 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 white sample is a solid reference group in that racial/ethnic composition of the native population has changed dramatically in recent years.

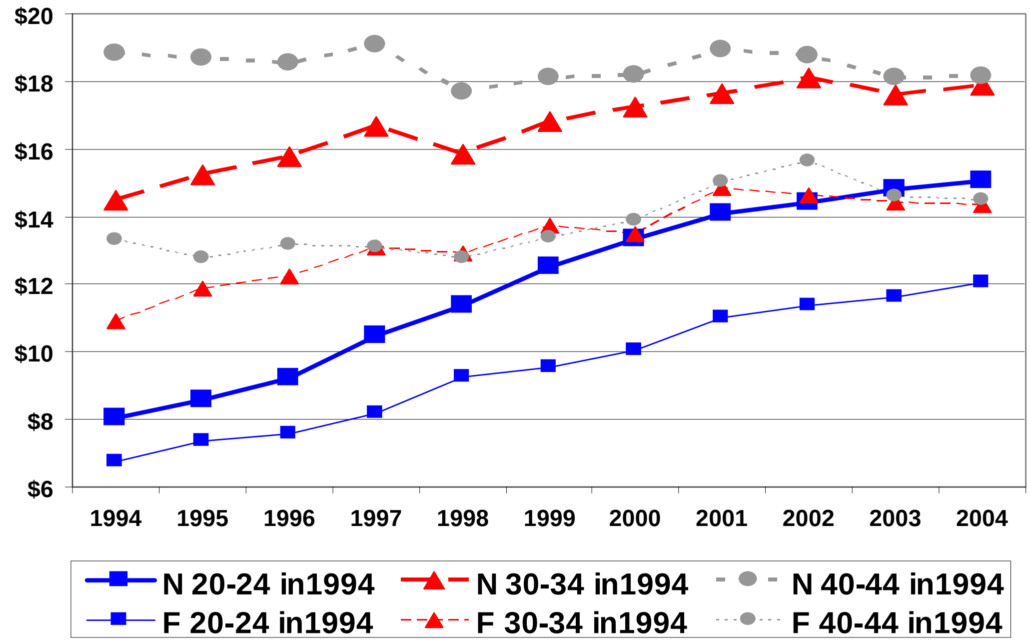

Figure 2: Average Wages (in 1994 Dollars) of Native-Born and Foreign-Born Workers

The foreign-born workers in the "40-44 in 1994" cohort experience economic assimilation over the 1994-2004 period as the wage gap narrows. Later in this paper, we show that panel estimates are consistent with the pattern in Figure 2, but repeated cross-section estimates are not.

2.2 Empirical Specification and Identification

Based on the model given in (1) and (2), 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

$$y_{it}^{imm} = (\alpha_{nat} + \alpha)\, age_{it} + \delta ysm_{it} + (\beta_{nat} + \beta)\, edu_i + \mu_i + \gamma_{imm,t} + \varepsilon_{it}, \quad (4)$$

$$y_{it}^{nat} = \alpha_{nat} age_{it} + \beta_{nat} edu_i + \mu_i + \gamma_{nat,t} + \varepsilon_{it}, \quad (5)$$

where μ_i is involves ability or skill endowment, and γ_t reflects business cycles, and ε captures idiosyncratic shocks.[10] The IH model in (4) and (5) allows fixed unobserved heterogeneity such as variation in skill endowments within the groups of individuals who entered in the same year. The empirical findings of this paper suggest positive correlation between ability and age at migration. Estimation of the IH model

[10] A more general model is estimated in this analysis allowing for nonlinearities in the age and the number of years since migration. These generalizations do not affect the discussion of identification issues.

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

$$\begin{aligned} y_{it}^{imm} &= (\alpha_{nat} + \alpha)\, age_{it} + \delta ysm_{it} + (\beta_{nat} + \beta)\, edu_i + \mu_c + \lambda_a am_i + bc_i' \lambda_b + \gamma_{imm,t} + \varepsilon_{it}, && (6) \\ y_{it}^{nat} &= \alpha_{nat} age_{it} + \beta_{nat} edu_i + \gamma_{nat,t} + \varepsilon_{it}, && (7) \end{aligned}$$

where μ_c is arrival year cohort effects, am is the age at migration, bc is a vector of birth country indicators. 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 (6) and (7) is the CH model. Estimation of the CH model requires repeated cross-sections.

The empirical measure of economic assimilation in (3) for both the IH and the CH models is given by

$$\begin{aligned} EA\,(age, ysm; t) &\approx \left(\alpha_{nat} + \alpha + \delta + \gamma_{imm,t+\triangle t} - \gamma_{imm,t}\right) - \left(\alpha_{nat} + \gamma_{nat,t+\triangle t} - \gamma_{nat,t}\right) \\ &= \alpha + \delta + \left(\gamma_{imm,t+\triangle t} - \gamma_{imm,t}\right) - \left(\gamma_{nat,t+\triangle t} - \gamma_{nat,t}\right). && (8) \end{aligned}$$

We address three issues regarding the identification of (8). First, we assume that the vector of the coefficients for the calendar year dummy variables common to foreign-born and native-born workers: $\gamma_{imm,t} = \gamma_{nat,t} = \gamma_t$ for all t.[11] With the restriction, the measure of economic assimilation in (8) is given by

$$EA\,(age, ysm; t) = \alpha + \delta.$$

It is crucial for identification as the number of years since migration, the arrival year, and the calendar year are collinear. So we need some restrictions for identification, although different restrictions lead to different estimates of the underlying parameters of interest. Under our identification restrictions, aggregate economic shocks affect the wages by the same percentage amount to foreign-born and native-born workers.[12]

Second, the IH model is estimated by taking the first difference, but α and δ are not separately

[11] This restriction is proposed by Borjas (1985). While LaLonde and Topel (1997) argue that this assumption is too restrictive, Borjas (1999) discusses that “[it] is less restrictive than it seems.”

[12] As Baker and Benjamin (1997) point out, the shocks will not be common, if immigrant and native workers differ in their sensitivity to the business cycle.

identified:

$$
\begin{aligned}
\Delta y_{it}^{imm} &= \alpha_{nat} + \alpha + \delta + \Delta\gamma_t + \Delta\varepsilon_{it}, \\
\Delta y_{it}^{nat} &= \alpha_{nat} + \Delta\gamma_t + \Delta\varepsilon_{it}.
\end{aligned}
$$

Therefore, $\alpha + \delta$ is identified by $(\alpha_{nat} + \alpha + \delta + \Delta\gamma_t) - (\alpha_{nat} + \Delta\gamma_t)$. Notice that $\alpha_{nat} + \alpha + \delta$ and α_{nat} are not identified due to $\Delta\gamma_t$. The assumption of a common γ plays a key role in identification of $\alpha + \delta$.

The third issue is the age at migration, am, in the CH model. The coefficients for age, years since migration, and age at migration are not separately identified: the three variables are collinear, $am = age - ysm$. To identify these coefficients, Borjas (1995), for instance, restricts the age coefficient for immigrants and natives to be common. However, we claim that restricting age coefficients does more harm than good if the parameter of interest is $\alpha + \delta$. Consider a CH model without the age at migration. Now we have the omitted variable bias: the probability limits of the coefficients for the age and the years since migration are $\alpha + \lambda_a$ and $\delta - \lambda_a$, respectively. In consequence, α and δ are not separately identified. It is worth noticing, however, that even if there is an omitted variable, the sum of the coefficients, $\alpha + \delta$, is identified because $\alpha + \delta = (\alpha + \lambda_a) + (\delta - \lambda_a)$.[13]

2.3 Advantages and Disadvantages of the Available 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-born and foreign-born populations permit one to track specific individuals over time and thus control for fixed unobserved heterogeneity. In practice, however, longitudinal analysis gives rise to an additional problem: sample attrition. 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.[14] Although panel analyses are robust to a link between outmigration and individual fixed effects, if the missing individuals are nonrandomly selected, longitudinal analyses must be approached cautiously.

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

[13] We may interact the cohort fixed effects with age at migration, $\mu_{c,am}$, but perfect correlation is still a problem. The point is that there is no need of further restrictions such as common age coefficients for immigrants and natives. Later we show that $\mu_{c,am}$ eliminates the bias in estimation of economic assimilation.

[14] If one is interested in the impact of immigrants on the U.S. economy rather than immigrants labor market performance, the Social Security records would be an ideal data set.

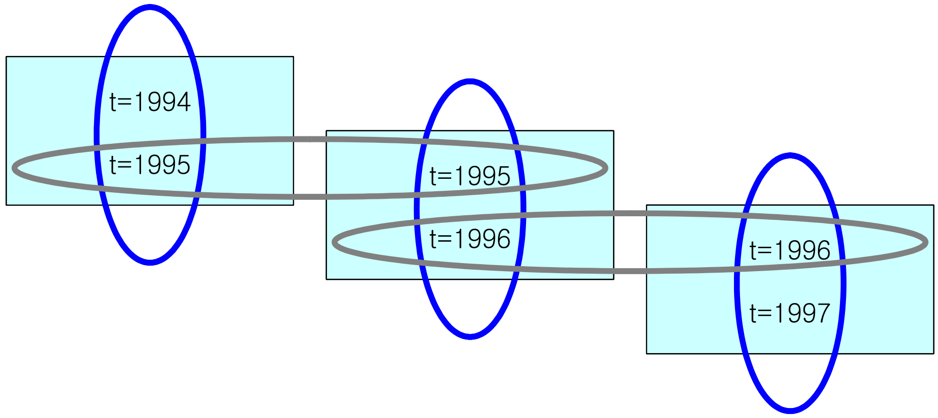

Figure 3: Data Structure of an Overlapping Rotation Panel Data Set

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. More recently, Hu (2000) uses the Health and Retirement Study (HRS) linked to the Social Security Earnings data and Lubotsky (2007) uses the Current Population Survey (CPS) and the Survey of Income and Program Participation (SIPP) linked to the Social Security Earnings data. They collect a sample of individuals with known social security numbers from cross-sections and connect time series of their past social security earnings. These linked data, different from the original CPS which includes illegal aliens, will underrepresent immigrants working in the uncovered sector or the underground economy. In addition, the data fail to include immigrants who left the United States before the cross-section interview period.[15]

Given that an ideal sample is not available, it is desirable to have a 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 MORG. The data structure of an overlapping rotating panel is depicted in Figure 3, where the short panels are represented by the blocks. Vertical circles symbolize the longitudinal feature of an overlapping rotating panel. Horizontal circles illustrate the overlapping feature of the short panels. As the sampling periods of two adjacent short panel data sets overlap, short panels can mimic a longitudinal sample if combined properly.

An overlapping rotating panel data set shares most of the advantages of usual panel data sets and is superior in some dimensions. First, the sample has a longitudinal feature. This means that usual panel data models, such as the first difference or the fixed effects models, can be used to control for individual

[15] The CPS data used in this study also does not have data on immigrants who left prior to each CPS round. For example, the sample of immigrants who arrived in, say, 1970 is limited to those who remain in the United States until the late-1990s. The CPS, however, allows one to see who left between year 1 and year 2.

specific permanent components. The estimators based on multiple short panels are consistent, although they may be less efficient than estimators based on a longer panel. Second, the rotating panel that we use, the CPS MORG, is large, which makes it even more powerful than a usual panel with smaller sample sizes, such as the PSID or the NLSY79. The loss due to using possibly less efficient estimator is offset by the benefit from much larger sample size. Sample sizes matter in immigration studies because foreign-born persons, after all, are minorities. Third, the sample serves as a representative cross-section of the target population for any given time period. This property is the key in identifying sample attrition and outmigration processes.

2.4 Accounting for Sample Attrition when Emigrants are Unobserved

In the matched CPS, the sample attrition problem is particularly severe as the survey does not follow households who move.[16] Suppose that there is no international migration. 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.[17] More specifically, when a foreign-born respondent is missing in the second period, it is not possible to 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 (2009) 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

[16] Sample attrition and outmigration rates for 1994 to 2004 are 22-40% and about 3% per year, respectively, and panel attrition has a larger impact on estimation results than outmigration does.

[17] This person might be missing because of decease, emigration, or other reasons, but these possibilities (especially emigration) are relatively low and negligible.

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. A more formal discussion is presented in the Appendix.

3 Data Description

This section introduces the structure of the data set. Then it reports summary statistics and presents estimates of outmigration rates.

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 to 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.[18]

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.[19] If a household is included in either the first or the last 4 months of the interview periods, it is said that the household is in the rotation group. Figure 4 demonstrates the sample

[18] 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.

[19] 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 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.

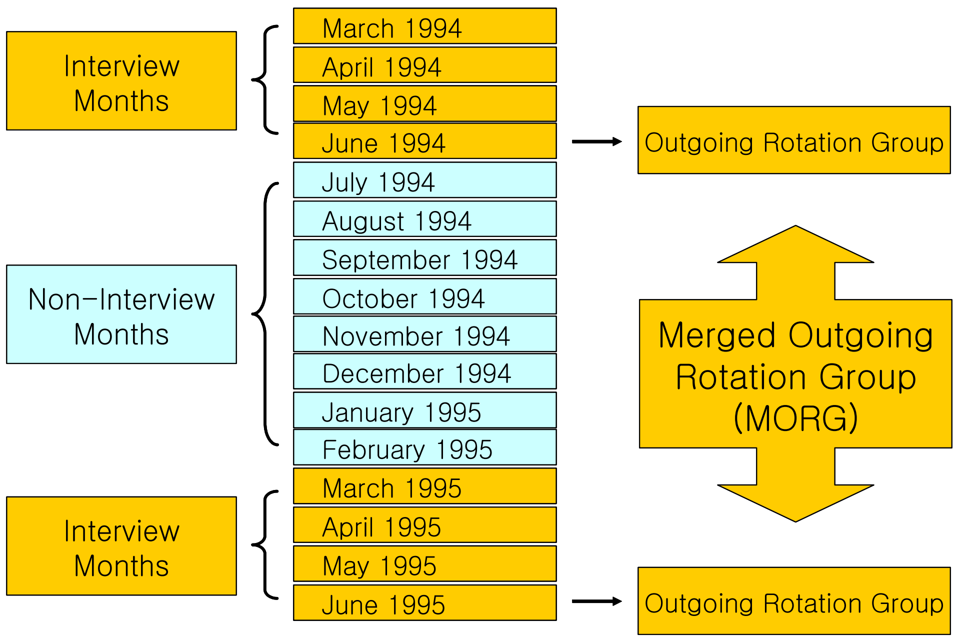

Figure 4: Sample Design of the CPS and its Merged Outgoing Rotation Group

design for a housing unit which, for instance, joins the survey on 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 4.) By construction, an 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.[20] 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 4 groups: immigrants from Central and South America, from Europe (including Australia, New Zealand, and Canada), from Asia, and from other countries.[21] The group of the 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 Central and South 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 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.[22] 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% 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

[20] 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 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.

[21] 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.

[22] 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.)

those in the entire sample, especially among foreign-born workers. For instance, the imputed wages for those from Central and South 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.[23] Consequently, our preferred way to handle the imputed wages is simply dropping them.[24]

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 Central and South 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.

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.[25] Between 1994 and 2004, the attrition rates are 28-40% among the immigrant samples and 22-32% among the native samples.[26] 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 Central and South 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.[27] We find substantial sample attrition bias.

[23] 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.

[24] 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.

[25] 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%.

[26] 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.

[27] 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).

Table 1. Summary Statistics

	Cross-Section Sample				Matched Sample			
	All		Reported Wage		All		Reported Wage	
	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)
C.S.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)
C.S.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)
Hours Worked per Week	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 (1 if married)	0.640	0.680	0.639	0.682	0.696	0.730	0.699	0.739
U.S. Citizen (1 if U.S. citizen)	1.000	0.385	1.000	0.387	1.000	0.440	1.000	0.434
C.S.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
N	872598	126240	578519	82630	254837	34018	167981	20718

Standard deviations are reported in parentheses. N: sample size

All: reported & imputed wages; Reported Wage: reported wages only

C.S.America: Central and South America; Europe: Europe, Australia, New Zealand, and Canada;

Asia: Asia; Others: Africa, Oceania, and other countries

The United States stopped collecting information on return migrants in 1957. To estimate outmigration rates, we exploit the structure of the CPS MORG. As housing units in the sample are kept fixed over the sampling period, the decrease in the sample size of immigrants will imply outmigration.[28] Using the panels prior to trimming individuals with extreme wages or negative experience, Table 2 provides the ratios of persons staying in the United States (one minus the outmigration rates) by year of entry.[29] For instance, the cell in the first row and first column indicates that in the 1st year of the 1994-1995 (unbalanced) panel, there were 5329 foreign-born persons in the United States. Then we count the number of foreign-born persons in the 2nd year of the 1994-1995 (unbalanced panel), which is 5331. We take the ratio between these numbers and get 1.00 (=5331/5329). This roughly means that little outmigration occurred during this period. Similarly in 1995-1996, the numbers of the foreign-born persons in the first and the second years are 5417 and 4605, respectively. It implies that about 15% (=1–4605/5417) of the foreign-born population in 1995 left the United States in 1996.[30]

Conceptually, it is impossible to have the stay rate exceed unity (or the outmigration rate lie below zero), but estimates above unity could arise from several reasons. On average fewer immigrants are expected to move in than move out due to outmigration, but the realized moving in immigrants may exceed moving out immigrants by chance provided the outmigration rate is low. This problem can be resolved by taking an overall average outmigration rate. Another reason for negative outmigration rates is misreporting.[31] For example, reentering foreign-born persons report their previous entry years or additional immigrants moving into existing households in year 2 report the same entry year as the host households. In the sample, values greater than unity are observed frequently, implying that sampling and measurement errors are relatively large.

Taking possible sampling errors into account, the last column reports the stay probability over the entire sample period. The last column of the first row reports that 25.2% (=1–0.768) of the foreign-born population who arrived in the United States in 1994 or before left the country by 2004.[32] On average, 2.6% (=1–0.974) of the foreign-born population outmigrates and the impact of outmigration is likely to be small. The stay probability by ethnic origin is available upon request.

[28] To be precise, it is not possible to separate outmigration from non-response, decease, or imprisonment in year 2.

[29] In the CPS, it is not possible to obtain outmigration rates for the most recent arrivals. For example, observations for the 1992-1993 arrivals not appear until the 1996-1997 period (and similarly for later arrivals). This is because there is inconsistency in coding the most recent arrivals in the CPS. It is explained in Appendix.

[30] The number of observations in year 2 beyond the first row are dropped to save space.

[31] See Lubotsky (2007) for further discussion on misreporting.

[32] This estimate is consistent with other empirical findings. For instance, Warren and Peck (1980) estimate that more than 1/6 of total immigrants admitted during the 1960s emigrated by the end of the decade.

Table 2. Stay Probability (One minus the Outmigration Rate) by Arrival Year

	1994 –1995	1995 –1996	1996 –1997	1997 –1998	1998 –1999	1999 –2000	2000 –2001	2001 –2002	2002 –2003	2003 –2004	[row total] row average
all foreign persons											
# in 2nd year	5331	4605	5011	5070	5398	5578	6299	6293	6831	6090	
# in 1st year	5329	5417	5121	5220	5527	5435	6060	6021	7001	6811	[0.768]
stay probability	1.000	0.850	0.979	0.971	0.977	1.026	1.039	1.045	0.976	0.894	0.974
before 1980 arrivals											
# in 1st year	2524	2417	2158	2078	2090	1936	1893	1793	1904	1745	[0.644]
stay probability	0.998	0.826	0.966	0.949	0.944	0.999	1.031	1.022	0.957	0.896	0.957
1980-1981 arrivals											
# in 1st year	517	615	511	520	467	474	458	534	457	483	[0.931]
stay probability	0.965	0.862	0.971	0.952	1.000	1.108	1.083	1.060	1.039	0.917	0.993
1982-1983 arrivals											
# in 1st year	323	343	282	317	329	294	321	313	349	338	[0.844]
stay probability	0.947	0.930	1.035	0.987	0.936	0.959	1.078	1.099	1.003	0.879	0.983
1984-1985 arrivals											
# in 1st year	456	521	411	451	401	389	429	395	444	447	[1.041]
stay probability	1.042	0.904	1.010	0.940	0.983	1.103	1.061	1.104	0.977	0.940	1.004
1986-1987 arrivals											
# in 1st year	400	433	421	405	353	357	375	353	426	409	[1.062]
stay probability	1.055	0.885	0.964	1.007	1.057	1.050	1.053	1.125	1.035	0.861	1.006
1988-1989 arrivals											
# in 1st year	567	545	473	529	528	596	502	497	498	527	[0.672]
stay probability	0.984	0.809	0.981	1.000	0.992	0.938	0.982	1.012	1.044	0.890	0.961
1990-1991 arrivals											
# in 1st year	542	543	491	437	478	476	536	587	588	542	[0.827]
stay probability	1.018	0.855	0.994	1.078	0.912	1.053	1.076	1.019	0.927	0.910	0.981
1992-1993 arrivals											
# in 1st year			374	483	424	442	458	450	481	477	[1.028]
stay probability			0.976	0.948	1.087	1.068	1.020	1.096	0.977	0.876	1.003
1994-1995 arrivals											
# in 1st year					457	471	542	572	520	488	[1.153]
stay probability					1.011	1.064	1.068	1.038	0.981	0.986	1.024
1996-1997 arrivals											
# in 1st year							546	527	566	575	[0.829]
stay probability							0.987	1.004	0.952	0.878	0.954
1998-1999 arrivals											
# in 1st year									768	780	[0.785]
stay probability									0.944	0.832	0.886

in 1st (2nd) year: the number of foreign-born persons in the 1st (2nd) year

stay probability: the (unconditional) ratio between the numbers of immigrants in the 2nd and in the 1st years

this value is less than or equal to unity in the population.

4 Empirical Evidence of Economic Assimilation

This section reports empirical findings. As our results are qualitatively different from the findings in the previous studies that use repeated cross-sections, we explore why they are different. Finally, this section presents assimilation estimates by ethnic origin.

4.1 Estimates of Economic Assimilation

We estimate the individual heterogeneity (IH) and the cohort heterogeneity (CH) models. The IH wage equations, (4) and (5), and the CH wage equations, (6) and (7), 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. For example, a quadratic IH or CH specification is given by

$$\begin{aligned} y_{it}^{imm} &= (\alpha_{nat}+\alpha)\,age_{it} + (\alpha_{2,nat}+\alpha_2)\,age_{it}^2 + \delta ysm_{it} + \delta_2 ysm_{it}^2 + (\beta_{nat}+\beta)\,edu_i + \cdots\,, \\ y_{it}^{nat} &= \alpha_{nat} age_{it} + \alpha_{2,nat} age_{it}^2 + \beta_{nat} edu_i + \cdots\,, \end{aligned}$$

and a cubic IH or CH specification is given by

$$\begin{aligned} y_{it}^{imm} &= (\alpha_{nat}+\alpha)\,age_{it} + (\alpha_{2,nat}+\alpha_2)\,age_{it}^2 + (\alpha_{3,nat}+\alpha_3)\,age_{it}^3 + \delta ysm_{it} + \delta_2 ysm_{it}^2 + \delta_3 ysm_{it}^3 + \cdots\,, \\ y_{it}^{nat} &= \alpha_{nat} age_{it} + \alpha_{2,nat} age_{it}^2 + \alpha_{3,nat} age_{it}^3 + \cdots\,. \end{aligned}$$

We estimate these equations by (a) accounting for sample attrition and outmigration and (b) accounting for sample attrition only as well as (c) without adjustment.[33] The coefficient estimates are reported in Tables A2-1, -2, -3, and -4 in the Appendix.

When age and years since migration enter as polynomials, it is difficult to read the implications of the coefficient estimates. Hence, Table 3 reports 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. Due to the polynomials of age and years since migration, the measure of economic assimilation,

[33] 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. In estimation of the matching functions, we use the matched longitudinal sample of individuals and cross-sections of all individuals including those not working, but we exclude extreme wage observations. Not-working individuals are included in this step in order to reflect market level changes, such as in the composition of natives, between consequtive years. In estimation of outmigration, we use the (unbalanced) panel of all individuals including extreme wage observations. In estimation of the outmigration process, labor market outcomes are not used as the variables must have known transition probabilities. To make sure that the foreign sample is large enough, we keep the largest available sample.

$EA\,(age, ysm)$, becomes a function of age and year since migration. For example, in a quadratic model, the measure of economic assimilation at $age = \overline{age}$ and $ysm = \overline{ysm}$ is $\alpha + 2\alpha_2\overline{age} + \delta + 2\delta_2\overline{ysm}$. Similarly, when a cubic specification is used, the measure of economic assimilation at $\overline{age}$ and $\overline{ysm}$ is $\alpha + 2\alpha_2\overline{age} + 3\alpha_3\overline{age}^2 + \delta + 2\delta_2\overline{ysm} + 3\delta_3\overline{ysm}^2$.

Table 3 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 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 % 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% points when sample attrition and outmigration are accounted for. This estimate is derived from the observation that immigrant wages grow annually by 2.38% and native wages by 2.13% under the assumption that year fixed effects on the level of wages are constant between two adjacent years. The difference is –0.25% points and is not statistically different from zero.[34].

The first three columns in the upper panel of Table 3 present the IH estimates of economic assimilation that accounts for sample attrition and outmigration. The sample attrition-outmigration-adjusted estimates from the quadratic specification suggest that wages of foreign-born workers grow slower than those of native-born workers by 1.17% 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% 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% points at age 24 and by 0.55% 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 sample attrition-outmigration-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 –0.0589. The solid line with circles in Figure 5 plots the foreign-native difference

[34] 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 $EA\,(age, ysm) < 0$.

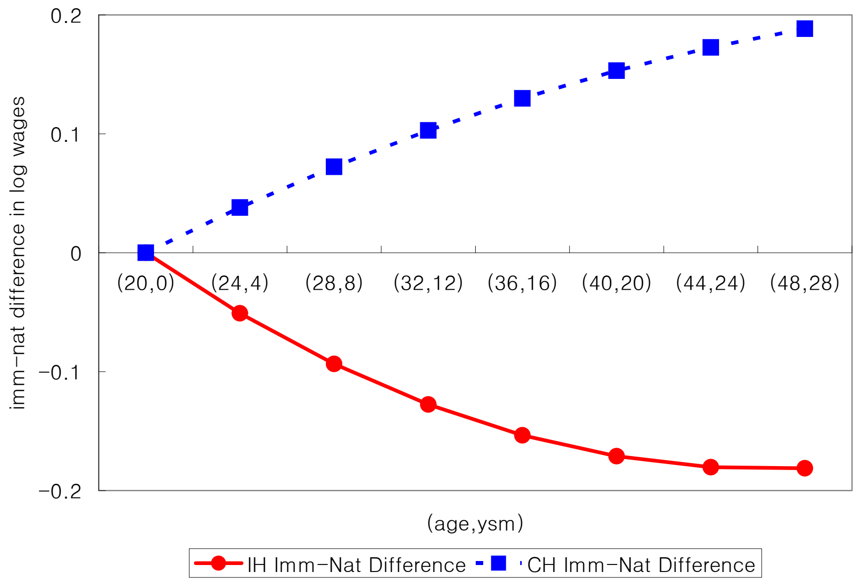

Figure 5: Immigrant-Native Difference in Simulated Log Wages (using the Quadratic Model Estimates)

in log wages. 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 11.7% points as reported in Table 3. According to Table 3, the slope estimate is negative and is statistically different from zero. Similarly, at age 32, the slope estimate is –0.0055 and is statistically different from zero. The foreign-native difference in wages stop widening above age 40. The slope or the measure of economic assimilation becomes close to zero.

The next three columns report sample attrition-adjusted estimates. These estimates are not very different from the sample attrition-outmigration-adjusted estimates. It suggests that the effect of outmigration is ineligible because the outmigration is not large between two adjacent years. The last three columns report unadjusted estimates. In general, the unadjusted estimates are greater than the sample attrition-adjusted ones, which implies that immigrants with slower wage growth are less likely to be observed in the second year panel than natives with slower wage growth. 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% points higher per year during the first 10 years and 0.38-0.50% points higher per year during the first 20 years based on CH models. To replicate the CH models using our sample, we drop the second period observations from the longitudinal samples and construct cross-section data.

The lower panel of Table 3 shows the economic assimilation estimates using the repeated cross-section approach using the same data. The results (misleadingly) suggest that there is significant economic assimilation, which is consistent with the previous repeated cross-section studies. In the first column applying the linear CH model, the estimate is 0.81 and is statistically different from zero at 1% confidence level. It implies that each additional year in the United States immigrant wages grow faster than native wages by 0.81% points when sample attrition and outmigration are accounted for. In the second column, the quadratic specification suggest that wages of a foreign-born worker grow at a faster rate than those of a native-born worker by 1.01% points per year at age 24. When they become 32, immigrant wages still grow at a faster rate than native wages by 0.70% points per year. At age 40, immigrant wages are growing 0.38% points faster than native wages. The cubic specification results in the third column suggest that wages of foreign-born workers grow slower than those of native-born workers by 1.30% points at age 24 and by 0.66% points at age 32. We find similar patterns in models that use reported wages only and models without attrition correcting weights.

The broken line in Figure 5 with squares uses the sample attrition-outmigration-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 0.0101 or 1.01% points per year. According to Table 3, the slope estimate is statistically different from zero. Similarly, at age 32, the slope estimate is 0.0070 and is statistically different from zero. In Figure 5 the simulated wage difference path is increasing even at age 48. In Table 3, the wages of 48 years old immigrants grow faster than that of 48 years old natives by 0.37% points per year, although this estimate is not statistically different from zero.

Table 3. Economic Assimilation Estimates in % points

		SATT-OUT-Adjusted			SATT-Adjusted			Not Adjusted		
		linear	quadra.	cubic	linear	quadra.	cubic	linear	quadra.	cubic
IH	Reported Wages									
	age=24, ysm=4	–0.25	–1.17**	–1.49**	–0.22	–1.13**	–1.46**	–0.18	–1.15**	–1.44**
		(0.31)	(0.55)	(0.68)	(0.31)	(0.55)	(0.68)	(0.30)	(0.54)	(0.68)
	age=32, ysm=12		–0.75**	–0.55		–0.72**	–0.52		–0.78**	–0.70*
			(0.35)	(0.39)		(0.36)	(0.38)		(0.35)	(0.38)
	age=40, ysm=20		–0.33	0.05		–0.31	0.08		–0.40	–0.16
			(0.32)	(0.47)		(0.32)	(0.47)		(0.32)	(0.47)
	age=48, ysm=28		0.08	0.33		0.10	0.35		–0.03	0.18
			(0.48)	(0.53)		(0.48)	(0.53)		(0.47)	(0.52)
IH	All Wages									
	age=24, ysm=4	–0.06	–1.33**	–1.32*	–0.03	–1.23**	–1.23*	0.20	–0.96	–0.91
		(0.34)	(0.59)	(0.73)	(0.34)	(0.59)	(0.73)	(0.34)	(0.59)	(0.73)
	age=32, ysm=12		–0.73*	–0.64		–0.68*	–0.58		–0.50	–0.49
			(0.38)	(0.43)		(0.38)	(0.43)		(0.39)	(0.43)
	age=40, ysm=20		–0.14	–0.07		–0.13	–0.05		–0.04	–0.10
			(0.36)	(0.53)		(0.36)	(0.53)		(0.35)	(0.54)
	age=48, ysm=28		0.45	0.39		0.42	0.37		0.42	0.27
			(0.53)	(0.58)		(0.53)	(0.58)		(0.54)	(0.58)
CH	Reported Wages									
	age=24, ysm=4	0.99***	0.93**	0.70	0.99***	0.91**	0.68	0.95***	1.05***	0.77
		(0.21)	(0.36)	(0.52)	(0.21)	(0.36)	(0.52)	(0.21)	(0.37)	(0.55)
	age=32, ysm=12		0.74***	0.69***		0.73***	0.69***		0.83***	0.76***
			(0.24)	(0.24)		(0.24)	(0.24)		(0.25)	(0.25)
	age=40, ysm=20		0.56***	0.64**		0.56***	0.65**		0.60***	0.69**
			(0.21)	(0.27)		(0.21)	(0.27)		(0.21)	(0.27)
	age=48, ysm=28		0.37	0.56*		0.38	0.57*		0.37	0.56
			(0.30)	(0.33)		(0.30)	(0.33)		(0.30)	(0.33)
CH	All Wages									
	age=24, ysm=4	0.81***	1.01***	1.30**	0.82***	1.02***	1.30**	0.94***	1.32***	1.71***
		(0.15)	(0.25)	(0.37)	(0.15)	(0.25)	(0.38)	(0.15)	(0.27)	(0.39)
	age=32, ysm=12		0.70***	0.66***		0.70***	0.67***		0.97***	0.96***
			(0.17)	(0.17)		(0.17)	(0.17)		(0.17)	(0.17)
	age=40, ysm=20		0.38***	0.23		0.39***	0.24		0.63***	0.45**
			(0.14)	(0.19)		(0.15)	(0.19)		(0.15)	(0.19)
	age=48, ysm=28		0.07	0.02		0.07	0.02		0.29	0.18
			(0.20)	(0.22)		(0.20)	(0.22)		(0.21)	(0.22)

IH: Individual Heterogeneity Model Estimates; CH: Cohort Heterogeneity Model Estimates

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

SATT-Adjusted: Sample Attrition-Adjusted; SATT-OUT-Adjusted: Sample Attrition-Outmigration-Adjusted

Standard errors for adjusted estimates do not account for sampling error in weighting functions estimation.

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

4.2 Discussion

Why are the estimates from the IH and the CH models different? It is because the cohort fixed effects fail to fully reflect individual heterogeneity. The empirical findings in Table 3 indicate a positive correlation between μ_i and age at migration conditional on age, year of entry, and other observables. It implies that among the immigrants of the same year of entry, older persons have higher μ_i than younger ones.[35]

There are possibly two explanations for the positive correlation between μ_i and age at migration. The first one is based on a search model and the Roy model. 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 human capital investment is more difficult in the U.S. 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.

There is a study by Bleakley and Chin (2004) which shows that among the immigrants who immigrated to the U.S. as children, those who entered as younger children assimilate faster than those who entered as older children due to language proficiency. To see whether our findings conflict with theirs, we estimate economic assimilation after dropping immigrants who entered as children and find more negative assimilation estimates. (See the last three columns of Table A1-2 in the Appendix.) This implies that our findings are consistent with Bleakley and Chin and that the positive correlation between μ_i and age at migration is dominated by adult immigrants.

Our findings, however, do not imply that previous results based on repeated cross-sections are incorrect. Our results are based on a much more recent sample and whether there is a correlation between μ_i and other regressors is an empirical question. While it is not possible to test the positive correlation for earlier periods using our CPS data, other papers applying IH type models for earlier periods report existence of economic assimilation. For example, Lubotsky (2007) uses time series linked to cross-section data 1951-1997. He finds that the earnings of immigrants have grown 0.50-0.65% points per year during the first twenty years relative to the earnings of native-born workers of similar characteristics. Also, Duleep and Regets (1997a), matching the June 1987 and June 1988 CPS, provide a similar table for 1987-1988. They find that the wage growth of foreign-born workers exceeds that of native-born workers

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

by 0.3% points.

To summarize, assimilation estimates based on CH models for recent years are biased upward due to fixed unobserved heterogeneity within an immigration year cell even after country of origin is controlled for, but may not be biased for earlier years. The results in this paper suggest that the trend in assimilation rates may have changed in the last decade and that using cross-section data to estimate assimilation for this period is misleading. It is an empirical question whether there is bias due to unobserved heterogeneity and the answer to the question will vary by time and location. In consequence, 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.

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 5 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 Central and South America. From the sample attrition-outmigration-adjusted estimates of nonlinear specifications, we learn that their wages grow slower than native wages by 1.41-2.23% points at age 24 and by 0.39-0.76% 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, reports 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 5. Economic Assimilation Estimates in % (by Origin): Reported Wages Only

Individual Hetero.	SATT-OUT-Adjusted			SATT-Adjusted			Not Adjusted		
	linear	quadra.	cubic	linear	quadra.	cubic	linear	quadra.	cubic
C.S.America									
age=24, ysm=4	0.10	–1.41**	–2.23***	0.11	–1.40**	–2.22***	0.12	–1.33**	–2.36**
	(0.37)	(0.64)	(0.78)	(0.37)	(0.64)	(0.78)	(0.37)	(0.63)	(0.77)
age=32, ysm=12		–0.76*	–0.39		–0.74*	–0.38		–0.82**	–0.57
		(0.41)	(0.47)		(0.41)	(0.48)		(0.41)	(0.46)
age=40, ysm=20		–0.11	0.66		–0.09	0.68		–0.31	0.44
		(0.41)	(0.59)		(0.41)	(0.59)		(0.41)	(0.58)
Europe									
age=24, ysm=4	–1.18	–0.96	1.80	–1.17	–0.88	1.86	–1.09	–1.16	2.54
	(0.86)	(1.74)	(2.49)	(0.86)	(1.74)	(2.49)	(0.84)	(1.77)	(2.63)
age=32, ysm=12		–0.85	–1.21		–0.79	–1.14		–0.95	–1.00
		(1.20)	(1.19)		(1.20)	(1.20)		(1.23)	(1.23)
age=40, ysm=20		–0.73	–2.64**		–0.71	–2.59**		–0.74	–2.68**
		(0.86)	(1.29)		(0.85)	(1.29)		(0.87)	(1.23)
Asia									
age=24, ysm=4	–0.51	–0.84	–0.27	–0.48	–0.76	–0.17	–0.36	–1.12	–0.51
	(0.64)	(1.37)	(1.84)	(0.64)	(1.37)	(1.83)	(0.62)	(1.30)	(1.72)
age=32, ysm=12		–0.52	–0.38		–0.46	–0.32		–0.60	–0.47
		(0.82)	(0.87)		(0.82)	(0.87)		(0.79)	(0.85)
age=40, ysm=20		–0.19	–0.29		–0.17	–0.27		–0.08	–0.19
		(0.76)	(1.05)		(0.76)	(1.05)		(0.75)	(1.04)
Others									
age=24, ysm=4	–0.66	–0.58	–3.18	–0.50	–1.30	–2.92	–0.04	–0.11	–1.73
	(1.75)	(3.26)	(4.05)	(1.72)	(3.15)	(3.90)	(1.61)	(2.92)	(3.70)
age=32, ysm=12		–0.03	1.35		0.25	1.62		0.34	0.98
		(2.03)	(2.28)		(1.96)	(2.22)		(1.83)	(2.00)
age=40, ysm=20		0.52	2.93		0.80	3.21		0.79	2.17
		(1.94)	(2.68)		(1.88)	(2.65)		(1.88)	(2.55)

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

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

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

A caveat is that the findings of economic assimilation might be sensitive to the choice of the definition of economic assimilation. Especially for immigrants from Central and South America, perhaps we would obtain different results if we compare them to the natives within their ethnic groups using the CPS MORG. Immigrants from Europe and Asia are of little interest because they perform better than white natives. This is an important point as there exists difference in economic performance among natives by country of ancestry and race, but it is difficult to precisely measure assimilation. Borjas (1995) finds that Mexicans fail to catch up to natives of similar ancestry. However, he does not make a strong argument out of it because the composition of non-white natives has changed over time. In addition to this, we address the importance of measurement errors in the survey of ethnic origin among natives. If there is a systematic pattern of misreporting ancestry, the measurement errors pose another difficulty in the approach of comparison with natives by country of ancestry and race.

5 Concluding Remarks

This study reexamines the evidence of wage convergence of immigrants using a novel research design. This paper provides a longitudinal analysis of assimilation by exploiting the two-year panel aspect of the CPS to account for individual fixed effects. The resulting overlapping rotating panel data set enables one to account for the problems of sample attrition as well as outmigration. By comparing cross-section and panel estimates of economic assimilation for 1994-2004, we find that repeated cross-section estimates are biased upward by fixed unobserved heterogeneity and controlling for the heterogeneity reverses the conventional result of economic assimilation.

The empirical findings suggest little evidence of economic assimilation. The growth rate of hourly wages of immigrants from Central and South America at age 24 is 1.41-2.23% slower than that of native-born workers. At age 32 the gap in growth rates is between 0.39-0.76% points. New immigrants from Central and South America earn lower wages than natives, and this gap widens with time 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. These results are qualitatively different from the findings in the previous studies that use repeated cross-sections.

Our results of no assimilation do not imply that previous results based on repeated cross-sections are incorrect. It is an empirical question whether individual heterogeneity is correlated with other covariates after controlling for country of birth and year of entry. Our study, however, suggests that one should be aware of the fact that the repeated cross-section approach may yield biased estimates when the study focuses on different time periods in the United States or on different countries other than the United States.

6 References

Ai, Chunrong and Xiaohong Chen (2003): "Efficient Estimation of Models with Conditional Moment Restrictions containing Unknown Functions," *Econometrica*, 71(6), 1795-1843.

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 2A, 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 (2009): "Sample Attrition in the presence of Population Attrition," Working Paper, University of Washington.

Kitamura, Yuichi, Gautam Tripathi, and Hyungtaik Ahn (2004): "Empirical Likelihood-Based Inference in Conditional Moment Restriction Models," *Econometrica*, 72(6), 1667-1714.

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 3B, 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. (2001): "The Effect of Changes in the U.S. Wage Structure on Recent Immigrants' Earnings," Working paper.

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," Working Paper.

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

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 are 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.[36] Workers with extreme wages (less 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 let 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

[36] 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 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.

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 panel: codes 1-12 (1990-1991)

1995-1996 panel: codes 1-12 (1990-1991)

1996-1997 panel: codes 1-13 (1992-1993)

1997-1998 panel: codes 1-13 (1992-1993)

1998-1999 panel: codes 1-14 (1994-1995)

1999-2000 panel: codes 1-14 (1994-1995)

2000-2001 panel: codes 1-15 (1996-1997)

2001-2002 panel: codes 1-15 (1996-1997)

2002-2003 panel: codes 1-16 (1998-1999)

2003-2004 panel: 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

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

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.[37]

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

$$\Pr\left(D_S = 1 | U_1 = u_1, U_2 = u_2, V = v\right) = g\left(v'\phi_0 + u_1'\phi_1 + u_2'\phi_2\right), \tag{9}$$

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

$$C\left(u_1, u_2, v\right) = \frac{\Pr\left(D_S = 1\right)}{g\left(v'\phi_0 + u_1'\phi_1 + u_2'\phi_2\right)}. \tag{10}$$

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

In the presence of population attrition, one additional step is required prior to the above procedure. The population attrition function, $\Pr\left(D_P = 1 | u_2, v\right)$, can be nonparametrically identified when population attrition is solely determined by variables of known transition probability.[38] Suppose that the transition probability is given by $P\left(Z_2 = z_2 | Z_1 = z_1\right)$, where z is a vector of variables of known transition probabil-

[37] 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, $D_P \cdot D_S = 1$.

[38] This assumption is strong but necessary because we do not know who emigrated from the United States.

ity.[39] 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

$$\begin{aligned} \Pr(D_P = 1|u_2, v) &= \Pr(D_P = 1|z_2) \\ &\equiv k(z_2'\psi), \end{aligned} \tag{11}$$

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.[40] Intuitively, weight the individuals in the population (or more precisely the cross-section) with the inverse of one minus the probability of population attrition, $1/k(z_2'\psi)$.

The weights in (10) can be estimated by the conditional moment restrictions given by

$$\begin{aligned} 1 &= E\left[\frac{D_S}{g(v'\phi_0 + u_1'\phi_1 + u_2'\phi_2)}|u_1, v\right] \quad \text{w.p.1}, \\ \frac{1}{k(z_2'\psi)} &= E\left[\frac{D_S}{g(v'\phi_0 + u_1'\phi_1 + u_2'\phi_2)}|u_2, v, D_P = 1\right] \quad \text{w.p.1}. \end{aligned} \tag{12}$$

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 (12) and obtain (10). Finally, use (10) 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.

7.3 Analyzing the Difference between IH and CH Estimates

In order to see why the IH and the CH estimates are different, it is useful to discuss how IH and CH models control for heterogeneity. Assume that the true data generating process is given by

$$\begin{aligned} y_{it} &= \alpha age_{it} + \delta ysm_{it} + \beta edu_i + \mu_i + \varepsilon_{it} \\ &= \alpha age_{it} + \delta(t - c) + \beta edu_i + \mu_i + \varepsilon_{it}, \end{aligned} \tag{13}$$

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

[40] 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.

for an individual i in an arrival year c. We drop year fixed effects as they are identified from the native equation. For the time being, assume that there is no sample attrition nor outmigration. Further, assume that all immigrants are from a common source country.

First of all, single cross-section analyses fail to identify $\alpha + \delta$ when the skill composition of new immigrants change over time. Assume that the individual fixed effects, μ_i, in (13) can be replaced with the arrival year cohort fixed effects, μ_c. Then we have

$$\begin{aligned} E\left[y_{it}|c_1, t, age_{it}, edu_i\right] &= \alpha age_{it} + \delta\left(t - c_1\right) + \beta edu_i + \mu_{c1}, \\ E\left[y_{jt}|c_2, t, age_{jt}, edu_j\right] &= \alpha age_{jt} + \delta\left(t - c_2\right) + \beta edu_j + \mu_{c2}. \end{aligned}$$

In this case $\alpha + \delta$ is not identified unless $\mu_{c1} = \mu_{c2}$, which is exactly the same argument made by Borjas (1985).

When repeated cross-sections are available, it is possible to identify assimilation, but simply including entry year fixed effects is not enough. One needs to employ a pseudo-panel approach. Since the pseudo-panel approach requires grouping of individuals, the sample size becomes much smaller: the number of arrival year cohort times the number of cross-sections. Suppose that we have 2 periods, and i and j are in the same arrival year cohort but in different cross-sections. From (13), we have

$$\begin{aligned} E\left[y_{it}|c, t, age_{it}, edu_i\right] &= \alpha age_{it} + \delta\left(t - c\right) + \beta edu_i + E\left[\mu_i|c, t, age_{it}, edu_i\right], \\ E\left[y_{jt'}|c, t', age_{jt'}, edu_j\right] &= \alpha age_{jt'} + \delta\left(t' - c\right) + \beta edu_j + E\left[\mu_j|c, t', age_{jt'}, edu_j\right], \end{aligned}$$

where $t' = t + 1$. Now $\alpha + \delta$ is identified under, for instance, the cohort heterogeneity assumption

$$E\left[\mu_i|c, t, age_{it}, edu_i\right] = \mu_c \quad \text{w.p.1 for all } i \in c \text{ and for } t, t'.$$

This identification restriction is employed in almost all the repeated cross-section studies. The constraint, however, is not likely to hold if age at migration, $age_{it} - (t - c)$, is correlated with μ_i conditional on the year of entry and other observable variables. For instance, suppose the correlation between μ_i and age at migration is given by

$$E\left[\mu_i|c, t, age_{it}, edu_i\right] = \mu_c + \eta_a\left(age_{it} - (t - c)\right). \tag{14}$$

Under (13), assume that we follow a group of persons of the same age at migration:

$$age_{it} - (t - c) = age_{jt'} - \left(t' - c\right).$$

Then, we have

$$\begin{aligned} E\left[y_{it}|c, t, age_{it}, edu_i\right] &= \alpha age_{it} + \delta\left(t - c\right) + \beta edu_i + \left[\mu_c + \eta_a\left(age_{it} - (t - c)\right)\right], \\ E\left[y_{jt'}|c, t', age_{jt'}, edu_j\right] &= \alpha age_{jt'} + \delta\left(t' - c\right) + \beta edu_j + \left[\mu_c + \eta_a\left(age_{jt'} - \left(t' - c\right)\right)\right]. \end{aligned}$$

Therefore, $\alpha + \delta$ is identified.

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

Individual Hetero.	SATT-OUT-Adjusted			SATT-Adjusted			Not Adjusted		
	linear	quadra.	cubic	linear	quadra.	cubic	linear	quadra.	cubic
C.S.America									
age=24, ysm=4	–0.01	–1.52**	–2.09**	0.01	–1.46**	–2.03**	0.17	–1.25*	–1.98**
	(0.41)	(0.70)	(0.85)	(0.41)	(0.70)	(0.85)	(0.41)	(0.71)	(0.86)
age=32, ysm=12		–0.91**	–0.66		–0.88*	–0.63		–0.78	–0.63
		(0.45)	(0.52)		(0.45)	(0.52)		(0.45)	(0.52)
age=40, ysm=20		–0.30	0.24		–0.30	0.24		–0.31	0.19
		(0.46)	(0.65)		(0.46)	(0.65)		(0.46)	(0.65)
Europe									
age=24, ysm=4	–1.69	–3.17*	–1.94	–1.68	–3.09*	–1.92	–1.39	–3.14*	–2.21
	(0.90)	(1.76)	(2.30)	(0.90)	(1.77)	(2.31)	(0.90)	(1.79)	(2.38)
age=32, ysm=12		–2.24*	–2.29*		–2.20*	–2.23*		–2.19*	–2.09*
		(1.22)	(1.26)		(1.23)	(1.27)		(1.24)	(1.26)
age=40, ysm=20		–1.32	–2.13		–1.30	–2.06		–1.25	–1.68
		(0.90)	(1.47)		(0.91)	(1.48)		(0.91)	(1.47)
Asia									
age=24, ysm=4	0.55	–0.01	0.96	0.57	0.15	1.15	0.91	0.79	2.47*
	(0.69)	(1.40)	(1.81)	(0.69)	(1.40)	(1.82)	(0.69)	(1.38)	(1.74)
age=32, ysm=12		0.36	0.07		0.43	0.13		0.86	0.41
		(0.85)	(0.92)		(0.86)	(0.92)		(0.84)	(0.91)
age=40, ysm=20		0.72	0.08		0.71	0.04		0.92	–0.23
		(0.80)	(1.14)		(0.80)	(1.14)		(0.80)	(1.14)
Others									
age=24, ysm=4	0.93	–1.15	–0.42	1.17	–0.73	–0.09	1.80	–0.29	0.87
	(1.58)	(2.84)	(3.44)	(1.58)	(2.84)	(3.43)	(1.54)	(2.83)	(3.54)
age=32, ysm=12		–0.02	–0.43		0.34	0.01		0.74	0.26
		(1.83)	(2.01)		(1.83)	(2.02)		(1.81)	(1.95)
age=40, ysm=20		1.11	0.27		1.42	0.71		1.77	0.59
		(1.70)	(2.51)		(1.70)	(2.51)		(1.69)	(2.46)

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

Sample sizes: Native (156241), C.S.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 Hetero.	SATT-OUT-Adjusted			Not Adjusted			A-O-Adjusted, enter ≥18		
	linear	quadra.	cubic	linear	quadra.	cubic	linear	quadra.	cubic
C.S.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), C.S.America (6438), Europe (1689), Asia (2657), Others (492)

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

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

Table A2 -1. Wage Equation (in First Differenced) Estimates: Reported & Imputed Wages

IH	SATT-OUT-Adjusted		SATT-Adjusted		Not Adjusted	
	quadratic	cubic	quadratic	cubic	quadratic	cubic
Constant	0.101***	0.187***	0.101***	0.187***	0.091***	0.179***
	(0.006)	(0.015)	(0.006)	(0.015)	(0.007)	(0.015)
$\frac{1}{10}$Age	–0.019***	–0.066***	–0.019***	–0.066***	–0.019***	–0.065***
	(0.001)	(0.007)	(0.001)	(0.007)	(0.001)	(0.007)
$\frac{1}{100}$Age^2		0.006***		0.006***		0.006***
		(0.001)		(0.001)		(0.001)
Imm.	–0.032***	–0.065	–0.030**	–0.062	–0.026**	–0.058
	(0.012)	(0.040)	(0.012)	(0.040)	(0.013)	(0.041)
$\frac{1}{10}$Age$_i$	0.008**	0.029	0.007**	0.028	0.007*	0.029
	(0.004)	(0.022)	(0.004)	(0.022)	(0.004)	(0.022)
$\frac{1}{100}$Age$_i^2$		–0.003		–0.003		–0.003
		(0.003)		(0.003)		(0.003)
$\frac{1}{10}$Ysm	–0.000	–0.008	–0.000	–0.008	–0.001	–0.012
	(0.004)	(0.010)	(0.004)	(0.011)	(0.004)	(0.011)
$\frac{1}{100}$Ysm^2		0.002		0.002		0.002
		(0.002)		(0.002)		(0.002)
Year	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors are reported in parentheses. N: sample size = 177695

linear specification results are available upon request

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

Ysm: years since migration; Year: calendar year fixed effects

Table A2 -2. Wage Equation (in First Differenced) Estimates: Reported Wages Only

IH-R	SATT-OUT-Adjusted		SATT-Adjusted		Not Adjusted	
	quadratic	cubic	quadratic	cubic	quadratic	cubic
Constant	0.096^{***}	0.178^{***}	0.096^{***}	0.178^{***}	0.108^{***}	0.194^{***}
	(0.006)	(0.014)	(0.006)	(0.014)	(0.006)	(0.014)
$\frac{1}{10}$Age	-0.019^{***}	-0.064^{***}	-0.019^{***}	-0.064^{***}	-0.018^{***}	-0.064^{***}
	(0.001)	(0.007)	(0.001)	(0.007)	(0.001)	(0.007)
$\frac{1}{100}\text{Age}^2$		0.006^{***}		0.006^{***}		0.006^{***}
		(0.001)		(0.001)		(0.001)
Imm.	-0.019^{*}	–0.036	–0.018	–0.034	-0.021^{*}	–0.028
	(0.012)	(0.039)	(0.012)	(0.039)	(0.012)	(0.038)
$\frac{1}{10}\text{Age}_i$	0.003	0.008	0.002	0.008	0.003	0.005
	(0.003)	(0.021)	(0.003)	(0.021)	(0.003)	(0.020)
$\frac{1}{100}\text{Age}_i^2$		–0.001		–0.001		–0.000
		(0.003)		(0.003)		(0.002)
$\frac{1}{10}$Ysm	0.002	0.010	0.003	0.011	0.001	0.008
	(0.003)	(0.009)	(0.003)	(0.009)	(0.003)	(0.009)
$\frac{1}{100}\text{Ysm}^2$		–0.002		–0.002		–0.001
		(0.002)		(0.002)		(0.002)
Year	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors are reported in parentheses. N: sample size = 100393

linear specification results are available upon request

Imm.: indicator variable of a foreign-born person; Age_i: age of foreign-born persons

Ysm: years since migration; Year: calendar year fixed effects

Table A2 -3. Wage Equation (in Level) Estimates: Reported & Imputed Wages

CH	SATT-OUT-Adjusted		SATT-Adjusted		Not Adjusted	
	quadratic	cubic	quadratic	cubic	quadratic	cubic
Constant	–0.441***	–1.092***	–0.441***	–1.092***	–0.380***	–1.143***
	(0.016)	(0.043)	(0.016)	(0.043)	(0.016)	(0.046)
Age	0.078***	0.134***	0.078***	0.134***	0.076***	0.139***
	(0.001)	(0.004)	(0.001)	(0.004)	(0.001)	(0.004)
$\frac{1}{100}\text{Age}^2$	–0.081***	–0.228***	–0.081***	–0.228***	–0.079***	–0.242***
	(0.001)	(0.009)	(0.001)	(0.009)	(0.001)	(0.010)
$\frac{1}{1000}\text{Age}^3$		0.012***		0.012***		0.013***
		(0.001)		(0.001)		(0.001)
Imm.	0.566***	0.313**	0.562***	0.320**	0.516***	0.281**
	(0.046)	(0.129)	(0.046)	(0.129)	(0.048)	(0.135)
Age_i	–0.019***	0.001	–0.019***	–0.000	–0.019***	–0.001
	(0.002)	(0.010)	(0.002)	(0.010)	(0.002)	(0.011)
$\frac{1}{100}\text{Age}_i^2$	0.015***	–0.037	0.015***	–0.035	0.015***	–0.030
	(0.003)	(0.027)	(0.003)	(0.027)	(0.003)	(0.028)
$\frac{1}{1000}\text{Age}_i^3$		0.004*		0.004*		0.004
		(0.002)		(0.002)		(0.002)
Ysm	0.025***	0.026***	0.025***	0.026***	0.027***	0.030***
	(0.003)	(0.005)	(0.003)	(0.005)	(0.003)	(0.006)
$\frac{1}{100}\text{Ysm}^2$	–0.035***	–0.043*	–0.035***	–0.043*	–0.037***	–0.053**
	(0.007)	(0.023)	(0.007)	(0.026)	(0.007)	(0.027)
$\frac{1}{1000}\text{Ysm}^3$		0.001		0.001		0.002
		(0.004)		(0.004)		(0.004)
Educ	0.092***	0.092***	0.092***	0.092***	0.095***	0.094***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Educ_i	–0.037***	–0.036***	–0.037***	–0.036***	–0.036	–0.036***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)

Standard errors are reported in parentheses. N: sample size = 177695

Imm.: indicator variable of a foreign-born person; Age_i: age × Imm.

Ysm: years since migration; Educ: years of schooling; Educ_i: years of schooling × Imm.

Fixed Effects: birth country, arrival year, calendar year

Table A2 -4. Wage Equation (in Level) Estimates: Reported Wages Only

CH-R	SATT-OUT-Adjusted		SATT-Adjusted		Not Adjusted	
	quadratic	cubic	quadratic	cubic	quadratic	cubic
Constant	–0.493***	–1.087***	–0.493***	–1.088***	–0.468***	–1.134***
	(0.020)	(0.055)	(0.020)	(0.055)	(0.021)	(0.059)
Age	0.081***	0.132***	0.081***	0.132***	0.078***	0.134***
	(0.001)	(0.005)	(0.001)	(0.005)	(0.001)	(0.005)
$\frac{1}{100}$Age^2	–0.085***	–0.219***	–0.085***	–0.219***	–0.083***	–0.225***
	(0.001)	(0.012)	(0.001)	(0.012)	(0.001)	(0.013)
$\frac{1}{1000}$Age^3		0.011***		0.011***		0.012***
		(0.001)		(0.001)		(0.001)
Imm.	0.589***	0.281	0.590***	0.288*	0.580***	0.337*
	(0.063)	(0.172)	(0.063)	(0.172)	(0.065)	(0.182)
Age$_i$	–0.019***	0.010	–0.019***	0.009	–0.020***	0.003
	(0.003)	(0.014)	(0.003)	(0.014)	(0.003)	(0.015)
$\frac{1}{100}$Age$_i^2$	0.013***	–0.064*	0.013***	–0.063*	0.014***	–0.047
	(0.004)	(0.037)	(0.004)	(0.037)	(0.004)	(0.039)
$\frac{1}{1000}$Age$_i^3$		0.006**		0.006**		0.005
		(0.003)		(0.003)		(0.003)
Ysm	0.024***	0.015**	0.024***	0.015**	0.026***	0.017**
	(0.004)	(0.007)	(0.004)	(0.007)	(0.004)	(0.008)
$\frac{1}{100}$Ysm^2	–0.024**	0.027	–0.024**	0.028	–0.028***	0.020
	(0.010)	(0.035)	(0.010)	(0.035)	(0.011)	(0.037)
$\frac{1}{1000}$Ysm^3		–0.007		–0.007		–0.006
		(0.005)		(0.005)		(0.005)
Educ	0.094***	0.093***	0.094***	0.093***	0.096***	0.096***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Educ$_i$	–0.038***	–0.038***	–0.038***	–0.038***	–0.037***	–0.037***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)

Standard errors are reported in parentheses. N: sample size = 100393

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

Ysm: years since migration; Educ: years of schooling; Educ$_i$: years of schooling × Imm.

Fixed Effects: birth country, arrival year, calendar year